Table 1. Nomenclature for Antiviral Resistance

Term	Definition
Primary treatment failure (nonresponse)	Inability of nucleoside/tide analogue treatment to reduce serum HBV DNA by ≥ 1 $\log_{10}$ IU/ml after the first 6 months of treatment
Secondary treatment failure (virologic breakthrough)	Increase in serum HBV DNA by ≥ 1 $\log_{10}$ above nadir on ≥ 2 occasions 1 month apart, while on treatment, after achieving initial response in a medication compliant patient
Biochemical breakthrough	Elevation in serum alanine aminotransferase (ALT) while on treatment, after achieving normalization in a medication compliant patient
Genotypic resistance	Detection of viral populations bearing amino acid substitutions in the reverse transcriptase region of the HBV genome that have been shown to confer resistance to antiviral drugs in phenotypic assay, during antiviral therapy. These mutations are usually detected in patients with virologic breakthrough but they can also be present in patients with persistent viremia and no virologic breakthrough.
Phenotypic resistance	Decreased susceptibility of an HBV polymerase to an antiviral treatment *in vitro*
Cross resistance	Decreased susceptibility to more than one antiviral drug conferred by the same amino acid substitution or combination of amino acid substitutions

define genotypic, phenotypic, and clinical resistance to NA therapy. In this review, we propose definitions of terminologies, briefly describe available methods for detecting and quantifying drug resistance, and discuss the interpretation of drug resistance data and its current and future application in clinical practice.

(a) Clinical Classification of Antiviral Resistance

At the 2006 National Institutes of Health Workshop on HBV, standardized definitions of response to antiviral therapy were proposed (Table 1).[1] More detailed clarification of terminologies that are commonly used in describing antiviral resistance is provided here. A sensitive and specific HBV DNA assay with a wide dynamic range of quantification, calibrated to express results in WHO international units per ml (IU/ml), should be used to quantify serum HBV DNA levels prior to treatment, to assess response, and to detect virologic breakthroughs.

(i) Primary Antiviral Treatment Failure (or Nonresponse). Primary nonresponse is defined as the inability of NAs to reduce serum HBV DNA by ≥ 1 $\log_{10}$ IU/ml after the first 6 months of treatment. This definition was chosen as it exceeds variability in HBV DNA assays and reflects a true virologic response but a decrease in HBV DNA level by 1 $\log_{10}$ IU/ml is not a meaningful clinical response. Primary nonresponse may be due to factors related to the host, virus, or drug. Polymorphisms in enzymes involved in converting prodrugs to the active compounds or in phosphorylating NAs to their triphosphates (active moiety) may contribute to primary nonresponse.[2,3] Certain viral strains may be less susceptible to one or more antiviral therapies as suggested by a recent case report on primary nonresponse to adefovir.[4] Potency and dose of the antiviral therapy may also be important. For example, the approved dose of adefovir (10 mg daily) is not as potent as higher doses and may be the most important factor in the high rate of primary nonresponse to adefovir.[5]

Monitoring for primary nonresponse is important because a high residual viral level after the first 6-12 months of therapy has been demonstrated to be associated with increased risk of antiviral resistance.[6-8]

(ii) Secondary Antiviral Treatment Failure (or Virologic Breakthrough). Virologic breakthrough, which is usually associated with drug resistance, is defined as a ≥ 1 $\log_{10}$ IU/ml increase in serum HBV DNA level from nadir in two consecutive samples 1 month apart in patients who have responded and have been compliant with antiviral medication(s).[9,10] Confirmation of the increase in serum HBV DNA in a second sample is not necessary in patients with accompanying flare in aminotransferase level. Serum HBV DNA levels tend to be low initially because most antiviral-resistant HBV mutants have decreased replication fitness compared to wild-type HBV.[11,12] However, compensatory mutations that can restore replication fitness frequently accumulate during continued treatment leading to *viral rebound* — progressive increase in serum HBV DNA level that may exceed pretreatment value.[10]

(iii) Biochemical Breakthrough. Biochemical breakthrough is defined as elevation in serum aminotransferase level during treatment in a patient who had achieved initial normalization. Serum aminotransferases may remain normal for a few weeks or a few years after virologic breakthrough. Biochemical breakthrough often coincides with a viral rebound, in some cases a marked increase in aminotransferases occur resulting in a *hepatitis flare* (aminotransferase >5 times the upper limit of normal) and rarely hepatic decompensation.[13,14]

2. Definition of Genotypic Antiviral Resistance

A fundamental issue in antiviral resistance is the criteria for defining drug-resistant mutations. When a muta-

tion occurs during replication, it results in a nucleotide substitution. The substitution can be synonymous (not associated with an amino acid change) or non-synonymous (associated with an amino acid change). A resistance mutation induces an amino acid change that decreases the sensitivity to an antiviral drug.

Genotypic antiviral resistance designates the presence of unique nucleotide and corresponding deduced amino acid mutations in the drug target gene (the HBV polymerase gene in the case of HBV treatment with NAs) that have been previously demonstrated to be associated with antiviral resistance. Ideally, to identify potential genotypic resistance, the nucleotide and deduced amino acid sequence of HBV isolated from the patient during virologic breakthrough should be compared to the sequence of HBV isolated from a pretreatment sample from the same patient. When pretreatment samples are not available for analysis, sequence data at the time of virologic breakthrough should be compared to consensus published sequence(s) of the same HBV genotype.

Primary drug resistant mutations cause an amino acid substitution that result in reduced susceptibility to an antiviral agent while **secondary compensatory mutations** cause amino acid substitutions that restore functional defects in viral polymerase activity (i.e., replication fitness) associated with primary drug resistance.

For example, primary lamivudine resistance associated changes occur at codon 204 and result in amino acid changes within the tyrosine-methionine-aspartate-aspartate (YMDD) motif - rtM204V/I (methionine to valine or isoleucine substitution). These changes cause a greater than 100-fold decrease in susceptibility to lamivudine in phenotypic assays. The most common compensatory mutation associated with lamivudine resistance, rtL180M (leucine to methionine substitution) restores replication fitness of HBV polymerase that harbors the rtM204V/I mutation.[12]

While no true competitive replication fitness assay has been developed, the current *in vitro* assays can determine relative replication yield phenotype to compare HBV DNA encoding specific mutations to the reference, consensus, or wild-type genetic framework.[15,16] For example, HBV replication competent clones encoding the rtV173L (valine to leucine substitution) + rtL180M + rtM204V were demonstrated to have an increased replication yield phenotype but no change in sensitivity to lamivudine relative to HBV clones encoding rtL180M + rtM204V.[17] Therefore, the rtV173L change should be considered a compensatory mutation in the HBV DNA genome.

New HBV nucleotide and corresponding deduced amino acid mutations detected in 2 or more patients undergoing the same antiviral treatment with virologic breakthrough despite medication compliance can be defined as a "putative" resistance mutation, and given a provisional status until confirmatory *in vitro* phenotype testing can be performed. Thus, there need to be sufficient data to determine if the selection of a particular nucleotide mutation and deduced amino acid change is unique to the antiviral selection pressure. In addition, the mutation should not be detected in patients who have continued to respond to the same antiviral treatment or be present prior to treatment using readily available assays that detect mutants comprising >5% of the viral population. Although antiviral-resistant mutations can be detected using ultra-sensitive assays (see below) from pretreatment samples and/ or before virologic breakthrough, the likelihood of HBV encoding the mutation(s) as the dominant quasispecies in pretreatment samples without selection pressure is very low.

Confirmation of genotypic resistance is based on 2 methods: (a) *in vitro* phenotypic analysis and (b) virtual phenotypic analysis which is the correlation of patient treatment and response data with HBV sequence data.

(i)* In Vitro *Phenotypic Analysis. *In vitro* phenotypic assay is the "gold standard" to confirm genotypic antiviral resistance. Unfortunately, the methodology is time consuming and labor intensive due to the lack of a convenient cell culture system and the need to use specific HBV replication competent clones. In addition, multiple substitutions or sequences elsewhere in the HBV genome may influence the result.

(ii) Virtual Phenotypic Analysis: Correlation of Patient Treatment and Response Data with HBV Sequence Data. This method relies on relational HBV databases with both clinical, virological, and HBV sequence information that are integrated and analyzed statistically via linkage. [18] Large numbers of patients with virologic breakthrough during treatment are required. The input clinical and sequence data are compared to the database to determine the best match and most likely treatment response.

(a) Nomenclature

After the initial detection and reporting of antiviral resistance, there was confusion in the naming of drug-resistant mutations as the HBV genotypes vary in genomic length. In 2001, Stuyver and colleagues overcame this problem by dividing the HBV polymerase into four different functional units and re-numbering each functional unit.[19] The reverse transcriptase (rt) region of the polymerase gene is common for all genotypes. Mutations within this region are prefixed with the letters rt followed by the original deduced amino acid, the codon number relative to the start of the rt region, followed by

the deduced amino acid derived by the mutation. For example, the primary LAM resistance associated changes are defined as rtM204I and rtM204V using this nomenclature (substitution of methionine at codon 204 in the reverse transcriptase region of the HBV polymerase gene for isoleucine or valine). These LAM-resistant changes should not be referred to as "a YMDD mutant" as only the methionine of the conserved YMDD locus is changed.

(b) Technical Issues Associated with the Genotypic Identification of New Antiviral Resistance Mutations

The accepted standard for characterizing the sequence of HBV polymerase in order to identify known or new resistance mutations is double-stranded (or bi-directional) sequencing of the PCR amplified product. Reports on newly identified antiviral resistance mutations should include changes in nucleotide and amino acid sequences in the reverse transcriptase region of the HBV polymerase as well as any deduced amino acid change in the overlapping envelope reading frame and the sequences should be deposited in Genbank. Where more than one nucleotide change (i.e., mixed HBV population) is detected at the same position, both deduced amino acids using the International Union of Pure and Applied Chemistry (IUPAC) codes should be listed. For example, a mixed population of wild type sequence and LAM-resistant mutation would be reported as rtM204M/V and the corresponding change from isoleucine at codon 195 of the overlapping HBV surface protein to a mixture of methionine and isoleucine indicated as sI195M/I.

(c) Methods to Detect Genotypic Resistance Mutations

Assays available to identify resistance mutations include direct sequencing of PCR products, PCR amplification followed by sequence analysis of multiple clones derived from the amplicon, real time PCR formats with specific probes including allele specific PCRs, hybridization methods such as the line probe assay, restriction fragment length polymorphism (RFLP), and more recently matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) based restriction fragment mass polymorphism (RFMP) which can detect mutants that comprise <1% of the viral quasispecies.[20,21] The more sensitive assays can detect HBV DNA encoding the resistance mutations that represent 5%-10% of the entire HBV quasispecies enabling earlier identification of patients with genotypic resistance mutations prior to or at the time of virologic breakthrough. However, the utility of ultra-sensitive assays that can detect mutants comprising <1% of the viral population in predicting drug resistance remains to be determined.

(i) Direct PCR Sequencing. Direct PCR sequencing is the least sensitive in detecting minor populations of resistant mutants as HBV DNA encoding the resistance mutations needs to reach approximately 20% of the total HBV quasispecies pool before these mutations can be observed. Direct sequencing is also not amenable to high throughput screening. However, direct PCR sequencing does allow for all mutations to be identified, including additional potential compensatory mutations and new undefined mutations. For new therapies, or the identification of new mutations associated with resistance to existing therapies, *in vitro* phenotypic confirmatory assays are required.

