Characteristics of Hepatitis B Virus Isolates of Genotype G and Their Phylogenetic Differences from the Other Six Genotypes (A through F)

Hideaki Kato,1,2 Etsumi Orito,1 Robert G. Gish,3 Fuminaka Sugachi,1,2 Seiji Suzuki,1,2 Ryujo Ueda,1 Yuzo Miyakawa,4 and Masashi Mizokami2*

Department of Internal Medicine and Molecular Science1 and Department of Clinical Molecular Informative Medicine,2 Nagoya City University Graduate School of Medical Sciences, Nagoya, and Miyakawa Memorial Research Foundation, Tokyo,3 Japan, and Department of Hepatology and Gastroenterology, California Pacific Medical Center, San Francisco, California3

Received 2 January 2002/Accepted 14 March 2002

Eight hepatitis B virus (HBV) isolates of genotype G were recovered from patients and sequenced over the entire genome. Six of them had a genomic length of 3,248 bp and two had genomic lengths of 3,239 bp (USG15) and 3,113 bp (USG18) due to deletions. The 10 HBV/G isolates, including the 8 sequenced isolates as well as the original isolate (AF160501) and another isolate (B1-89), had a close sequence homology of 99.3 to 99.8% among themselves (excluding USG18 with a long deletion) but of <88.7% to any of the 68 HBV isolates of the other six genotypes with the full-length sequence known. The eight HBV/G isolates possessed an insertion of 36 bp in the core gene and two stop codons in the precore region, as did the AF160501 and B1-89 isolates. The 10 HBV/G isolates clustered on a branch separate from those bearing the other six genotypes (A through F) in the phylogenetic tree constructed from full-length sequences of 78 HBV isolates as well as in those constructed from the core, polymerase, X, and envelope genes. Despite two stop codons in the precore region that prohibited the translation of the HBV e antigen (HBeAg), all of the eight patients with HBV/G infection possessed the HBeAg in serum. By restriction fragment length polymorphism of the surface gene, all of the eight patients were found to be coinfected with HBV of genotype A (HBV/A), which would be responsible for the expression of HBeAg in them. It is worthy of examination to determine how coinfection occurs and whether HBV/G needs HBV/A for replication.

Hepatitis B virus (HBV) persistently infects 350 million people worldwide and can induce a spectrum of acute and chronic liver diseases (12). Based on sequence divergence in the entire genome of >8%, HBV has been classified into six genotypes named with capital letters A through F (A–F) (19, 22). The six major genotypes of HBV have distinct geographic distributions (14, 15) and are associated with different clinical diseases (8, 16, 24).

Recently, a seventh HBV genotype was proposed for an HBV isolate (AF160501) recovered in France and named G, based on a sequence divergence of >11.8% from HBV isolates of the other six genotypes (26). This HBV isolate of genotype G (HBV/G) has a genomic length of 3,248 bp, a little longer than HBV isolates of the other six genotypes, which are composed of 3,182 to 3,221 bp. The longer length of the HBV/G genome is attributed to an insertion of 36 bp at codon 2 of the core (C) gene (26). Other remarkable features of HBV/G are two stop codons at positions 2 and 28 of the pre-C region that prohibit the translation of the HBV e antigen (HBeAg) (2).

For the specific detection of HBV/G, eight sera were identified, by PCR with heminested primers deduced from the 36-bp insertion, that contained HBV of this genotype. The entire nucleotide sequences were determined for the eight HBV/G isolates recovered from sera, and along with the original HBV/G isolate (AF160501) and another HBV isolate (B1-89), the genotype of which is deduced to be G (11), the HBV/G genomes were compared to one another and against 68 HBV isolates of the major six genotypes (A, 9 isolates; B, 16 isolates; C, 17 isolates; D, 18 isolates; E, 2 isolates; and F, 6 isolates). Remarkably, all of the eight patients from whom HBV/G isolates were recovered were found to be coinfected with HBV of genotype A (HBV/A).

