Identification of hepatitis B virus indigenous to chimpanzees

Xiaolei Hu*, Harold S. Margolis*, Robert H. Purcell†, James Ebert*, and Betty H. Robertson‡

*Hepatitis Branch A33, Division of Viral and Rickettsial Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Atlanta GA 30333; and †Hepatitis Viruses Section, Building 7, Room 201, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892

Contributed by Robert H. Purcell, December 9, 1999

Hepatitis B virus (HBV) and related viruses, classified in the Hepadnaviridae family, are found in a wide variety of mammals and birds. Although the chimpanzee has been the primary experimental model of HBV infection, this species has not been considered a natural host for the virus. Retrospective analysis of 13 predominantly wild-caught chimpanzees with chronic HBV infection identified a unique chimpanzee HBV strain in 11 animals. Nucleotide and derived amino acid analysis of the complete HBV genome and the gene coding for the hepatitis B surface antigen (S gene) identified sequence patterns that could be used to reliably identify chimpanzee HBV. This analysis indicated that chimpanzee HBV is distinct from known human HBV genotypes and is closely related to HBVs previously isolated from non-human primates.

DNA Extraction, Amplification, and Sequence Determination and Analysis. HBV DNA was extracted from serum specimens by using a modification of the Master Pure procedure (TaqCentre Technologies, Madison, WI) in which 50 μl of serum was mixed with an equal amount of 2× T & C lysis solution containing 333 μg/ml proteinase K and incubated at 65°C for 1 hr. After proteinase K digestion, samples were placed on ice for 5 min, 50 μl of MPC protein precipitation reagent was added, the mixture was vortexed and centrifuged at 10,000 × g for 10 min at 4°C, and the supernatant was transferred to a new Eppendorf tube. DNA was precipitated in 500 μl of ice-cold isopropanol by centrifugation at 10,000 × g for 10 min at 4°C. The DNA pellet was rinsed twice with 500 μl of ice-cold 75% ethanol, was air-dried, and was resuspended in 20 μl of water.

The complete HBV genome from CH109 and CH926 was PCR amplified with eight primer pairs selected from conserved genome regions such that the resulting fragments overlapped at ampiclons by at least two-thirds of their length. The primers (shown below from 5’ to 3’ direction) were CHB2853P TACCACATATTCTTGGAACACA; CHB2853N TGTTCCAAAGAGATTGGTTGA; CHB58CP CCTGCTGTGCTCCATAGTT; CHB58CN GAACCTGGACCACCAACCAG; CHB409P CATCTCCTGCTCCTCACCT; CHB409N AGATGAGACCCATCATCT; CHB730P AGTGCCATTGTTTCCAGTGTT; CHB730N AGTGCCATTGTTTCCAGTGTT; CHB1101P CTCCGCAACCTTACAAAGCCTTTTC; CHB1101N GAAAGGCCTTTGAAAGTGGCGAG; CHB1450P TACGTCCCGTCGCGCTGAATC; CHB1450N GATCCCGCCCGAGCAG; CHB1860P GGTCTTTCAAGCCCTCAAGCTG; CHB1860N CAAGCGATGGTGCGCTACAG; CHB2440P GCCCGGTCGCCAGAGAGATCTCAAA; CHB2440N TTGAGATCTTCTTGGCGAGCGGC. The entire 20 μl of resuspended HBV viral DNA solution was added to a mixture containing 20 μl of 5× Invitrogen Buffer, 4 μl of 10 mM dNTPs (Invitrogen), 1 μl of 10 μM positive sense primer (HBVxxxxP), 1 μl of 10 μM negative sense primer (HBVxxxxN), 1.6 μl of TaqStart Dilution Buffer (CLONTECH), 1.6 μl of Taq polymerase antibody (CLONTECH), 2 units of Taq polymerase (Boehringer Mannheim), and 51.6 μl of water. PCR amplification was performed by using a modification of a previously described method (17) with cycling conditions as follows: denaturation at 94°C for 3 min; 45 cycles at 94°C for 45 sec and 55°C for 45 sec; ramping to 72°C within 1 min and extension at 72°C for 3 min; and a final extension cycle at 72°C for 7 min. PCR-amplified products were examined by agarose gel electrophoresis and then were purified by using a QIAquick Purification Kit (Qiagen, Valencia, CA). Cycle sequencing was performed by using dRhodamine dideoxynucleotides (Perkin-Elmer Corporation, CA) with 20 μl of PCR products.