(ii) RFLP Analyses. RFLP analyses can detect viral mutants that constitute as little as 5% of the total viral population.[22] However, separate sets of endonuclease reactions must be designed specifically for each mutant of interest. Some mutations result in a new restriction site and RFLP is therefore an easy method; some other mutations destroy a restriction site and in this case RFLP analysis should be used with caution as lack of enzyme digestion may be due to loss of a restriction site or technical problems with the assay. RFLP analysis may not be possible for all resistant mutations as endonucleases specific for such sequences may not exist.

(iii) Reverse Hybridization Assay LiPA DR. The commercially available *reverse hybridization assay LiPA DR* (Innogenetics, Belgium) can detect single nucleotide mismatches and the assay contains a series of short membrane-bound oligonucleotide probes. LiPA assays can detect emerging viral resistance when HBV encoding the resistance mutations constitute 5% or more of the total viral population.[21,23-25] Their major limitation is that new sets of specific probes are required for every mutant and due to genotype variability a number of probes may be required to detect a single nucleotide change.

(iv) Sequencing with Microchip-Based Technology. Sequencing with microchip-based technology using oligonucleotide microarrays may be used to detect "new" mutants. This technology is expensive and not widely available.[26]

(v) MALDI-TOF MS. MALDI-TOF MS is based on mass spectrometric analysis of small DNA fragments containing sites of variation.[20] This assay has been shown to be very sensitive and can detect mutants that constitute only 1% of the viral population.[21] However, a new set of primers must be designed to detect each new mutation and access to mass spectrometer is required.

(vi) Single Genome Sequencing. Single genome sequencing has been used in studying antiviral-resistant HIV mutations.[27,28] This method is tedious and does not

Table 2. *In vitro* and *in vivo* Significance of Antiviral-Resistant Mutations

	Lamivudine	Clevudine	Telbivudine	Entecavir	Adefovir	Tenofovir
HBV	-Fold Resistance	-Fold Resistance	-Fold Resistance	-Fold Resistance	-Fold Resistance	-Fold Resistance
Wild-type	1	1	1	1	1	1
M204I	>100 [a, h]	>100 [a]	4 [d]	1 [d]	<1-8 [a, b, c]	<1 [d]
L180M + M204V	>100 [a, h]	>100 [a]	NA	5-6 [e, f]	<1-4 [a, b, c]	3-6 [f, g]
A181 T/V	1-2 [d, h]	NA	5-6 [d]	1-4 [d]	1-3 [d]	1 [d]
N236T	1 [f, h]	NA	3 [d]	<1 [f]	3 [f]	5 [f]
I169T + V173L + M250V	>1000 [d, h]	NA	>1000 [d]	>700 [e]	1 [d]	<1 [d]
T184G + S202I *	>1000 [d, h]	NA	35 [d]	>700 [e]	2 [d]	6 [d]
A194T	NA	NA	NA	NA	NA	2 [h]

* (+ L180M + M204I/V); NA = Not Available

2 - 9 fold → no or low level of resistance; 10 - 99 fold → medium level of resistance; >100 fold → high level of resistance

References: [a] Chin et al.[53]; [b] Delaney et al. (2001)[34]; [c] Ono et al.[73]; [d] Sozzi et al.[74]; [e] Tenney et al.[47]; [f] Brunelle et al.[48]; [g] Sheldon et al.[58]; [h] Delaney et al. (2006).[59]

provide information on the absolute frequency of the mutation. In addition, given the high rate of spontaneous mutations during HBV replication, the clinical significance of mutants that may be present in <0.1% of the viral population is uncertain.

With four approved NAs for chronic hepatitis B and several more in development, designing assays that would permit detection of all the mutations known to confer resistance to these medications is increasingly difficult. Any of the above assays can be used for research as long as the method is specified in the report. The most commonly used methods in clinical practice include direct sequencing and line probe assay.

3. Definition of *In Vitro* Phenotypic Antiviral Resistance

Demonstration that a given amino acid substitution confers resistance is based on the use of phenotypic assays that demonstrate *in vitro* a reduced susceptibility of a replication competent clone with that substitution to an antiviral agent against HBV as compared to a replication competent clone with the "wild-type" sequence (without the substitution). *In vitro* phenotype testing is based on the determination of changes to the effective concentration of the drug required to inhibit 50% of the target (EC_{50} or IC_{50}) relative to the "wild-type" reference HBV.

Clinically, antiviral drug resistance is commonly described as high- (>100-fold increase), intermediate- (10-99–fold increase) or low-level (2-9–fold increase) (Table 2). Unfortunately, such ranking of drug resistance as determined from *in vitro* phenotypic tests does not readily translate to what is observed clinically. For example, a small decrease in *in vitro* ADV susceptibility (2-9–fold increase in EC_{50}) may confer resistance *in vivo*.

(a) Methods Used for Phenotyping

Antiviral susceptibility testing involves assaying the activity of the polymerase enzyme or of HBV replication. These methods are time consuming and require a high level of technical expertise.[29]

(i) Phenotyping Based on Enzymatic Assays. Currently, the only assays for the study of HBV polymerase activity are based on the polymerase expressed in insect cells using a baculovirus vector, and on HBV polymerase expression in purified HBV nucleocapsids.[30,31] A cell-free assay has been developed for duck hepatitis B virus (DHBV) but HBV polymerase gene contains a large insert not present in DHBV or other hepadnavirus polymerase genes, and this insertion may affect the interpretation of phenotype testing results.[32,33]

(ii) Phenotyping Based on Transient Transfection of Hepatocyte-Derived Cell Lines. Two approaches based on transient transfection have been used for phenotyping. The first relies on site-directed mutagenesis to generate point mutations that may be associated with drug resistance in recombinant, well-characterized "laboratory" HBV replication competent clones. This approach may be an advantage for research purposes as specific mutation(s) can be determined as being or not being associated with reduced antiviral susceptibility. However, in some instances multiple mutations and/or the broader genetic framework of the HBV DNA from a patient may be required to confer antiviral resistance. The second method relies on amplified full-length HBV genomes from clinical isolates (rather than laboratory generated mutants).[16] However, as a number of mutations and/ or clonal variants may be present; the effect may be a culmination of all the changes in the reverse transcriptase region as well as other aspects of the HBV genome sequence unique to that particular patient and not simply the effect of one particular mutation.

(iii) Phenotyping Based on Transduction of Recombinant Baculovirus/HBV into Hepatocyte-Derived Cell Lines. Recombinant baculovirus/HBV has also been used for phenotyping.[15,34] In this system, HBV rep-

lication is driven by endogenous promoters and therefore studies on relative replication yield phenotype can be performed; however, this technique is tedious.[15]

(iv) Phenotyping Based on Continuous Cell Lines Containing Stably-Integrated HBV Genomes. Stably-transfected HBV-expressing cell lines have been created specifically for the investigation of drug resistance.[35] The advantage of using stable cell lines for phenotype testing is the ability to perform cross-resistance testing in a consistent environment but a new cell line needs to be created for each new mutant.[36] The integration site of HBV DNA within the cellular chromosome may affect HBV replication and cellular function; therefore, these cell lines cannot be used to determine the relative replication efficiency of HBV encoding the antiviral resistance mutations.

4. HBV Resistance Databases and Virtual Phenotyping

Virtual Phenotyping involves assigning a phenotype for a clinical isolate based on the correlations from large databases containing genotypic, phenotypic, and clinical information.

The analytical program searches linked databases for the best matches among sequences known to confer particular phenotypes. Virtual phenotyping should be considered as an adjunct and not a substitute to *in vitro* phenotype testing.

The first program developed to correlate HBV patient clinical, virological, and HBV sequence information is SeqHepB http://www.seqvirology.com/genome7/index.htm. Registered users can access the program online, and input either genomic (nucleic acid) or amino acid sequences of clinical HBV isolates for analysis. SeqHepB defines all amino acid variations within the input sequences compared to reference sequence(s) of the same genotype as mutations[37] and correlate the results with a database of HBV genotype and phenotype data and clinical histories. Various data mining algorithms and functions are being developed that facilitate rapid and efficient identification of new markers of drug resistance. One such function includes the localization of deduced amino acid changes on a molecular model of the HBV reverse transcriptase.[38] This function can provide insights into the significance of a mutation in relation to antiviral drug resistance and potential mechanism for antiviral resistance.[38,39]

The European Network of Excellence on antiviral drug resistance management, ViRgil, is also developing a database using a common clinical record form and a centralized virology laboratory.

The Hepatitis Virus Database (HVDB) in Japan is open to the public http://s2as02.genes.nig.ac.jp. It is updated 4 times a year using the newest release of DDBJ (DNA Data Bank of Japan) and currently contains 10,892 HBV entries. The HVDB is mainly geared towards phylogenetic analyses.

If properly designed, a database can greatly facilitate the tracking of known resistance mutations and the development of new ones. It would also allow more fundamental research on mutation interactions, epidemiological studies on the spread of resistant mutations, and systematic study of the risk factors associated with the emergence of resistant mutations and the clinical outcomes of patients with antiviral resistance. Additional advantages are an easily accessible overview of known resistance mutations for treating physicians and timely dissemination of information on resistance profile changes.

5. Mutations Associated with Antiviral Resistance

The incidence of genotypic resistance is related to viral (pretreatment serum HBV DNA level, pre-existing antiviral-resistant mutations), host (immune status, pharmacodynamics), and treatment characteristics (potency, genetic barrier to resistance [number of mutations required to produce a marked decrease in susceptibility to the antiviral drug], and duration of treatment). The incidence of genotypic resistance also varies with the sensitivity of the methods used for detection of resistant mutations (see section 2: Definition of Genotypic Antiviral Resistance) and the patient population being studied. Thus, clinical studies have varied from testing samples from all patients with detectable serum HBV DNA by PCR assay using sensitive methods such as RFLP or reverse hybridization to testing only samples from patients with viral rebound (such as >5 $\log_{10}$ copies/ml or >4.3 $\log_{10}$ IU/ml) using less sensitive methods such as direct sequencing (Table 3). These approaches result in reports of LAM-associated resistance mutations after 1 year of therapy varying from 7% to 23%.[14,40] Figure 1 illustrates mutations associated with approved HBV therapies.

(a) Resistance to Monotherapies

(i) Lamivudine and Other L-Nucleosides. The primary LAM resistance mutation maps to codon rtM204V/I in the YMDD motif.[12,41-43] *In vitro* studies showed that these mutations decrease sensitivity to lamivudine by >100-fold (Table 3). The molecular mechanism of LAM resistance is steric hindrance caused by the β-branched side group of the valine or isoleucine amino acids colliding with the oxathiolane ring of LAM within the dNTP binding site.[38] The rtL180M is the main compensatory change.[12] Other compensatory mutations in-

Table 3. Rates of Antiviral-Resistant HBV Mutations Reported in Clinical Trials

LAM-R pts	Rates of genotypic resistance	Patients tested	Method used
LAM	15-30% after 1 yr [a b c] 70% after 5 yr [d]	All pts PCR+	RFLP
ADV	0% after 1 yr [e] 29% after 5 yr [f]	All pts PCR+	Direct sequencing
LAM-R pts	~20% after 2 yr [g h]	All pts PCR+	Line-probe, MALDI-TOF
ETV	0% after 1 yr <1% after 2 & 3 yr [i]	All pts PCR+	Direct sequencing
LAM-R pts	1%, 9%, ~17% after 1, 2 & 3 yr [j]	Pts with viral rebound	Direct sequencing
LdT (LAM)	2-3% after 1 yr [k] 7-8% after 1 yr [k]	Pts with viral rebound	Direct sequencing

Abbreviations: LAM-R, lamivudine resistance; pts, patients.