MATERIALS AND METHODS

Patients. Sera were obtained from patients who were persistently infected with HBV and taken care of in the Second Department of Medicine, Nagoya City University Medical School, Nagoya, Japan, and those in the Hepatology and Gastroenterology Division of the California Medical Center at San Francisco. There were eight patients from San Francisco who were found to be infected with HBV/G by PCR with specific primers that can detect 10 copies of HBV/G per test (10); HBV/G was not detected in any patients infected with HBV from Japan. The eight sera were tested for HBV markers, and HBV/G isolates in them were sequenced over the entire genome. The study design was approved by the Ethics Committee of each institution, and informed consent was obtained from each patient.

Determination of HBV/G. Nucleic acids were extracted from serum samples (100 μl) with Smitest EX R&D (Genome Science, Fukushima, Japan) in accordance with the recommendation from the manufacturer and subjected to PCR with heminested primers designed from the 36-bp insertion in the C gene of HBV/G genomes (10). In brief, extracted nucleic acids were subjected to the first round of PCR for 40 cycles with BHHKF1 (sense, 5′-ACG GGG GCC ACC TCT CTT TAC-3′; nucleotides [nt] 1519 to 1539) and BHHKR2 that involved the presence of AmpliTaq Gold (Applied Biosystems, Foster City, Calif.). The second round of PCR was performed for 40 cycles on the product of the first-round PCR with BHHKF2 (sense, 5′-GCA CTT CGT TAC TCT GCA-3′, nt 1581 to 1601) and BHHKR2. Then the products were examined for fragments of 357 bp.

* Corresponding author. Mailing address: Department of Clinical Molecular Informative Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya 467-8601, Japan. Phone: 81-52-853-8292. Fax: 81-52-842-0021. E-mail: mizokami@med.nagoya-cu.ac.jp.

Downloaded from http://jvi.asm.org/ on March 11, 2021 by guest
**RESULTS**

Eight HBV/G isolates. The entire nucleotide sequences of the eight HBV/G isolates have been deposited in the DDBJ/GenBank/EMBL database. They were compared to one another and against the sequences of two other HBV/G strains, i.e., AB160501 (26) and B1-89 (11). Six of the eight HBV/G isolates examined had a genomic length of 3,248 bp, which is identical to that of AB160501 and B1-89. USG15 had a shorter length of 3,229 bp because of a deletion of 9 bp in the pre-S2 region, and USG18 had a length of 3,113 bp due to a deletion of 87 bp in the pre-S1 and pre-S2 regions as well as a deletion of 48 bp in the C gene, together accounting for 135 bp. As was the case for AB160501 and B1-89, the eight HBV/G isolates had 29 amino acids (aa) encoded by the pre-C region, 195 aa encoded by the C gene, 842 aa encoded by the polymerase (P) gene (839 aa for USG15 and 805 aa for USG18), 154 aa encoded by the X gene, and 118 aa encoded by the pre-S1 region (99 aa for USG18). The subtype of the HBsAg was deduced to be adw in all of the eight HBV/G isolates by the presence of lysine residues at positions 122 and 160 (21); it was also adw in isolates AB160501 and B1-89 (11, 26). The YMDD motif that is prone to mutation during lamivudine therapy (3, 13) was identified at aa 548 to 551 (aa 545 to 548 in USG15 and aa 511 to 514 in USG18), with methionine at position 549 (position 546 in USG15 and position 512 in USG18).

Including the two reported isolates (AB160501 and B1-89), a total of 10 HBV/G isolates were compared to one another. They were 99.3% (85.3% when USG18, with a long deletion, was included) to 99.8% homologous within the full-length sequence. The 10 HBV/G isolates were examined for homology to the 68 HBV isolates of six distinct genotypes (A–F), the full-length sequences of which are deposited in the database (Table 2). They showed sequence homologies of 83.4 to 88.7% in the entire genome, 80.6 to 90.0% in the C gene, 73.1 to 90.8% in the P gene, 80.9 to 87.3% in the X gene, 78.1 to 92.0% in the large S gene (pre-S1 and pre-S2 regions included), and 88.9 to 97.5% in the S gene.