Materials and Methods

Chronic Carrier Chimpanzees. The following wild-caught animals were included in the study (identification number, sex, estimated date of birth, institution): NIH28, male (M), unknown, National Institutes of Health (NIH); NIH29, female (F), 1970, NIH; NIH39, F, 1968, NIH; NIH40, M, 1969, NIH; NIH44, ?, unknown, NIH; NIH56, M, 1974, NIH; NIH57, F, 1973, NIH; NIH814, 1970, F, NIH; NIH821, M, 1969, NIH; CH1109, 1958, M, Centers for Disease Control and Prevention; CH926, M, 1974, Centers for Disease Control and Prevention. In addition, the following animals born in captivity were included: NIH800, F, 1973, NIH; NIH904, M, 1975, NIH.

Abbreviations: HBV, hepatitis B virus; gnty, genotype; HBsAg, hepatitis B surface antigen; M, male; F, female; NIH, National Institutes of Health.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF22311-AF22323).

*To whom reprint requests should be addressed. E-mail: bjr1@cdc.gov.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.
damine terminators, and the products were electrophoresed on a Perkin–Elmer 377 sequencer (PE Biosystems, Foster City, CA). Archived serum specimens from chimpanzees NIH28, NIH29, NIH39, NIH40, NIH44, NIH56, NIH57, NIH800, NIH814, NIH821, and NIN904 were used to amplify and sequence the S gene as described above by using primers HBV2853P or HBV58P and 1101N.

A consensus nucleic acid and amino acid sequence was generated for each of the six human HBV genotypes by using the GCG PRETTY program (Genetics Computer Group, Madison, WI). These consensus sequences, designated genotype a (gntya) through gntyf, were derived by using the individual sequences (Fig. 1), which are representative of complete genome sequences available in GenBank. In addition, a consensus gibbon HBV sequence was derived by the same approach.

Sequence analysis and alignments were performed by using the PILEUP program within the University of Wisconsin GCG SEQUENCE ANALYSIS computer software package (Genetics Computer Group). Genetic distances were calculated by using the Jukes and Cantor correction within DNADIST of PHYLIP 3.5C (18), and the results are graphically illustrated as a neighbor-joining tree. Bootstrap values representing 1,000 replicates were determined by using SEQBOOT, DNADIST, NEIGHBOR, and CONSENSE in the PHYLIP package.

Results

The complete genome was sequenced from two animals, and the gene coding for HBsAg was sequenced from 11 additional animals. All animals were found to be HBsAg positive upon arrival at either the Centers for Disease Control and Prevention or the National Institutes of Health (NIH), where they were used for experimental transmission of infections with hepatitis viruses. The serum samples tested in this study were obtained before experimental inoculation of material containing hepatitis viruses. Although the precise date and place of birth for most of these animals were unknown, all but two were known to have been captured in Africa as infants before chimpanzees were listed as a threatened species. Two chimpanzees (CH800 and CH904) were born and reared in captivity; their father was HBsAg-positive, but the HBsAg status of the mother was unknown.

The complete HBV genome was amplified and sequenced from chimpanzees CH109 and CH926. The complete HBV genome from each animal was composed of 3,182 nucleotides and had a genetic organization identical to human HBV, with four ORFs encoding the S, X, C, and P genes and a characteristic 11-bp direct repeat motif located at positions 1827–1837 (DR1) and 1593–1603 (DR2). These two isolates had a 93.7% identity with each other and a 93–98% identity with HBV from a wild-caught chimpanzee at the London zoo (D00220, Table 1). Chimpanzee HBV isolates had a 87.5–91.2% nucleotide sequence identity with consensus sequences of each of the six human HBV genotypes (19, 20); this degree of variation is similar to that observed between the human HBV genotypes (Table 1).

The phylogenetic relationship of chimpanzee HBV to the six respective human HBV genotypes and HBV isolates from other non-human primates is depicted in Fig. 1A. Phylogenetic comparison of complete HBV genome nucleotide sequences of chimpanzees, humans, and other non-human primates. The respective human genotypes are indicated by the letters A–F on the corresponding branches. GenBank accession numbers for the representative sequences used in the phylogenetic analyses were as follows: Genotype A: V00866, X51970, X70185, S50225, Z35717, M57663; Genotype B: U95551, J02203, X65257, X85254, X97848, X59759, X02496, Y07587, a043593, Z35716, X80925, X97849, X72702, M32138, L27106, X65259, X68292; Genotype C: X75655, X76665, D28880, D00630, D23680, D23681, D23684, M38366, X01587, X00867, X04615, D23682, D23686, X22939, D25017, D25019, D25020, D16665, D25089, D16666, D16667, D12980, M12906; Genotype D: D00521, D00522, D00329, D23679, D23678, D00331, M54923, D00330, X97851, X97850; Genotype E: X75657, X75664; Genotype F: X75658, X75663, X69798; Gibbon: aj131568, aj131574, aj131560; Chimpanzees: D00220 (B) Phylogenetic comparison of the S gene region (nucleotides 157–673). It includes the additional chimpanzees with chronic HBV infection. The respective human genotypes are indicated by the letters A–F. The two chimpanzee isolates apparently infected with human HBV are indicated by numbers in the boxes.