References: [a] Dienstag et al.[75]; [b] Lai et al. (1998)[76]; [c] Schalm et al.[77]; [d] Lok et al.[14]; [e] Westland et al.[78]; [f] Hadziyannis et al.[79]; [g] Fung et al.[57]; [h] Lee et al.[80]; [i] Colonno et al.[62]; [j] Colonno et al.[81]; [k] Lai et al. (2005).[82]

clude the rtV173L and rtL80I changes.[9,17] The rtA181T change has been reported to occur in the absence of rtM204I/V and is considered a primary resistance mutation.[44,45] The rtA181T change is also selected during ADV treatment.[46] LAM-resistant mutations rtM204V/I +/− L180M decrease susceptibility to ETV.[36,47]

The primary LAM resistance mutations −rtM204V/I are cross-resistant with other L-nucleosides such as emtricitabine (FTC), telbivudine (beta-L-thymidine, LdT), beta-L-2′-deoxycytidine (LdC), elvucitabine and clevudine (L-FMAU; 2′-fluoro-5-methyl-beta-L-arabinofuranosyluracil), and the rtM204V/I changes have been selected during treatment with these compounds except for telbivudine where only rtM204I but not rtM204V has been observed.[7,48-51]

Telbivudine is not active against HBV encoding rtM204I or rtM204V + rtL180M in cell culture (www.fda.gov). Lamivudine, telbivudine, and clevudine exhibit modest activity against HBV encoding rtM204V alone in cell culture[52,53] but rtM204V mutation alone is rarely detected in patients. Thus, neither clevudine nor telbivudine is expected to be efficacious in patients with lamivudine-resistant HBV.

(ii) Acyclic Phosphonates. *Adefovir (ADV)* — The primary ADV-resistance mutations are rtN236T and /or rtA181T/V.[39,46,54] Isolates of HBV with the rtN236T change are susceptible to LAM while isolates with rtA181T/V changes have decreased susceptibility to LAM (Table 3). These ADV-associated mutations result in only a modest (2-9–fold) increase in EC_{50} but viral rebound, hepatitis flares and hepatic decompensation have been observed in patients.[55] ADV-associated mutations are partially cross-resistant with tenofovir. Decrease in serum HBV DNA levels observed when patients with virologic breakthrough due to ADV resistance are switched to tenofovir treatment is likely related to the higher dose of tenofovir than adefovir (300 mg versus 10 mg) used in clinical practice. The mechanism by which rtN236T confer resistance to ADV is thought to be due to indirect perturbation of the tri-phosphate binding site of the HBV pol.[38,54] Isolates of HBV with rtN236T and rtA181V changes are susceptible to entecavir *in vitro*.[39,48,56] Case reports have confirmed the *in vivo* efficacy of lamivudine and entecavir in the suppression of adefovir resistant HBV.[39,55,57]

Tenofovir (TDF) — There was a recent report of resistance to TDF associated with HBV encoding changes in HBV polymerase at rtL180M + rtA194T (alanine to threonine substitution) + rtM204V in 2 HIV / HBV co-infected patients.[58] HBV DNA was persistently detected in these 2 patients during TDF and LAM therapy but only 1 patient had an increase in serum HBV DNA. Aminotransferase levels remained normal in both patients and results of phenotypic studies are conflicting.[59] Data on TDF resistance in patients who received tenofovir monotherapy are not available because tenofovir has been used predominantly in patients with HIV/HBV coinfection, in combination with lamivudine or emtricitabine.

(iii) Entecavir. Two different ETV resistance genotypic profiles have been reported and confirmed *in vitro*.[47] The first pattern of ETV resistance includes: rtI169T + rtL180M + rtM204V + rtM250V and the second pattern includes: rtL180M + rtT184G + rtS202I +rtM204V (Table 3). Other patterns including triple ETV-associated mutations have also been

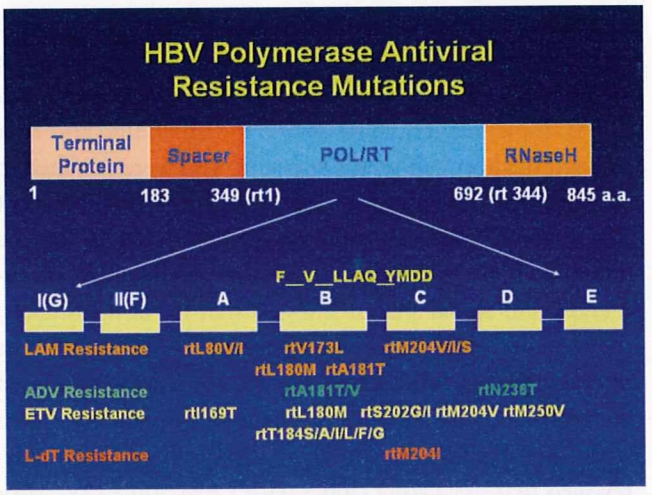

Fig. 1. The hepatitis B virus (HBV) polymerase open reading frame showing the conserved domains and location of the primary antiviral drug resistance mutations within the conserved domains. Rt, reverse transcription, LAM, lamivudine, ADV, adefovir dipivoxil, ETV, entecavir, LdT, telbivudine.

observed.[60] In the absence of the LAM-associated resistance mutations, the rtM250V change causes a 9-fold increase in EC_{50} while the changes at rt 169 (I169T), 184 (T184A/F/G/I/L/S) or 202 (S202G/I) changes have little effect (www.fda.gov). However, in the presence of LAM-associated resistance mutations, the ETV-associated mutations decrease susceptibility to ETV by >100-fold, particularly when 2 or more ETV-associated mutations are present.[47] HBV isolates with ETV-associated mutations are sensitive to adefovir and tenofovir *in vitro* and *in vivo* data confirming the efficacy of adefovir has been reported in one patient.[61]

Recent data suggest that pre-existence of mutations such as rtL180M + rtM204V even when present in <0.1% of the viral population in patients who have not received lamivudine treatment may increase the risk of selection for ETV-associated mutations in patients receiving entecavir treatment.[62] In addition, virologic breakthrough may occur in nucleoside-naïve patients receiving entecavir treatment due to selection of LAM-associated mutations alone. These data suggest that resistance to entecavir occurs through a 2-hit mechanism. Initially, LAM-associated mutants (rtM204V/I) are selected because they are less sensitive to ETV than wild type HBV. Virologic breakthrough usually occurs only after emergence of additional ETV-associated mutations, but rarely virologic breakthrough may occur after emergence of LAM-associated mutations alone.[62]

(b) Multidrug Resistance

Sequential treatment with NA monotherapy has resulted in the sequential selection of mutations conferring resistance to the initial therapy and subsequently the rescue therapy. For example, sequential resistance to lamivudine and later adefovir has been reported in patients who were switched to adefovir monotherapy for LAM-resistant HBV.[60] Recent studies reported that multidrug resistance changes can be detected in patients who received sequential NA monotherapy and clonal analysis showed that in most instances, the mutations associated with both treatments reside in the same clone.[56,60] The collocation of mutations associated with resistance to different treatments on the same genome is worrisome because *in vitro* analysis of antiviral sensitivities revealed that replicating clones with LAM and ADV- associated mutations had >50-fold reduced susceptibility to combination of LAM and ADV indicating that combination therapy of the 2 drugs may not be effective in suppressing multidrug resistant HBV.[54,56]

Table 4. Strategies for treating Antiviral-Resistant HBV

Type of Resistance	Rescue Therapy
Lamivudine/telbivudine resistance	•Add adefovir or tenofovir •Switch to emtricitabine + tenofovir •Switch to entecavir (risk of subsequent entecavir resistance and multidrug resistance)
Adefovir resistance*	•Add lamivudine or switch to emtricitabine + tenofovir •Add entecavir (if no prior lamivudine resistance)
Entecavir resistance*	•Add adefovir or tenofovir
Multidrug resistance#	•MDR to LAM + ADV: consider tenofovir + emtricitabine, tenofovir + entecavir •MDR to LAM + ETV: consider tenofovir or tenofovir + emtricitabine

Abbreviation: MDR, multidrug resistance.
*Limited *in vivo* data, available data indicate that addition of rescue therapy is less likely to result in sequential drug resistance than switching to rescue therapy.
In vivo data lacking

6. Monitoring and Treatment of Antiviral-Resistant HBV

(a) Monitoring for Virologic Response and Breakthrough

All patients receiving NA therapy for hepatitis B should be closely monitored for virologic response and breakthrough during treatment and for durability of response and viral relapse after treatment has stopped. Serum HBV DNA should be tested prior to treatment and then every 3 months during treatment. Patients with primary nonresponse should be considered for alternative treatment to facilitate clinical response and to minimize subsequent antiviral resistance. Patients with virologic breakthrough should be questioned about medication compliance. Tests for antiviral-resistant mutations should be performed whenever possible to confirm genotypic resistance, and to determine the pattern of mutations. The latter is particularly important as an increasing number of patients have been exposed to more than one anti-HBV NA.

(b) Treatment of Antiviral-Resistant HBV (Table 4)

Recommendations on treatment of patients with antiviral-resistant HBV depend on knowledge of the history of HBV treatments, virologic response to these treatments, the pattern of mutations detected at the time of virologic breakthrough, and *in vitro* data on antiviral activity of various HBV NAs against HBV isolates that harbor the mutations detected. Recent data suggest that initiating rescue therapy when virologic breakthrough is detected is more effective than delaying rescue therapy until viral rebound or biochemical breakthrough.[63]

(i) Lamivudine, Telbivudine and Other L-Nucleoside Resistance. *In vitro* studies have demonstrated that adefovir, tenofovir, and entecavir have antiviral activity against LAM- and other L-nucleoside-resistant HBV mutants, but the activity of entecavir against these mutants is substantially lower than for wild-type HBV (Table 3).[36] A pilot study in patients with lamivudine-resistant HBV reported that adefovir monotherapy resulted in similar rates of decrease in serum HBV DNA levels as combination therapy of lamivudine and adefovir.[64] However, combination of lamivudine and adefovir is more effective in preventing subsequent adefovir resistance.[57,65] Tenofovir has also been reported in clinical studies to be effective in suppressing lamivudine-resistant HBV.[66,67] Although tenofovir is more potent than adefovir, it is best used in combination with lamivudine or emtricitabine to prevent drug resistance. Entecavir has been shown in clinical trials to be effective in suppressing lamivudine-resistant HBV but a higher dose (1.0 mg daily) should be used.[68] Preexisting LAM-resistant mutations increase the risk of entecavir resistance[47,62]; therefore, entecavir is not an optimal treatment for patients with lamivudine-resistant HBV. If entecavir is used, lamivudine should be discontinued.

Based on *in vitro* data and the detection of rtM204I in patients with telbivudine resistance, the approach described above can be applied to patients with telbivudine-resistant HBV.

(ii) Adefovir. *In vitro* studies demonstrated that lamivudine and entecavir have antiviral activity against ADV-resistant HBV mutants (Table 3).[39,54] Case studies confirmed that lamivudine is effective in suppressing serum HBV DNA levels in patients with ADV-resistant HBV.[39,54] However, the durability of response, particularly in patients with prior lamivudine resistance is unknown. Furthermore, in the latter patients, a rapid reemergence of LAM-resistant mutations has been observed on reintroduction of lamivudine.[59,60] Case studies have reported that patients with primary nonresponse to adefovir experience further viral suppression when treatment is switched to tenofovir[69] presumably due to the higher dose of tenofovir used in clinical practice. For the same reason, tenofovir may result in some degree of viral suppression in patients with ADV-resistant HBV but the efficacy is likely limited due to *in vitro* evidence of cross resistance. Entecavir has been reported to be efficacious in 2 patients with ADV-resistant HBV.[57]

(iii) Entecavir. *In vitro* studies demonstrated that adefovir and tenofovir have antiviral activity against ETV-resistant HBV mutants[47,48] (Table 2) but clinical data on the efficacy of these treatments in patients with ETV-resistant HBV are not yet available.