Phylogenetic differences between 10 HBV/G isolates and 68
HBV isolates of the other six genotypes. Phylogenetic trees were constructed based on the entire genome of 78 HBV isolates (including the 10 HBV/G isolates) as well as individual genes (C, X, P, large S, and S) (Fig. 1a through f). The 10 HBV/G isolates clustered on a branch separate from those bearing the other six genotypes in all six of the phylogenetic trees. Of the 10 HBV/G isolates, all except USG18 were particularly close to each other, with a homology of 99.3 (excluding USG18 with a long deletion) to 99.6% in the entire genomic sequence (Fig. 1a). A close relatedness of the nine HBV/G isolates (USG18 excluded) was maintained within the C, P, X, large S, and S genes as well (Fig. 1b through f). The sequence of the C gene is available for an additional HBV/G isolate (adw2 variant) (27), and it was most divergent among the 11 C gene sequences of HBV/G isolates (Fig. 1b). HBV isolates of genotypes B and C did not cluster on two separate branches in the phylogenetic analysis of the C gene, unlike in phylogenetic analyses of the entire genome or of the other genes (Fig. 1b). Notably, a clear branching of the seven genotypes (A–F and G) in the entire genomic sequence (Fig. 1a) was reproduced only in the P gene (Fig. 1c). Primary and secondary nodes distinct from those in the phylogenetic tree of the entire genome were obvious in the trees for the C, X, and S genes (Fig. 1b, d, and f).

Insertion and deletion of nucleotides in the C gene and pre-S1 region as well as two stop codons in the pre-C region. The eight HBV/G isolates possessed an insertion of 36 nt at the fifth nucleotide in the C gene (Fig. 2a) as in the original HBV/G isolate (AF160501) (26) and B1-89 (11). Hence the 36-bp insertion would be the hallmark of genotype G, distinguishing it from the other six genotypes. Likewise, a 2-aa deletion at the carboxy terminus of the product of the C gene was preserved in all 10 of the HBV/G isolates (Fig. 2) as in any HBV isolates except for HBV/A isolates.

The two stop codons at positions 2 and 28 in the pre-C region of the original HBV/G isolate (AF160501) (26) were invariably preserved in the eight HBV/G isolates (Fig. 2). The stop codon at position 2 (TAA) is converted to the codon for glutamine (CAA) in B1-89 (11) while the stop codon at position 28 (TAG) is retained in it.

The 1-aa deletion at position 3 in the pre-S1 region in the original HBV/G isolate (AF160501) (26) was possessed in common by the eight HBV/G isolates examined and B1-89 (11); it is shared by HBV isolates of genotype E (19).

HBsAg detected in sera from eight patients who were coinfecting with HBV/G and HBV/A. The eight patients in San Francisco from whom HBV/G isolates were recovered were predominantly Caucasian and male (six patients were male and six patients were Caucasian). They were 27 to 49 years of age. Except for one patient, from whom USG2638 was recovered, they all had elevated levels of alanine aminotransferase in the serum (>50 U/liter). Remarkably, although the HBV/G isolates had two stop codons in the pre-C region, HBeAg was detected in sera from all eight of the patients (Fig. 2).

For evaluating the possibility of coinfection with HBV of the other genotypes with an HBeAg-positive phenotype, HBV DNA samples from sera of the eight patients were genotyped by restriction fragment length polymorphism of the S gene (18). The pattern of HBV/A was detected in all of them; it was distinguished from that of HBV/G (see Materials and Methods). Hence, HBV/A coinfecting the patients with HBV/G would have been coding for HBeAg and responsible for serum HBeAg in all of them.

**DISCUSSION**

A seventh genotype of HBV (G) has been proposed for a French HBV isolate named AF160501 (26) based on a sequence divergence of >8.0% from the entire genome of HBV isolates of the six major genotypes (A–F) (19, 22). A search for sequence homology in the DNA database identified another HBV/G isolate (B1-89) previously reported from France (11), albeit a particular genotype was not assigned for it. In order to confirm the genotype G as a separate entity from the six major genotypes (19, 22) and characterize it virologically, the entire genomes of eight additional HBV/G isolates were sequenced and features specific to this genotype were sorted out.