Fig. 1. (A) Phylogenetic comparison of complete HBV genome nucleotide sequences of chimpanzees, humans, and other non-human primates. The respective human genotypes are indicated by the letters A–F on the corresponding branches. GenBank accession numbers for the representative sequences used in the phylogenetic analyses were as follows: Genotype A: V00866, X51970, X70185, S50225, Z35717, M57663; Genotype B: U95551, J02203, X65257, X85254, X97848, X59759, X02496, Y07587, a043593, Z35716, X80925, X97849, X72702, M32138, L27106, X65259, X68292; Genotype C: X75655, X76665, D28880, D00630, D23680, D23681, D23684, M38366, X01587, X00867, X04615, D23682, D23686, X22939, D25017, D25019, D25020, D16665, D25089, D16666, D16667, D12980, M12906; Genotype D: D00521, D00522, D00329, D23679, D23678, D00331, M54923, D00330, X97851, X97850; Genotype E: X75657, X75664; Genotype F: X75658, X75663, X69798; Gibbon: aj131568, aj131574, aj131560; Chimpanzees: D00220. (B) Phylogenetic comparison of the S gene region (nucleotides 157–673). It includes the additional chimpanzees with chronic HBV infection. The respective human genotypes are indicated by the letters A–F. The two chimpanzee isolates apparently infected with human HBV are indicated by numbers in the boxes.

sequence analysis and alignments were performed by using the PILEUP program within the University of Wisconsin GCG SEQUENCE ANALYSIS computer software package (Genetics Computer Group). Genetic distances were calculated by using the Jukes and Cantor correction within DNADIST of PHYLIP 3.5C (18), and the results are graphically illustrated as a neighbor-joining tree. Boot strap values representing 1,000 replicates were determined by using SEQBOOT, DNADIST, NEIGHBOR, and CONSENSE in the PHYLIP package.
non-human primates is shown in Fig. 1A. Each non-human primate HBV clusters on a distinct phylogenetic branch, and, with the exception of the wooly monkey isolate, all of the non-human primate isolates originate from a node that is distinct from any human HBV genotype (Fig. 1A).

When the derived amino acid sequence from the three complete chimpanzee genomes were compared with each of the human HBV genotypes and the other non-human primate genome sequences, certain patterns appeared to identify certain non-human primate HBVs (Figs. 2 and 3). These included (i) a glutamic acid at position 16 of the preS region in chimpanzee, gorilla, and gibbon isolates (Fig. 2, preS); (ii) three amino acid changes (L133-IY134 and A177) in the S gene of the chimpanzee, gorilla, and gibbon HBV isolates (Fig. 2, HBsAg); (iii) a glutamine at position 113 coding for hepatitis B core antigen of the chimpanzee and gorilla HBVs that replaced a leucine in human HBV (Fig. 2, Core); and (iv) an 11-amino acid deletion in the polymerase (P) gene of the non-human primate isolates that was also found in human genotype D (Fig. 3). In addition, a precore stop codon (TTG → TAG) at position 28 (Fig. 2, Core) was found in two chimpanzees; a leucine was present in the remaining non-human primate isolates, in contrast to the tryptophan found in all human isolates. Three amino acid changes in the X gene product of chimpanzees (T11, K107, T110; Fig. 2) were found, two of which were also found in the gorilla sequence (K107, T110; Fig. 2) whereas multiple changes within the preS1 and preS2 overlap region (data not shown) were also present.

Nucleotide sequences were obtained from the HBsAg coding region (S gene) on an additional 11 chimpanzees with chronic HBV infection. The isolates from nine animals were closely related to those of CH109, CH926, and D00220 (Fig. 1B) whereas two chimpanzees (Fig. 1B, NIH28 and NIH44) had isolates related to human HBV genotypes A and C, respectively. The derived amino acid sequences indicated that the chimpanzee isolates would be classified as subtype adw (21) whereas the gibbon and woolly monkey isolates would be classified as subtype ayw. The translated amino acids from the S gene sequences from the additional chimpanzee HBV strains also contained the three

| Table 1. Percent nucleotide identity between chimpanzee hepatitis B virus strains and consensus sequences for human hepatitis B virus genotypes |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|                             | CH926 | D00220 | gntya | gntyb | gntyc | gntyd | gntye | gntyf |
| CH109                      | 93.7  | 93.8  | 89.8  | 89.5  | 89.5  | 89.8  | 90.6  | 87.5  |
| CH926                      | 98.6  | 90.9  | 90.8  | 90.6  | 90.7  | 91.2  | 88.1  |       |
| D00220                     | 90.7  | 90.4  | 90.5  | 90.6  | 91.0  | 88.1  |       |       |
| Gnty A                     | 92.4  | 91.9  | 91.2  | 90.6  | 86.8  |       |       |       |
| Gnty B                     |       |       | 92.1  | 91.0  | 89.9  | 87.5  |       |       |
| Gnty C                     |       |       | 91.0  | 90.3  | 87.1  |       |       |       |
| Gnty D                     |       |       |       | 93.3  | 87.4  |       |       |       |
| Gnty E                     |       |       |       |       | 86.8  |       |       |       |