(iv) Multidrug-Resistant HBV. The most effective treatment of multidrug-resistant HBV is prevention through judicious use of NA therapy and avoidance of sequential NA monotherapies. Thus, patients with minimal disease and those who are unlikely to achieve sustained response (such as inactive carriers and HBeAg-positive patients in the immune tolerance phase) should not be treated with NA therapy, particularly if they are young. When possible, the most potent NA with the lowest rate of genotypic resistance should be administered and compliance reinforced. Response should be closely monitored and modification of treatment considered in patients with primary nonresponse. *De novo* combination therapy of lamivudine plus pegylated interferon or adefovir has been shown to be associated with lower rates of virologic breakthrough compared to lamivudine monotherapy,[70-72] but resistance was not completely prevented. Studies on other *de novo* combination therapies are needed to determine the optimal combination of drugs and its cost-effectiveness.

7. Conclusions

Antiviral resistance and noncompliance to therapy are the most important cause of treatment failure in patients with hepatitis B. As more treatments become available, the complexity of antiviral-resistant mutations and the options for primary as well as rescue therapy increase. Therefore, there is an urgent need for standardization of (i) nomenclature on antiviral resistance, (ii) assays used in detection or confirmation of resistance, and (iii) format for reporting both *in vitro* and *in vivo* resistance data. This document represents the collaborative efforts of investigators from North America, Europe, Asia, and Australia. The group recognizes the importance of information sharing not only among investigators but also between investigators and practicing physicians such that new information generated from research can be rapidly disseminated and observations in clinical practice can be validated. To this end, the group hopes to establish a database on hepatitis B virus drug resistance that will be freely accessible and to develop means for sharing technology, clinical samples, and HBV isolates.

Acknowledgment: The authors express their appreciation for the valuable comments from other members of the Hepatitis B Virus Drug Resistance Working Group, in particular Drs. Rajen Koshy, T. Jake Liang and William Mason.

References

1. Hoofnagle JH, Doo E, Liang TJ, Fleischer R, Lok ASF. Management of Hepatitis B: Summary of a Clinical Research Workshop. HEPATOLOGY 2007;45:1056-1075.

2. Eriksson S, Wang L. The role of cellular deoxynucleoside kinases in the activation of nucleoside analogs used in chemotherapy. Recent Advances in Nucelosides: Chemistry and Chemotherapy, Chu C.K. ed. Amsterdam: Elsevier, 2002:455-475.
3. Hulgan T, Haas DW. Toward a pharmacogenetic understanding of nucleotide and nucleoside analogue toxicity. J Infect Dis 2006;194:1471-1474.
4. Schildgen O, Sirma H, Funk A, Olotu C, Wend UC, Hartmann H, et al. Variant of hepatitis B virus with primary resistance to adefovir. N Engl J Med 2006;354:1807-1812.
5. Marcellin P, Chang TT, Lim SG, Tong MJ, Sievert W, Shiffman ML, et al. Adefovir dipivoxil for the treatment of hepatitis B e antigen-positive chronic hepatitis B. N Engl J Med 2003;348:808-816.
6. Yuen MF, Sablon E, Hui CH, Yuan HJ, Decraemer H, LaibCL. Factors preceding hepatitis B virus DNA breakthrough in patients receiving prolonged lamivudine therapy. HEPATOLOGY 34:785-791 2001:785-791.
7. Lai CL, Leung N, Teo EK, Tong M, Wong F, Hann HW, et al. A 1-year trial of telbivudine, lamivudine, and the combination in patients with hepatitis B e antigen-positive chronic hepatitis B. Gastroenterology 2005; 129:528-536.
8. Locarnini S, Qi X, Arterburn S, Snow A, Brosgart CL, Currie G, et al. Incidence and predictors of emergence of Adefovir resistant HBV during four years of adefovir dipivoxil (ADV) therapy for patients with chronic hepatitis B (CHB) [Abstract]. J Hepatol 2005;42:17.
9. Ogata N, Fujii K, Takigawa S, Nomoto M, Ichida T, Asakura H. Novel patterns of amino acid mutations in the hepatitis B virus polymerase in association with resistance to lamivudine therapy in japanese patients with chronic hepatitis B. J Med Virol 1999;59:270-276.
10. Locarnini S, Hatzakis A, Heathcote J, Keeffe EB, Liang TJ, Mutimer D, Pawlotsky JM, et al. Management of antiviral resistance in patients with chronic hepatitis B. Antivir Ther 2004;9:679-693.
11. Zhou T, Saputelli J, Aldrich CE, Deslauriers M, Condreay LD, Mason WS. Emergence of drug-resistant populations of woodchuck hepatitis virus in woodchucks treated with the antiviral nucleoside lamivudine. Antimicrob Agents Chemother 1999;43:1947-1954.
12. Ono SK, Kato N, Shiratori Y, Kato J, Goto T, Schinazi RF, et al. The polymerase L528M mutation cooperates with nucleotide binding-site mutations, increasing hepatitis B virus replication and drug resistance. J Clin Invest 2001;107:449-455.
13. Nafa S, Ahmed S, Tavan D, Pichoud C, Berby F, Stuyver L, et al. Early detection of viral resistance by determination of hepatitis B virus polymerase mutations in patients treated by lamivudine for chronic hepatitis B. HEPATOLOGY 2000;32:1078-1088.
14. Lok AS, Lai CL, Leung N, Yao GB, Cui ZY, Schiff ER, et al. Long-term safety of lamivudine treatment in patients with chronic hepatitis B. Gastroenterology 2003;125:1714-1722.
15. Chen R, Edwards R, Shaw T, Colledge D, Delaney IV WE, Isom H, et al. Effect of the G1896A precore mutation on drug sensitivity and replication yield of lamivudine-resistant HBV *in vitro*. HEPATOLOGY 2003;37:27-35.
16. Durantel D, Carrouee-Durantel S, Werle-Lapostolle B, Brunelle MN, Pichoud C, Trepo C, et al. A new strategy for studying in vitro the drug susceptibility of clinical isolates of human hepatitis B virus. HEPATOLOGY 2004;40:855-864.
17. Delaney IV WE, Yang H, Westland CE, Das K, Arnold E, Gibbs CS, et al. The hepatitis B virus polymerase mutation rtV173L is selected during lamivudine therapy and enhances viral replication in vitro. J Virol 2003; 77:11833-11841.
18. Gonzales MJ, Wu TD, Taylor J, Belitskaya I, Kantor R, Israelski D, et al. Extended spectrum of HIV-1 reverse transcriptase mutations in patients receiving multiple nucleoside analog inhibitors. AIDS 2003;17:791-799.
19. Stuyver LJ, Locarnini SA, Lok A, Richman DD, Carman WF, Dienstag JL, et al. Nomenclature for antiviral-resistant human hepatitis B virus mutations in the polymerase region. HEPATOLOGY 2001;33:751-757.
20. Hong SP, Kim NK, Hwang SG, Chung HJ, Kim S, Han JH, et al. Detection of hepatitis B virus YMDD variants using mass spectrometric analysis of oligonucleotide fragments. J Hepatol 2004;40:837-844.
21. Kim HS, Han KH, Ahn SH, Kim EO, Chang HY, Moon MS, et al. Evaluation of methods for monitoring drug resistance in chronic hepatitis B patients during lamivudine therapy based on mass spectrometry and reverse hybridization. Antivir Ther 2005;10:441-449.
22. Allen MI, Gauthier J, DesLauriers M, Bourne EJ, Carrick KM, Baldanti F, et al. Two sensitive PCR-based methods for detection of hepatitis B virus variants associated with reduced susceptibility to lamivudine. J Clin Microbiol 1999;37:3338-3347.
23. Stuyver L, Van Geyt C, De Gendt S, Van Reybroeck G, Zoulim F, Leroux-Roels G, Rossau R. Line probe assay for monitoring drug resistance in hepatitis B virus-infected patients during antiviral therapy. Journal of Clinical Microbiology 2000;38:702-707.
24. Lok AS, Zoulim F, Locarnini S, Mangia A, Niro G, Decraemer H, et al. Monitoring drug resistance in chronic hepatitis B virus (HBV)-infected patients during lamivudine therapy: evaluation of performance of INNO-LiPA HBV DR assay. J Clin Microbiol 2002;40:3729-3734.
25. Hussain M, Fung S, Libbrecht E, Sablon E, Cursaro C, Andreone P, et al. Sensitive line probe assay that simultaneously detects mutations conveying resistance to lamivudine and adefovir. J Clin Microbiol 2006;44:1094-1097.
26. Tran N, Berne R, Chann R, Gauthier M, Martin D, Armand MA, et al. European multicenter evaluation of high-density DNA probe arrays for detection of hepatitis B virus resistance mutations and identification of genotypes. J Clin Microbiol 2006;44:2792-2800.
27. Palmer S, Wiegand A, Maldarelli F, Bazmi H, Mican J, Polis M, et al. New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. J Clin Microbiol. 2003;41:4531-4536.
28. Palmer S, Kearney M, Maldarelli F, Halvas EK, Bixby CJ, Bazmi H, et al. Multiple, linked human immunodeficiency virus type 1 drug resistance mutations in treatment-experienced patients are missed by standard genotype analysis. J Clin Microbiol 2005;43:406-413.
29. Zoulim F. In vitro models for studying hepatitis B virus drug resistance. Semin Liver Dis 2006;26:171-180.
30. Seifer M, Hamatake R, Bifano M, Standring DN. Generation of replication-competent hepatitis B virus nucleocapsids in insect cells. J Virol 1998; 72:2765-2776.
31. Seifer M, Hamatake RK, Colonno RJ, Standring DN. In vitro inhibition of hepadnavirus polymerases by the triphosphates of BMS-200475 and lobucavir. Antimicrob Agents Chemother 1998;42:3200-3208.
32. Jacquard AC, Brunelle MN, Pichoud C, Durantel D, Carrouee-Durantel S, Trepo C, et al. In vitro characterization of the anti-hepatitis B virus activity and cross-resistance profile of 2′,3′-dideoxy-3′-fluoroguanosine. Antimicrob Agents Chemother 2006;50:955-961.
33. Zoulim F, Seeger C. Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase. J Virol 1994;68:6-13.
34. Delaney WET, Edwards R, Colledge D, Shaw T, Torresi J, Miller TG, et al. Cross-resistance testing of antihepadnaviral compounds using novel recombinant baculoviruses which encode drug-resistant strains of hepatitis B virus. Antimicrob Agents Chemother 2001;45:1705-1713.
35. Ladner SK, Otto MJ, Barker CS, Zaifert K, Wang GH, Guo JT, et al. Inducible expression of human hepatitis B virus (HBV) in stably transfected hepatoblastoma cells: a novel system for screening potential inhibitors of HBV replication. Antimicrob Agents Chemother 1997;41:1715-1720.
36. Yang H, Qi X, Sabogal A, Miller M, Xiong S, Delaney WET. Cross-resistance testing of next-generation nucleoside and nucleotide analogues against lamivudine-resistant HBV. Antivir Ther 2005;10:625-633.
37. Yuen LK, Ayres A, Littlejohn M, Colledge D, Edgely A, Maskill WJ, et al. SEQHEPB: A sequence analysis program and relational database system for chronic hepatitis B. Antiviral Res 2006.
38. Bartholomeusz A, Tehan BG, Chalmers DK. Comparisons of the HBV and HIV polymerase, and antiviral resistance mutations. Antivir Ther 2004;9:149-160.
39. Villeneuve JP, Durantel D, Durantel S, Westland C, Xiong S, Brosgart CL, et al. Selection of a hepatitis B virus strain resistant to adefovir in a liver transplantation patient. J Hepatol 2003;39:1085-1089.