In addition to the two documented isolates (AF160501 and B1-89), eight HBV/G isolates were compared with 68 HBV isolates of the major six genotypes (A–F), the sequences of which are retrievable from DNA databases. The 10 HBV/G isolates were >99.3% (>85.7% when USG18 with a long deletion was included) homologous along the entire genomic sequence but differed from 68 HBV isolates of the other genotypes by >11.3%, which cleared the threshold of >8% separating HBV genotypes (19, 22). Taken along with the two previously reported results (11, 26), therefore, the eight HBV/G isolates examined in the present study would establish G as a seventh genotype in addition to the major six genotypes (A–F).

Two remarkable traits have been pointed out in the original HBV/G isolate (AF160501) that are not shared by any HBV isolates of the other six genotypes (A–F). Most remarkably, HBV/G has an insertion of 36 bp at the fifth nucleotide in the C gene (26), which is exhibited by another isolate (11) and shared by all eight of the HBV/G isolates studied herein (Fig.

### TABLE 2. Sequence homology of 10 HBV/G isolates to previously reported HBV isolates of the other six genotypes

<table>
<thead>
<tr>
<th>Part of HBV</th>
<th>A (9)</th>
<th>B (16)</th>
<th>C (17)</th>
<th>D (18)</th>
<th>E (2)</th>
<th>F (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full genome</td>
<td>87.3–88.1</td>
<td>84.9–86.8</td>
<td>86.1–87.3</td>
<td>85.2–87.5</td>
<td>88.2–88.7</td>
<td>83.4–85.0</td>
</tr>
<tr>
<td>C gene</td>
<td>81.5–89.6</td>
<td>82.9–89.7</td>
<td>81.9–89.7</td>
<td>80.6–89.3</td>
<td>82.3–90.0</td>
<td>81.0–88.9</td>
</tr>
<tr>
<td>F gene</td>
<td>76.2–89.8</td>
<td>77.8–88.4</td>
<td>75.3–88.9</td>
<td>75.3–90.8</td>
<td>75.3–90.8</td>
<td>73.1–86.5</td>
</tr>
<tr>
<td>X gene</td>
<td>84.9–87.3</td>
<td>84.3–86.5</td>
<td>80.9–85.8</td>
<td>84.3–86.2</td>
<td>85.2–86.0</td>
<td>83.7–84.9</td>
</tr>
<tr>
<td>Large S gene</td>
<td>81.7–91.7</td>
<td>79.6–89.7</td>
<td>78.1–90.9</td>
<td>80.5–92.0</td>
<td>80.0–92.2</td>
<td>78.6–87.8</td>
</tr>
<tr>
<td>S gene</td>
<td>94.6–97.5</td>
<td>93.4–94.7</td>
<td>88.9–94.3</td>
<td>94.6–95.7</td>
<td>94.6–95.2</td>
<td>92.2–93.4</td>
</tr>
</tbody>
</table>
FIG. 1. Phylogenetic trees constructed for HBV isolates of seven genotypes on the full genome (a) and on the C (b), P (c), X (d), large S (e), and S (f) genes. Confidence values calculated by bootstrap analysis are indicated after each genotype.
The peculiar 36-bp insertion has been taken advantage of, in detecting HBV/G isolates, by use of its sequence as a type-specific primer in PCR (10).

The second conspicuous brand of HBV/G isolates are two stop codons in the pre-C region at positions 2 and 28, either of which prohibits the translation of the HBeAg precursor (2, 23). The two stop codons were preserved in the original HBV/G isolate and all eight isolates examined in the present study; the one at position 2 is missing in B1-89 while the one at position 28 is retained (11). Thus, all 10 of the HBV/G isolates would have an HBeAg-negative phenotype, which influences the replication and disease-inducing capacity of HBV (17). HBV/G infection would need to be looked into with special reference to the acute and chronic liver diseases it induces, in view of HBV genotypes associated with the severity of hepatitis B (8, 16, 24).

HBV/G may have a distinct geographic distribution just as the major six genotypes do (14, 15). Originally, HBV/G was identified in 11 of 82 (13%) HBV carriers in Georgia and 2 of 39 (5%) carriers from France (26). It does not appear to be a coincidence that the B1-89 isolate (11) is reported to be from France, the genotype of which is deduced to be G because of the 36-bp insertion in the C gene, and that an adw variant of probable genotype G, judged by the peculiar 36-bp insertion, was isolated from a homosexual carrier from San Francisco also infected with human immunodeficiency virus type 1 (1). Furthermore, all eight of the HBV/G isolates in the present study were recovered from patients living in San Francisco. By contrast, HBV/G was not detected in any HBV carriers from Nagoya, Japan. It seems as if HBV/G prevails in restricted areas in the world, possibly via particular routes of transmission.