Fig. 2. Comparison of regions from Pre S, HBsAg, Core, and X coding regions with translated amino acid sequence differences. The following sequences are shown: (i) individual sequences from chimpanzees CH109, CH926, and D00220; (ii) gorilla HBV (aj131567); (iii) a consensus sequence of gibbon HBV (gib); (iv) a consensus sequence for each genotype (gntya, gntyb, gntyc, gntyd, gntye, and gntyf); and (v) woolly monkey HBV (af046996) (15). Dashes indicate conserved amino acids while changed positions are indicated in lowercase letters. A master sequence based on this comparison is depicted on the bottom line in upper-case letters. Shaded regions are discussed in the text. A comparison of the complete amino acid sequences is published as supplemental data on the PNAS web site, www.pnas.org.
isolation from other primates, the chimpanzee HBV has unique features that are not shared with human HBV isolates. A complete sequence of HBV from orangutans with chronic HBV infection has been obtained (S. Mishiro, personal communication). Similar sequence information has been obtained in another study (ii) the gibbon and chimpanzee nucleic acid sequences are more closely related to each other than to human genotypes, and these two types of great apes are widely separated geographically (Africa and Asia). (iii) the human genotypes found in Africa and Asia differ—Asia is predominantly B and C, and Africa is predominantly A and E; and (iii) the possible modes of virus transmission between non-human primates and humans are unclear because blood-borne viruses are primarily transmitted by sex, birth, or the direct inoculation of blood.

The extent to which chronic HBV infection occurs in free living non-human primate populations is unknown. However, it is unlikely that the presence of HBVs indigenous to non-human primates would have important public health implications. Because gibbon and woolly monkey HBV strains have been shown to produce infection in chimpanzees (15, 22), and human HBV causes disease in chimpanzees, the potential for zoonotic disease transmission exists where blood or body fluid exposure is common. Such scenarios could include chronically infected animals kept as family pets, close contact with caretakers, or in situations in which chimpanzees are slaughtered and used as bushmeat. Human HBV infection can be prevented by immunization with hepatitis B vaccine, which contains HBsAg to induce protective immunity through the amino acids that comprise the a determinant (underlined in Fig. 2). Among the Old World non-human primate HBVs, all contain the immunodominant glycine at position 145 of the a determinant, and other amino acid changes are limited, which indicates that available vaccines would probably cluster with the gibbon, chimpanzee, and gorilla HBV sequences.

**Discussion**

The HBVs isolated from 11 of the 13 chimpanzees with chronic infection appear to be genetically distinct from known human HBV genotypes and appear to represent a virus that produces an infection indigenous to chimpanzees. The two additional animals with chronic HBV infection had human HBV isolates, a finding that could be explained by the practice of injecting wild-caught infant animals with human serum. When compared with HBVs isolated from other primates, the chimpanzee HBV has unique nucleotide and amino acid differences throughout the entire genome and the S gene that allow the rapid and precise identification of this strain. In addition, our analysis confirms the existence of indigenous non-human primate HBV strains in both Old World great apes and New World monkeys.

The phylogenetic evolution of the non-human primate HBVs with respect to human HBV is not clear. The woolly monkey HBV and human genotype F appear to be related with respect to their phylogenetic and geographic (South American) proximity. However, although such phylogenetic patterns and geographic proximity may suggest possible zoonotic transmission of HBVs, a number of biologic and epidemiologic inconsistencies argue for independent evolution within various geographically separated species or populations. These inconsistencies include the following: (i) the gibbon and chimpanzee nucleic acid sequences are more closely related to each other than to human genotypes, and these two types of great apes are widely separated geographically (Africa and Asia); (ii) the human genotypes found in Africa and Asia differ—Asia is predominantly B and C, and Africa is predominantly A and E; and (iii) the possible modes of virus transmission between non-human primates and humans are unclear because blood-borne viruses are primarily transmitted by sex, birth, or the direct inoculation of blood.

Note Added in Proof.

Similar sequence information has been obtained in another study (S. Mishiro, personal communication) and has been deposited under GenBank accession numbers AB032431–AB032433. For the purposes of HBV nomenclature, we have proposed that the chimpanzee HBV be abbreviated as ChHBV.