40. Lai CL, Gane E, Liaw YF, Thongsawat S, Wang Y, Chen Y, et al. Telbivudine (LDT) Vs. Lamivudine for Chronic Hepatitis B: First - year results from the international phase III globe trial [Abstract]. HEPATOLOGY 2005; 42(Suppl):748A.
41. Allen MI, Deslauriers M, Andrews CW, Tipples GA, Walters KA, Tyrrell DL, et al. Identification and characterisation of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group. HEPATOLOGY 1998;27:1670-1677.
42. Ling R, Mutimer D, Ahmed M, Boxall EH, Elias E, Dusheiko GM, et al. Selection of mutations in the hepatitis B virus polymerase during therapy of transplant recipients with lamivudine. HEPATOLOGY 1996; 24:711-713.
43. Tipples GA, Ma MM, Fischer KP, Bain VG, Kneteman NM, Tyrrell DL. Mutation in HBV RNA-dependent DNA polymerase confers resistance to lamivudine in vivo. HEPATOLOGY 1996;24:714-717.
44. Yeh CT, Chien RN, Chu CM, Liaw YF. Clearance of the original hepatitis B virus YMDD-motif mutants with emergence of distinct lamivudine-resistant mutants during prolonged lamivudine therapy. HEPATOLOGY 2000;31:1318-1326.
45. Yatsuji H, Noguchi C, Hiraga N, Mori N, Tsuge M, Imamura M, et al. Emergence of a novel lamivudine-resistant hepatitis B virus variant with a substitution outside the YMDD motif. Antimicrob Agents Chemother 2006;50:3867-3874.
46. Lacombe K, Ollivet A, Gozlan J, Durantel S, Tran N, Girard PM, et al. A novel hepatitis B virus mutation with resistance to adefovir but not to tenofovir in an HIV-hepatitis B viurs-co-infected patient. AIDS 2006;20: 1-3.
47. Tenney DJ, Levine SM, Rose RE, Walsh AW, Weinheimer SP, Discotto L, et al. Clinical emergence of entecavir-resistant hepatitis B virus requires additional substitutions in virus already resistant to Lamivudine. Antimicrob Agents Chemother 2004;48:3498-3507.
48. Brunelle MN, Jacquard AC, Pichoud C, Durantel D, Carrouee-Durantel S, Villeneuve JP, et al. Susceptibility to antivirals of a human HBV strain with mutations conferring resistance to both lamivudine and adefovir. HEPATOLOGY 2005;41:1391-1398.
49. Lim SG, Ng TM, Kung N, Krastev Z, Volfova M, Husa P, et al. A double-blind placebo-controlled study of emtricitabine in chronic hepatitis B. Arch Intern Med 2006;166:49-56.
50. Seigneres B, Pichoud C, Martin P, Furman P, Trepo C, Zoulim F. Inhibitory activity of dioxolane purine analogs on wild-type and lamivudine-resistant mutants of hepadnaviruses. HEPATOLOGY 2002;36: 710-722.
51. Lai CL, Leung NW, Teo EK, Tong M, Wong F, Hann HW, et al. Phase IIb Extended-Treatment Trial of Telbivudine (LdT) Vs Lamivudine Bs combination treatment in Hepatitis B patients: Two Year Results [Abstract]. J Gastroenterol 2005;128(Suppl 2):A692.
52. Ladner S, Miller T, Otto M, King R. The hepatitis B virus M539V polymerase variation responsible for 3TC resistance also confers cross-resistance to other nucleoside analogues. Antivir Chem Chemother 1998;9:65-72.
53. Chin R, Shaw T, Torresi J, Sozzi V, Trautwein C, Bock T, et al. In vitro susceptibilities of wild-type or drug-resistant hepatitis B virus to (-)-beta-D-2,6-diaminopurine dioxolane and 2'-fluoro-5-methyl-beta-L-arabino-furanosyluracil. Antimicrob Agents Chemother 2001;45:2495-2501.
54. Angus P, Vaughan R, Xiong S, Yang H, Delaney W, Gibbs C, et al. Resistance to adefovir dipivoxil therapy associated with the selection of a novel mutation in the HBV polymerase. Gastroenterology 2003;125:292-297.
55. Fung SK, Andreone P, Han SH, Rajender Reddy K, Regev A, Keeffe EB, et al. Adefovir-resistant hepatitis B can be associated with viral rebound and hepatic decompensation. J Hepatol 2005;43:937-943.
56. Villet S, Pichoud C, Villeneuve J-P, Trepo C, Zoulim F. Selection of a multiple drug resistant hepatitis B virus strain in a liver transplanted patient. Gastroenterology 2006;131:1253-1261.
57. Fung SK, Chae HB, Fontana RJ, Conjeevaram H, Marrero J, Oberhelman K, et al. Virologic response and resistance to adefovir in patients with chronic hepatitis B. J Hepatol 2006;44:283-290.
58. Sheldon J, Camino N, Rodes B, Bartholomeusz A, Kuiper M, Tacke F, et al. Selection of hepatitis B virus polymerase mutations in HIV-coinfected patients treated with tenofovir. Antivir Ther 2005;10:727-734.
59. Delaney IV WE, Ray AS, Yang H, Qi X, Xiong S, Zhu Y, et al. Intracellular metabolism and in vitro activity of tenofovir against hepatitis B virus. Antimicrob Agents Chemother 2006;50:2471-2477.
60. Yim HJ, Hussain M, Liu Y, Wong SN, Fung S, Lok A. Evolution of multi-drug resistant hepatitis B virus during sequential therapy. HEPATOLOGY 2006;44:703-712.
61. Villet S, Ollivet A, Pichoud C, Barraud L, Villeneuve JP, Trepo C, et al. Stepwise process for the development of entecavir resistance in a chronic hepatitis B virus infected patient. J Hepatol 2007;46:531-538.
62. Colonno RJ, Rose R, Baldick CJ, Levine S, Pokornowski K, Yu CF, et al. Entecavir resistance is rare in nucleoside naïve patients with hepatitis B. HEPATOLOGY 2006;44:1656-1665.
63. Lampertico P, Vigano M, Manenti E, Iavarone M, Lunghi G, Colombo M. Adefovir rapidly suppresses hepatitis B in HBeAg-negative patients developin genotypic resistance to lamivudine. HEPATOLOGY 2005;42: 1414-1419.
64. Peters MG, Hann H, Martin P, Heathcote EJ, Buggisch P, Rubin R, et al. Adefovir dipivoxil alone or in combination with lamivudine in patients with lamivudine-resistant chronic hepatitis B. Gastroenterology 2004;126: 91-101.
65. Snow A, Thibault V, Qi X, Zhu Y, Westland C, Arterburn S, et al. Combination of Adefovir Dipivoxil (ADV) and Lamivudine (LAM) prevented emergence of ADV resistance mutations in chronic hepatitis B (CHB) patients with LAM-resistant HBV. Gastroenterology 2005;128:M945.
66. van Bommel F, Wunsche T, Mauss S, Reinke P, Bergk A, Schurmann D, et al. Comparison of adefovir and tenofovir in the treatment of lamivudine-resistant hepatitis B virus infection. HEPATOLOGY 2004;40:1421-1425.
67. Kuo A, Dienstag JL, Chung RT. Tenofovir disoproxil fumarate for the treatment of lamivudine-resistant hepatitis B. Clin Gastroenterol Hepatol 2004;2:266-272.
68. Chang T, Chish R, Hadziyannis S, Cianciara J, Rizzetto M, Schiff E, et al. A dose-ranging study of the efficacy and tolerability of entecavir in Lamivudine-refractory chronic hepatitis B patients. Gastroenterology 2005; 129:1198-1209.
69. van Bömmel S, Zöllner B, Sarrazin C, Spengler U, Huppe D, Möller B, et al. Tenofovir for patients with lamivudine-resistant hepatitis B virus (HBV) infection and high HBV DNA level during adefovir therapy. HEPATOLOGY 2006.
70. Lau GK, Piratvisuth T, Luo KX, Marcellin P, Thongsawat S, Cooksley G, et al. Peginterferon Alfa-2a, lamivudine, and the combination for HBeAg-positive chronic hepatitis B. N Engl J Med 2005;352:2682-2695.
71. Marcellin P, Lau GK, Bonino F, Farci P, Hadziyannis S, Jin R, et al. Peginterferon alfa-2a alone, lamivudine alone, and the two in combination in patients with HBeAg-negative chronic hepatitis B. N Engl J Med 2004; 351:1206-1217.
72. Sung JJY LJ, Zeuzem S, Chow WC, Heathcote E, Perrillo R, Brosgart C, et al. A randomized double-blind phase II study of lamivudine compared to lamivudine plus adefovir dipivoxil for treatment naïve patients with chronic hepatitis B: week 52 analysis [Abstract]. J Hepatol 2003:25.
73. Ono SK, Kato N, Shiratori Y, Carrilho FJ, Omata M. Novel nucleoside analogue MCC-478 (LY582563) is effective against wild-type or lamivudine-resistant hepatitis B virus. Antimicrob Agents Chemother 2002;46: 2602-2605.
74. Sozzi T, Edwards R, Shaw T, Xiong S, Borroto-Esoda K, Locarnini S. antiviral cross-resistance between clinically important hbv mutants: phoenotypic testing using the recombinant HBV-Baculovirus Assay System. Global Antiviral J 2005;1(Suppl 2):79-80.
75. Dienstag JL, Schiff E, Wright TL, Perrillo R, Hann HW, Goodman Z, et al. Lamivudine as initial treatment for chronic hepatitis B in the United States. N Engl J Med 1999;341:1256-1263.
76. Lai CL, Chien RN, Leung NW, Chang TT, Guan R, Tai DI, et al. A one-year trial of lamivudine for chronic hepatitis B. Asia Hepatitis Lamivudine Study Group. [see comments]. N Engl J Med 1998;339:61-68.

77. Schalm S, Heathcote J, Cianciara J, Farrell G, Sherman M, Willems B, et al. Lamivudine and alpha interferon combination treatment of patients with chronic hepatitis B infection: a randomised trial. Gut 2000;46:562-568.
78. Westland C, Yang H, Delaney IV WE, Gibbs C, Miller M, Wulfsohn M, et al. Week 48 resistance surveillance in two phase 3 clinical studies of adefovir dipivoxil for chronic hepatitis B. HEPATOLOGY 2003;38:96-103.
79. Hadziyannis S, Tassopoulos N, Chang T, Heathcote J, Kitis G, Rizzetto M, et al. Long-term adefovir dipivoxil treatment induces regression of liver fibrosis in patients with hbeag-negative chronic hepatitis b: results after 5 years of therapy [Abstract]. HEPATOLOGY 2005;42(Suppl):754A.
80. Lee Y, Suh D, Lim Y, Jung S, Kim K, Lee H, et al. Increased risk of adefovir resistance in patients with lamivudine-resistant chronic hepatitis B after 48 weeks of adefovir dipivoxil monotherapy. HEPATOLOGY 2006;43:1385-1391.
81. Colonno R, Rose F, CJ B. High barrier to resistance results in no emergence of entecavir resistance in nucloside-naive subjects during the first two years of therapy [Abstract 490]. J Hepatol 2006;44(Suppl 1):S182.
82. Lai C, Gane E, Liaw Y-F, Thongawat S, Wang Y, Heathcote E, et al. Telbivudine (LDT) vs. lamivudine for chronic hepatitis B: first-year results from the International Phase III Globe Trial [Abstract]. HEPATOLOGY 2005;42(Suppl):748A.