In the original report, HBeAg was detected in the serum from a carrier infected with HBV/G, in spite of an HBeAg-negative phenotype deduced from the two pre-C region stop codons (26). It may strike one as a big surprise that all eight of the HBV/G carriers examined in the present study possessed HBeAg in the serum as well. Moreover, HBeAg was detected in sera from two individuals from whom HBV/G isolates (B1-89 and an adw variant) were recovered (1, 11).

HBV/A with an HBeAg-positive phenotype was invariably detected, by restriction fragment length polymorphism, in the S gene (18) of the eight patients from San Francisco from whom HBV/G isolates were recovered. The coinfection with HBV/G and HBV/A is also deduced in an HBeAg-positive individual in France from whom the B1-89 isolate was obtained (11). Taken altogether, coinfection of HBV/G with HBV/A would be frequent. HBeAg in the sera of individuals infected with HBV/G examined in previous studies (11, 26) or in the present study, therefore, would be attributed to the HBV/A with which they were coinfected.

In a previous report (9), the sera of four of the eight patients (corresponding to isolates USG15, USG16, USG17, and USG18) from whom HBV/G isolates, whose entire genomes were sequenced in the present study, were recovered were examined for partial HBV DNA clones with special reference
to HBeAg and antibody to HBeAg in serum. The clones of HBV/G and those of HBV/A were recovered from them in various proportions. In addition, HBV DNA clones representing the recombination between HBV/G and HBV/A, within a partial sequence from the X gene to the C gene, were recovered from three of them (USG15, USG17, and USG18). In one of them (USG17), a shift from the clones of HBV/A to those of HBV/G occurred along with seroconversion from HBeAg to

FIG. 2. (a) Nucleotide sequences of the eight HBV/G isolates within the pre-C region and a 5’-terminal part of the C gene. Sequences of the original HBV/G strain (AF160501 [25]) and an isolate of retrospective genotype G (adw2 variant [27]) are shown for comparison. The 36-bp insertion in the C gene and two stop codons in the pre-C region are shaded. (b) Amino acid sequence of the C gene product in the eight HBV/G isolates. Sequences of the other six genotypes (A–F) in representative strains and the original HBV/G strain (AF160501 [26]) as well as an isolate with retrospective genotype G (adw2 variant [27]) are shown for comparison.
antibody to HBeAg in serum. These results would be taken as evidence to ascribe the presence of HBeAg to the HBV/A coinfecting virus with HBV/G, although transfection experiments are required to ascertain that HBV/G has an HBeAg-negative phenotype.

Given strong evidence for frequent coinfection of HBV/G with HBV/A, one wonders whether HBV/G is replication competent by itself or if it represents a defective virion in need of evidence to ascribe the presence of HBeAg to HBV/A, one wonders whether HBV/G is replication competent by itself or if it represents a defective virion in need of HBV/G and HBV/A genomes manifested itself with high homology within the S gene sequence up to 94.6 to 97.5% between HBV/G and HBV/A genomes manifested itself with high homology within the S gene sequence up to 94.6 to 97.5% that is closer than the divergence of >4% in the S gene separating the six major genotypes (A–F) (20). The similarity may well be due to coinfection with HBV/G and HBV/A and the recombination between them.

Replication competence of the HBV/G genome can be evaluated in transfection of HepG2 or Huh7 cell lines by it alone or with HBV/A and, better still, in experimental transmission to chimpanzees. The 36-bp insertion changes the conformation of the ϵ-encapsulation signal (7) and gives rise to an additional Watson–Crick pair between A1914 (the last nucleotide inserted) and U1817 (the fourth nucleotide in the pre-C region) that can consolidate it theoretically. Whether or not it might give HBV/G an advantage to overgrow HBV/A and thrive in some infected individuals (9) would be another matter of virological interest.

REFERENCES