Hepatology Research 2007 doi: 10.1111/j.1872-034X.2007.00262.x

Original Article

Development and public release of a comprehensive hepatitis virus database

Tadasu Shin-i,[1] Yasuhito Tanaka,[1] Yoshio Tateno[2] and Masashi Mizokami[1]

[1]*Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya and* [2]*Gene Function Research Laboratory, Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Mishima, Japan*

Aim: Currently, approximately 44 000 hepatitis C virus (HCV), 11 000 hepatitis B virus (HBV), and 1600 hepatitis E virus (HEV) sequences are available at the International Nucleotide Sequence Database Collaboration (INSDC, previously known as DDBJ/EMBL/GenBank), and the number of these virus sequences is growing rapidly. However, since INDSC is not specialized to hepatitis viruses, it is difficult to retrieve information of virological or clinical interests from it. Thus, it is quite worthwhile to construct a specialized database for the hepatitis virus sequences and to make it accessible to researchers worldwide.

Methods: We developed a WWW-based database hepatitis virus database (HVDB), which contains all the HCV, HBV, and HEV sequences available at INSDC. In the HVDB, all piece sequences obtained from INSDC are arranged to the genome sequence of each virus. Also given in the database are the phylogenetic relationships of each locus on the genome among variants for each virus.

Results: Users of the database can easily retrieve entries (sequences with annotations) of the specific genotype by referring to the phylogenetic relationships or those of specific loci by referring to the genome map information. HVDB provides users with a tool for phylogenetic analysis that can be used in combination with the data retrieval tools.

Conclusion: The latest release is publicly accessible at the HVDB website: http://s2as02.genes.nig.ac.jp.

Key words: database, genome, hepatitis virus, phylogenetic relationships.

INTRODUCTION

HEPATITIS C VIRUS (HCV) has infected approximately170 million people worldwide.[1,2] The majority of HCV infections result in persistent and chronic hepatitis, liver cirrhosis, and finally to hepatocellular carcinoma. HCV is a positive-stranded RNA virus of approximately 9400 nucleotides encoding a large polyprotein that is cleaved into structural and non-structural proteins.[3] The virus genome evolves so rapidly as to form quasi-species within the host. So far, six genotypes and their serial subtypes have been identified in the virus.[4] The genetic variation is not only of scientific interest, but also of clinical importance. For example, interferon-α treatment, which is the only applicable therapy for HCV patients at present, is more effective against genotypes 2, 3, and 5 than genotype 1.[5]

Hepatitis B virus (HBV) is also endemic around the world. The virus has infected more than two billion people worldwide, among which 350 million people are chronic carriers.[6–8] HBV has a partially double-stranded circular genome of approximately 3200 nucleotides, which contains four overlapping open reading frames (ORF). The virus has evolved over the years into many genotypes in humans and other primates, and eight genotypes have been identified in humans.[9] These genotypes have an uneven geographic distribution and only a few of them are prevalent in a given area of the world. Therefore, the information on each genotype can be used to trace back to its ancestor. The information is also useful for searching the transmission route of accidental exposures to HBV and evaluating infections from wildlife to humans or vice versa. Furthermore, the genotypes are strongly associated with the progression of liver

Correspondence: Dr Masashi Mizokami, Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Kawasumi 1, Mizuho, Nagoya 467-8601, Japan. Email: mizokami@med.nagoya-cu.ac.jp
Received 24 March 2007; revision 23 July 2007; accepted 1 August 2007.

diseases, as well as the response to antiviral therapies for HBV patients.[9]

Hepatitis E virus (HEV) is an unclassified virus that is the leading cause of enterically transmitted non-A, non-B hepatitis among adults in developing countries. The genome of HEV is a positive-stranded RNA of approximately 7200 nucleotides, and contains a short 5′ untranslated region (UTR), three ORF, and a short-3′UTR terminated by a poly(A) tract.[10,11]

So far, approximately 44 000 HCV, 11 000 HBV, and 1600 HEV sequences have been submitted to the International Nucleotide Sequence Database Collaboration (INSDC, previously known as DDBJ/EMBL/GenBank). However, since INSDC is not specialized to viruses, especially hepatitis viruses, it is difficult to retrieve the sequences of virological or clinical interest. Therefore, it is highly important to develop and release a well-organized database that contains all the available HCV, HBV, and HEV entries to researchers who are interested in the clinical treatment, transmission routes, relationships among hepatitis viruses, and/or their genomic information.

We had already constructed the HCV database, which has been open to the public since 1997. Recently, we added data about two more hepatitis viruses, HBV and HEV, the genomic locations of the sequences, and phylogenetic relationships among the variants, and called the extended database the hepatitis virus database (HVDB). In this paper, we describe the structure, function, application, and enhancement plan of HVDB.

METHODS

HVDB IS CURRENTLY composed of the HCV, HBV, and HEV databases. The basic contents of HVDB were obtained by extracting all the HCV, HBV, and HEV entries from the latest DDBJ release.[12] It is noted that DDBJ serves all the data collected by INSDC and regularly publicizes the DDBJ release four times a year. Once new DDBJ release is published, HVDB contents are fully reconstructed in 2 weeks.

The database schema for each of the three databases is made up of two aspects: the genomic location and the phylogenetic relationship. For the location contents, we aligned all the entry sequences to the reference genome. The alignment result was compiled to draw a map, in which the location of each entry is positioned on the genome, as shown in Figure 1. Then, we defined the pair of divisions for each locus on the genome. One division, the N-division, includes all nucleotide sequences that belong to a locus, irrespective of the sequence length, and the other, the A-division, includes all the translated amino acid sequences from the nucleotide sequences. Using the sequences, we construct a phylogenetic tree for each of the two divisions to show their evolutionary relationships. The phylogenetic trees for three N-divisions are shown in Figure 1. Each of the N- and A-divisions also includes a multiple alignment of the sequences and other pieces of information on the locus, as listed in Table 1. The information on the annotation is extracted from the DDBJ release. Each of a pair of divisions is also used as a unit of data analysis, so that users who want to analyze their own data in reference to the unit are able to make their own divisions; these are called "private divisions", and are described later.

HCV master database

The process of building the HCV master database is as follows:

1 All the entries are retrieved from the DDBJ release. This step is done by simple keyword searches in which the keywords, "HCV", "Hepatitis C", and "Hepatitis virus C" are used. The resulting data set is called the "source".
2 The reference genome sequence is selected from the extant HCV genome data. It is used as a template for locating all the entries. We chose the whole viral genome sequence of strain HCV-H77, AF009606[13] as the reference sequence, since it has been recommended to be used for the numbering of HCV sequences in the HCV studying community.[14]
3 The locations of loci and subregions are obtained from the reference sequence and compiled as the "reference map". In the case of the HCV database, the locations of seven loci (C, E1, NS1/E2, NS2, NS3, NS4, and NS5) and six subregions (E2, p7, NS4a, NS4b, NS5a, and NS5b) are positioned on the "reference map".
4 Each entry in the "source" is aligned against the reference sequence by using the LALIGN program in the FASTA[15] package. The results are compiled to create the "map information".
5 Divisions are created for all the loci and subregions presented on the "reference map". A pair of the N- and A-divisions is made for each of the seven loci. In addition, "full sequence division", which contains a set of the whole HCV genome sequences, is prepared.

Divisions

Each of N-divisions in the master database is made using the following process:

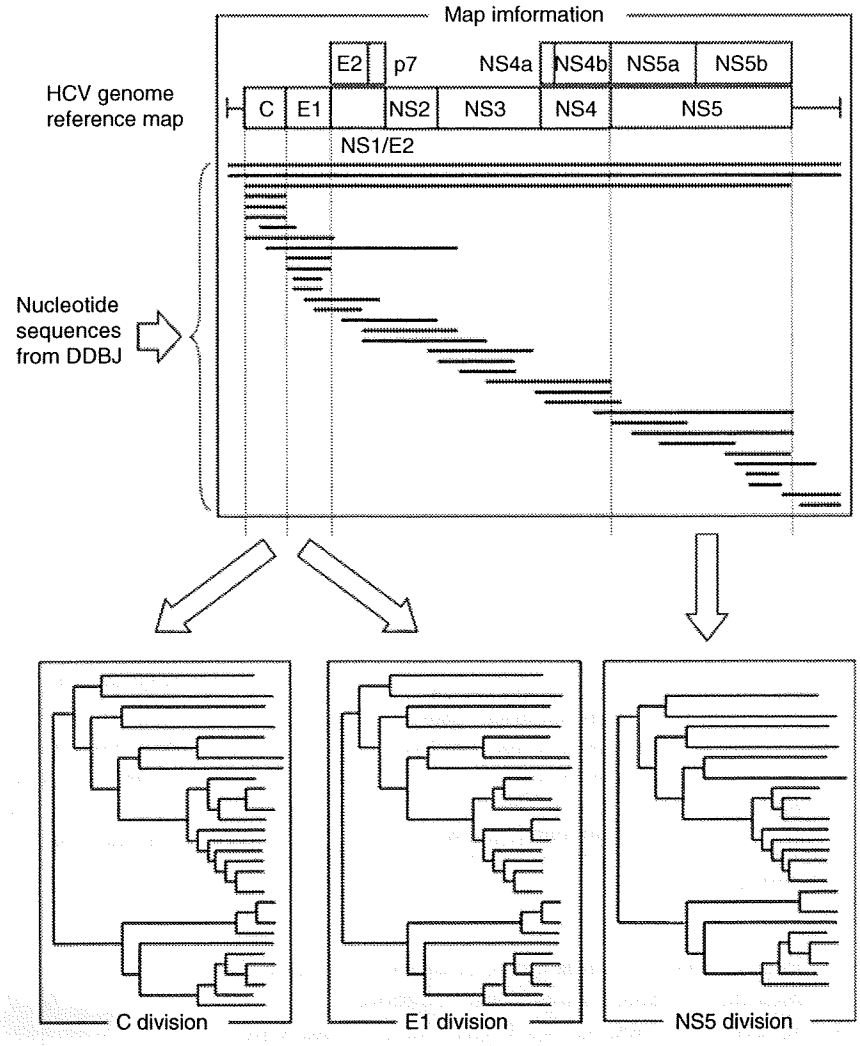

Figure 1 Schematic overview of the database structure. Figure shows the process of making the hepatitis C virus (HCV) master database, as an example. Each horizontal line below the HCV genome reference map shows the mapped position of an HCV entry obtained from the DDBJ.

1 All the entries that cover a locus are extracted from the master database by using the "map information" and "reference map". Each entry sequence is clipped at the both ends of the locus by referring to the "map information" if it exceeds either or the both ends. The results are then gathered to make the "sequence".
2 The pieces of annotation information for the entries are extracted from the "source" of the master database and recorded in the "annotation".
3 All the data in "sequences" are multiple-aligned using the CLUSTALW[16] program to create the "align".
4 The genetic distance matrix, the "matrix", is obtained from the "align" with the six-parameter method.[17]
5 A phylogenetic tree, the "tree", is constructed from the "matrix" using the Neighbor–Joining method.[18]

Each A-division is made from the corresponding N-division by translating the latter into the former. Then, a phylogenetic analysis is performed in almost the same manner as above, but Kimura's method[19] is used to estimate the evolutionary distance between the pair of translated amino acid sequences.

HBV master database

HBV has a circular genome, unlike HCV which has a linear one. We therefore adopted a different process for creating the HBV master database:

1 All the entries for this database are retrieved from the DDBJ database using "HBV", "Hepatitis B", and "Hepatitis virus B" as keywords, and thus the "source" is constructed.
2 Because the genotype G of HBV is known to have a 36 bp insertion in the precore region, it is suitable to choose the reference sequence from this genotype.

Table 1 Data types in a division

Type	Description	
Sequence	Nucleic or amino acids sequences by FASTA format.	
Annotation	Annotation of the entries. In the case of divisions in master database, they are described by DDBJ release format.	
Index	Table of entry ID, accession No., locus name, and sequence length.	
Amino index	CDS location and product length. (amino acids division only)	
Tag	Additional data. There may be one or more tags in a division.	Optional
Align	Multiple alignment of "sequence", by interleave format.	
Matrix	Genetic distance matrix calculated from "align".	
Tree	Phylogenetic tree calculated from "matrix".	
Bootstrap tree	Result of bootstrap resampling analysis.	Optional

CDS, coding sequence.

We then select the earliest one of them, AF160501,[20] which has a whole viral genome sequence.

3 The seven loci, PreC, C, Pol, PreS1, PreS2, S, and X are positioned in the reference sequence to create the "reference map".
4 The reference genome sequence for a circular genome needs a device to deal with the fragment to be aligned over the origin of the genome. In this case, a linear reference sequence is thus made by cleaving a pair of circular genomes at the origin and concatenating them together. Using the linear reference sequence, all the entry sequences in the "source" are located correctly on the circular genome by the alignment process. If there are two equivalent hits to the reference sequence, the first is chosen.
5 Divisions are created by the same process as for HCV.

HEV master database

We recently completed making the HEV master database in the same manner as for the HCV master database. In this database, M73218[11] was chosen as the reference sequence, because it is the earliest of all the available entries which have whole genome sequences. The three loci, ORF1, ORF2, and ORF3 in the reference are processed to make three pairs of divisions.

Querying the database

We developed a WWW-based interface to provide easy access to the HVDB. The following three tools are used for viewing and retrieving data via the interface.

Map viewer

This viewer displays "map information" as well as a "reference map" of the master database. In this view, filled boxes that show the loci are on the top, the ruler of the base position on the reference genome follows, and under the ruler are numerous horizontal lines, each of which shows the location of each entry (Fig. 2).

In the case of a circular genome (e.g. HBV), a circular-style map is displayed. In this view, the ruler of the base position, the map of loci, and numerous arcs and rings each of which shows the location of each entry are depicted concentrically (Fig. 3).

Users can employ the viewer to obtain a set of entries that cover a specific locus or a specific region on the genome. Users can choose a locus or a region by clicking

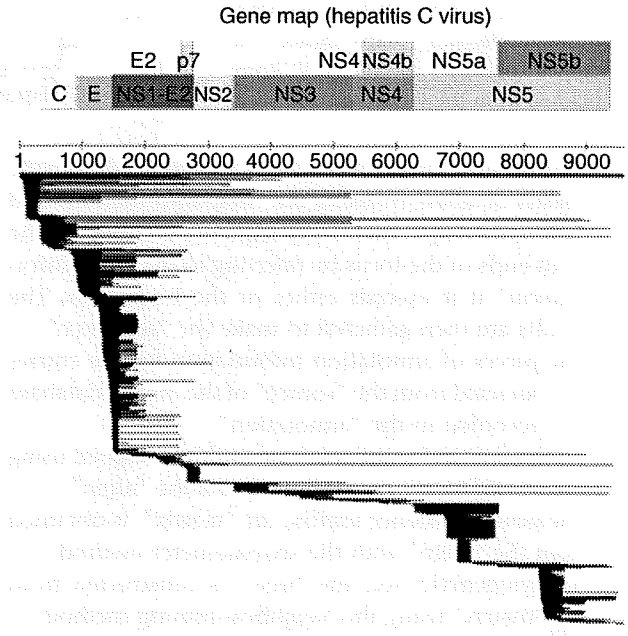

Figure 2 Map viewer for a linear genome. Figure shows the map information of the hepatitis C virus master database.

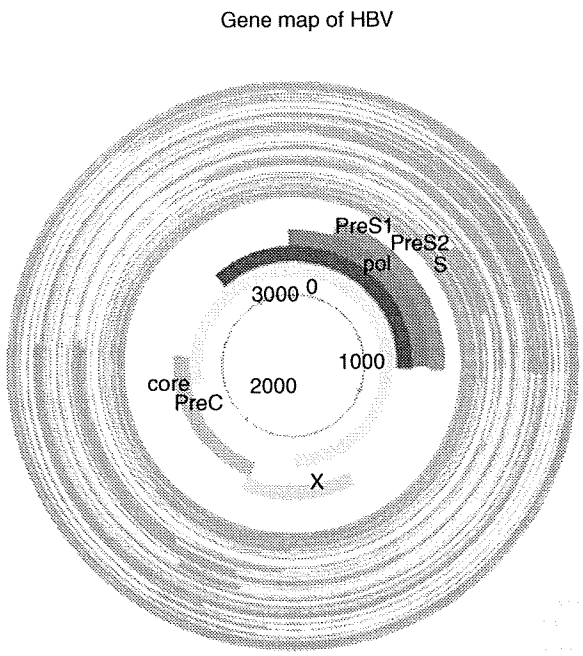

Figure 3 Map viewer for a circular genome. Figure shows the map information of the hepatitis B virus (HBV) master database.

the filled box of a locus, dragging a region on the map, or typing the start and end positions in the fields above the map. Data retrieval is thus performed by selecting the "entry" (annotations), "sequence" (nucleotide sequences), or "peptide" (amino acid sequences) data types. If "sequence" or "peptide" is selected, retrieved sequences are clipped at both ends at the specified region mentioned above. Once the annotations or sequences set is displayed, it can be downloaded by GenBank (annotations) or FASTA (sequences) format by clicking the "download" button at the bottom of the page.

Users can also make a private division from the retrieved entries employing the "division" menu.

Tree viewer

This viewer displays a "tree", or the phylogenetic tree of a division (Fig. 4). If the bootstrap resampling analysis[21] has been applied to the division, the value of each node can be displayed. Users can modify the view by changing the root position of the tree, exchanging two branches under the specified node, or zooming up to the specified portion of the tree. Descriptions of operational taxonomic unit (OTU) (leaves of the tree) names can be changed to data stored as a "tag" in the division, although the INSDC accession numbers are used by default. For example, this function can be used to show the phylogenetic relationship of the genotypes of HCV.

Users can employ the viewer to obtain a set of entries that belong to a specified node or OTU specific branch, which is selected by clicking a node (OTU or a branching point) on the tree. Then, the annotation or sequence information about the entries is displayed by selecting the "entry" (annotation) or "sequence" (nucleotide/amino sequence) data type (Fig. 4), which can be downloaded in the same way as described earlier. Users can also make a private division from the retrieved entries by using the "division" menu.

Align viewer

This viewer displays the "align", or a multiple alignment of all sequences of a division in interleaved format. The first line of each block of the alignment shows a consensus, followed by all the sequences in which only bases different from the consensus are identified as A, C, G, or T (Fig. 5).

Private divisions

Users can make private divisions to do their own analyses, as described above. There are several ways of making these:

1 By extracting partial data from a master database by referring to map information or a phylogenetic tree, as described above.
2 By uploading their own data. FASTA, GenBank, and EMBL format are supported.
3 By copying another division. Both master and private divisions can be copied.
4 By merging two or more private divisions into one. If there are duplicated entries, an editor page is displayed in which users can modify or remove entries.

Users may want to analyze data sets that contain both their own data and publicly available data. In such cases, they obtain a portion of the master database by (1) or (2), upload their own data by (2), and then merge them by (4).

After creating the "private divisions", users are able to execute phylogenetic analyses of them. For this purpose, we provide multiple alignment, genetic distance estimation, and phylogenetic tree construction services with various options. Bootstrap resampling is also available to estimate the confidence of a phylogenetic tree. The results of the analyses can be seen by use of the tree

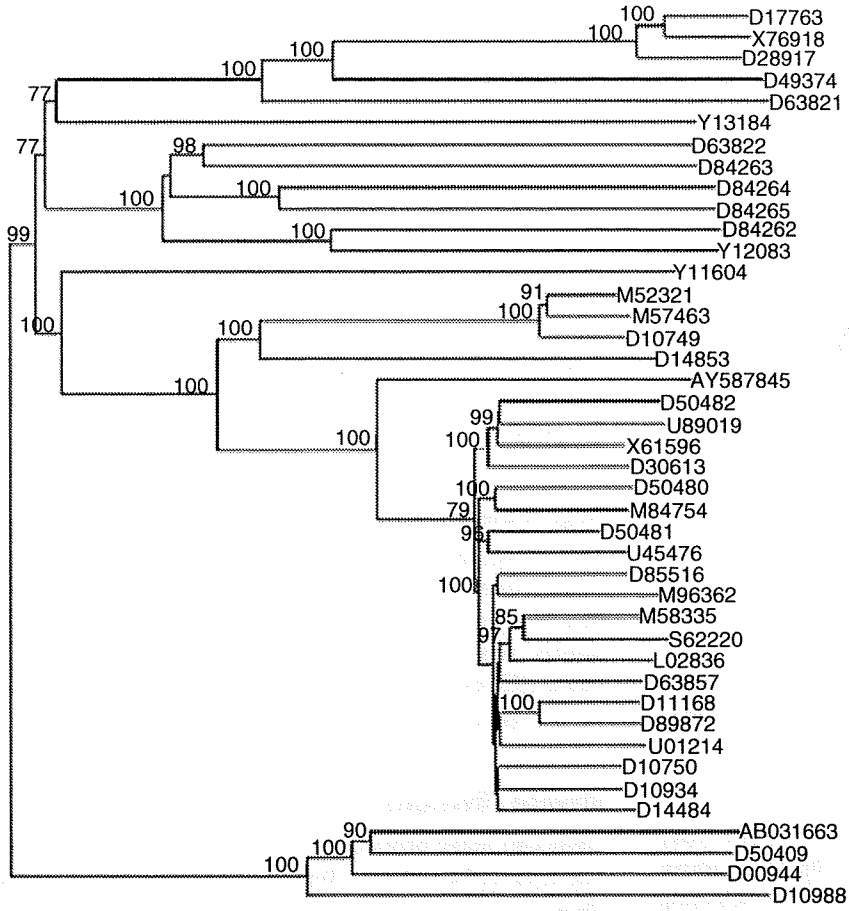

Figure 4 Tree viewer. Numbers on nodes show their bootstrap probabilities.

viewer (phylogenetic tree) or align viewer (multiple alignment).

Before making private divisions, users need to obtain their access accounts free of charge by submission through the WWW page mentioned above. The accounts are used to manage the divisions (who owns which division) in the database system.

HCV genotyping

Estimating the genotypes of sample sequences plays an important role in both epidemiological and phylogenetic studies. We therefore provide a service for automatically processing analyses of HCV sequences. For this purpose, we prepared a subset of the master database by extracting entries that have full genome sequences and whose genotypes are well identified. There are too many entries in some genotypes (e.g. HCV-1b). In such cases, we chose a few entries from them. The resultant subset is called a "typing database".

Users can analyze their own sequences (queries) automatically in nearly the same manner as the process of making divisions in the master database as follows:

1 A query is aligned to the reference sequence and the location of the query on the reference sequence is obtained.
2 All the sequences in the typing database are clipped at the location identified in (1); the resultant fragments are then compiled with the query into a private division.
3 The sequences are multiple aligned, the genetic distances are estimated, and finally a phylogenetic tree is constructed. Bootstrap resampling may then be executed.

Once users finish their analyses, they can visually refer to the phylogenetic relationships of the query and the existing genome types, as well as the location of the query on the reference genome.

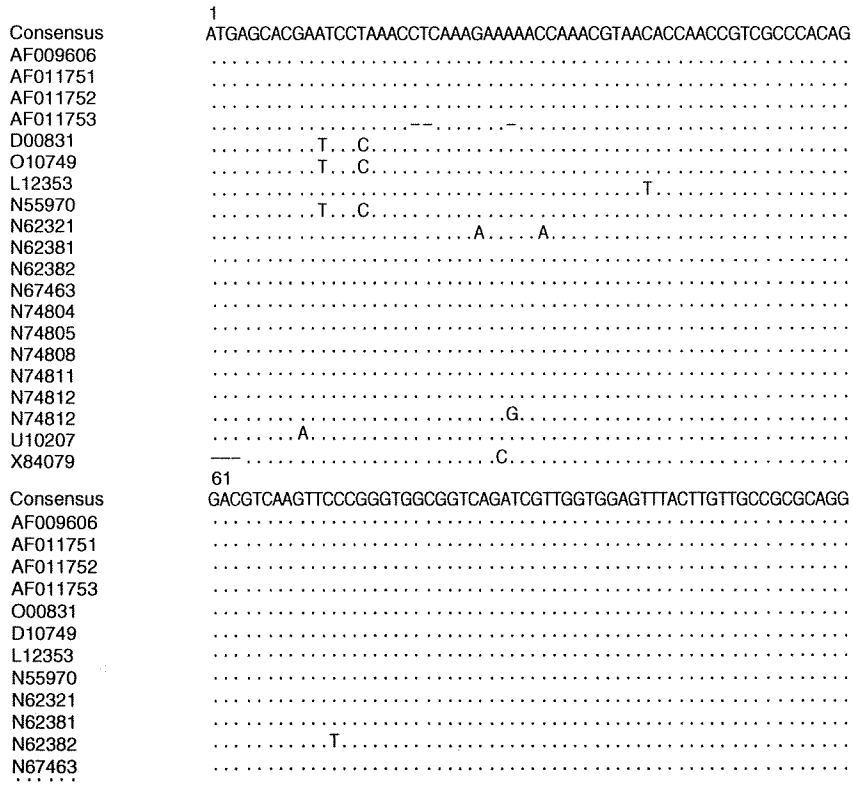

Figure 5 Align viewer. A gap is designated by a hyphen (-) in both consensus and entries. A base that is the same as the consensus sequence is designated by a period (.).

It is noted that the service is applied to "private division" architecture to manage the results, so that users need to have accounts, as described above.

Example of an analysis using the genotyping service

To confirm the effectiveness of HVDB, we analyzed genome sequences using the genotyping service of the database. The HCV strain RF1-2k/1b N687 (accession No. AY587845)[22] is known to be a recombinant, but the breakpoint is not described in the annotation. To find it, we executed the following analysis.

First, five fragments that represent five loci of HCV were manually extracted from the sequence of the entry, AY587845. The base positions of the five fragments are 350–900 (C locus), 950–1450 (E1 locus), 2550–3350 (NS2 locus), 3400–5150 (NS3 locus), and 7600–9300 (NS5b locus). Then, applying each fragment to the genotyping service, a set of its orthologous fragments was selected and a phylogenetic tree was constructed for the set. In this case, five trees were obtained for the five loci (Fig. 6). The five trees are clearly divided into two classes with respect to the phylogenetic position of the query. One class includes the three trees of C, E1, and NS2 loci, and the other contains the two trees of NS3 and NS5 loci. Therefore, together with the information on the locus arrangement on the genome, we can conclude that the breakpoint is located between the NS2 and NS3 loci on the genome (Fig. 6).

RESULTS

THE LATEST RELEASE of the HVDB contains all the public HCV, HBV, and HEV entries in the DDBJ release, as of December 2006. The divisions of the three viruses are listed in Table 2.

HVDB provides the "map information" that can be used to locate a given entry on the reference sequence. Users can retrieve all the entries that cover a specific region by referring to the "map information". Once a user specifies a region, the database system automatically retrieves the entries and clips the sequences that match the region. Users can also retrieve the entries that belong to a specified cluster in a phylogenetic tree.

A division is also used as a unit of analysis for users ("private division"). Users can easily create, remove, or merge divisions and execute their own analyses. From a technical point of view, many data analyzing

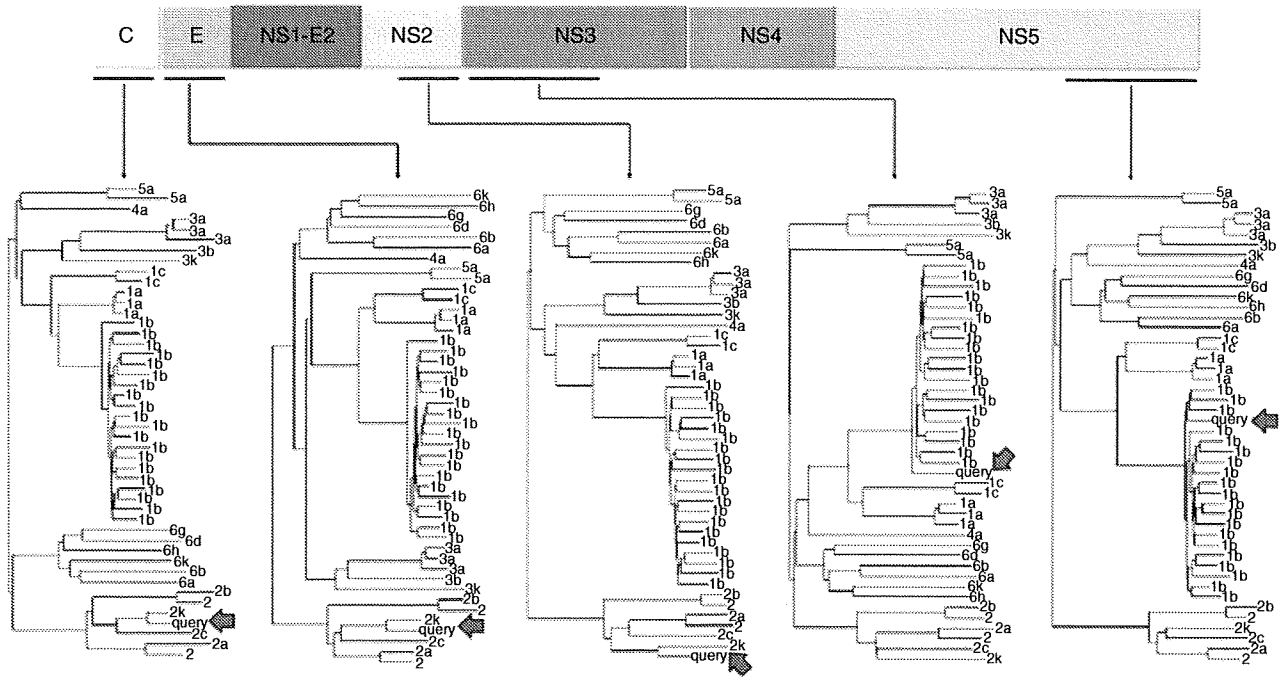

Figure 6 Result of genome typing of the recombinant hepatitis C virus genome: Horizontal lines below the locus map show the regions that were used as five queries. Each phylogenetic tree, which is estimated by genome typing service, shows the relation among the query (block arrows) and fragments whose genome types are well defined.

components and viewers can be shared among the master databases, making it possible to simplify the overall system structure.

Users may need multiple-step procedures in their data analyses. For example, genotyping requires mapping to the genome, the extraction of reference sequences, multiple alignment, genetic distance and phylogenetic tree estimation, and sometimes bootstrap resampling. The genotyping service in the HVDB is designed to eliminate this laborious process. It provides users with an easy-to-use analysis pipeline to estimate a genotype from a sequence fragment specified as a query.

Usage flow

Users can retrieve and analyze data in various ways as shown in Figure 7. Users who want data of the specific locus or region can retrieve them by selecting a region in "map viewer". Users who want data of the specific genotype or phylogenetic cluster can retrieve them by selecting a node in "tree viewer". The retrieved data can also be downloaded as GenBank or FASTA format and then be used by other software beyond the HVDB.

When users do their own analyses of the retrieved data, the data are compiled into a private division by selecting a corresponding option in the viewers. When users want to analyze their own data with published ones in the HVDB, they may upload their data and then merge the data into a private division retrieved from the HVDB as described above. Once private divisions are made, they can be phylogenetically analyzed by various algorithms and parameters, and the result prepared in the HVDB can be viewed by viewers.

DISCUSSION

RECENTLY, TWO HCV sequence databases have been opened to the public; one in the USA[23] and another in France.[24] The US database was constructed by collecting HCV sequences from INSDC and adding some information about patients to them. The French database is made up of HCV sequences of INSDC, the SRS (Sequence Retrieval System) keyword query tool, and sequence analysis tools. While neither database includes data of hepatitis virus other than HCV, the HVDB includes data of HBV and HEV, as well as HCV. The other property of the HVDB is that it has the precomputed "map information" and phylogenetic trees, which the US and French databases do not. The

Table 2 Number of entries that are contained in each master database. Each number shows the number of entries that cover whole area of the corresponding locus

Virus	Total entries	Locus	Nucleic acids	Amino acids
HCV	44 709	Whole genome	178	
		C	901	865
		E1	1456	1377
		NS1/E2	376	359
		E2	414	394
		p7	462	390
		NS2	370	351
		NS3	314	305
		NS4	284	275
		NS4a	545	472
		NS4b	284	275
		NS5	178	173
		NS5a	812	797
		NS5b	180	175
HBV	11 895	Whole genome	1009	
		P	1015	793
		X	1239	954
		C	1736	1444
		pre-C	1534	860
		S	2662	2488
		pre-S1	1033	635
		pre-S2	1287	893
HEV	1 611	Whole genome	77	
		ORF1	77	76
		ORF2	89	82
		ORF3	87	84

HBV, hepatitis B virus; HCV, hepatitis C virus; HEV, hepatitis E virus, ORF, open reading frame.

information enables users to obtain data of the specific locus or genotype very easily. As a functional property, the HVDB provides the genotyping service. Although the US database has a similar function, users have to process all the steps of analyses manually. On the other hand, the genotyping service of the HVDB can process the analyses automatically once users set their queries and parameters.

Enhancement plan

There are several other hepatitis-related viruses; the major ones are hepatitis A virus (HAV) and hepatitis D virus (HDV). To enrich HVDB, we are now preparing the HAV and HDV master databases in the same manner as HCV (linear genome) or HBV (circular genome).

So far the genotyping service can only be used for HCV sequences, since the strict classifications of other viruses are still arguable. We will add the service of other viruses once a consensus of classification of each virus is reached.

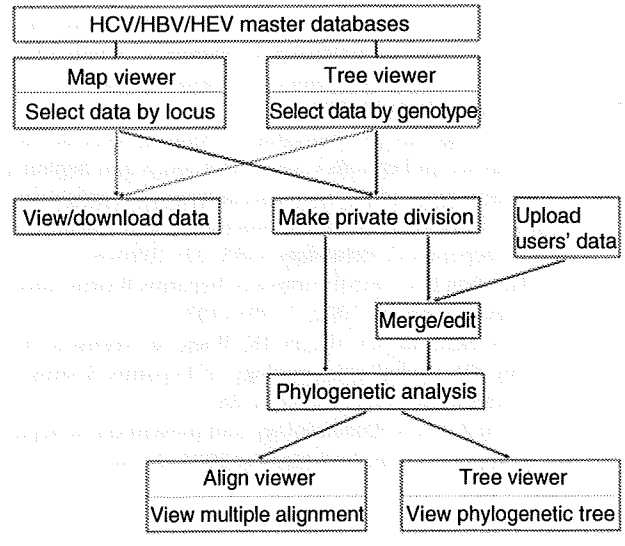

Figure 7 Flow chart of the hepatitis virus database usage. Dashed arrows show steps of data retrieval, and the solid arrows show steps of data analyses. HBV, hepatitis B virus; HCV, hepatitis C virus; HEV, hepatitis E virus.