Genetic Analysis of Hepatitis A Virus Strains That Induced Epidemics in Korea during 2007–2009

Hyeokjin Lee, Hyeseok Jeong, Heasun Yun, Kisang Kim, Jong-Hyun Kim, Jai Myung Yang, and Doo-Sung Cheon

Division of Enteric and Hepatitis Viruses, Center for Infectious Diseases, National Institute of Health, Korea Centers for Disease Control and Prevention, Chungbuk, Department of Life Science, Sogang University, Seoul, and Department of Pediatrics, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

Hepatitis A virus (HAV) is the only member of the genus Hepatovirus of the family Picornaviridae. HAV occurs worldwide, and humans are thought to be its natural host. The virus is transmitted via person-to-person contact, through the fecal–oral route, and by the ingestion of contaminated food and water (18, 26, 30, 31). The virion is a naked, spherical particle with a diameter of 27 to 32 nm. Its linear, single-stranded, positive-sense RNA genome is approximately 7.5 kb in length and can be divided into 3 functional regions, designated P1, P2, and P3. The P1 region encodes capsid polypeptides (VP1, VP2, VP3, and a putative VP4), whereas the P2 and P3 regions encode nonstructural proteins associated with viral replication (22, 26, 27). The amplification and sequencing of variable regions within the capsid proteins, that is, the VP3/VP1 and VP1/P2A junctions of wild-type HAV isolates from different regions of the world, revealed that there was significant nucleotide sequence heterogeneity but limited amino acid heterogeneity (1, 13, 27, 29).

HAV displays a high degree of antigenic (amino acid) and genetic (nucleotide) conservation throughout the genome. However, enough genetic diversity exists to define several HAV genotypes and subgenotypes. When these are defined by sequence variation within the VP1/P2A junction, there is 15% nucleotide variation between isolates and 7 to 7.5% nucleotide variation between subgenotypes. Of the 6 HAV genotypes identified thus far, 3 (genotypes I, II, and III) are of human origin, and 3 (genotypes IV, V, and VI) are of simian origin (2, 4, 6, 7, 19, 33).

In developing countries, HAV infection occurs early in life, and a large proportion of the population has HAV antibodies (anti-HAV) (11, 25). In Korea (comprising North and South Korea), economic growth has led to changes in the pattern of the age-specific seroprevalence of anti-HAV antibodies; specifically, the prevalence of anti-HAV antibodies in 1- to 20-year-olds declined from 60% in 1980 to 9% in 1995 (3, 14). As a result, this age group is now at a high risk for HAV infection, as reflected by the fairly continuous reports of small HAV outbreaks among adolescents and young adults in South Korea. Over 1,500 cases of clinically overt hepatitis A occurred in 1998, while only a few cases had been reported prior to the early 1990s (17).

Here, we report the findings of a phylogenetic analysis of HAV isolates from Korean patients with acute, IgM-positive hepatitis. The samples were collected during 5 separate outbreaks from 2007 to 2009.

MATERIALS AND METHODS

Clinical specimens from outbreaks. Serum samples and stool specimens were collected from 64 patients with acute hepatitis from 5 outbreaks in Korea (Fig. 1) during 2007 to 2009. The stool specimens were diluted to a 10% suspension with phosphate-buffered saline. Both sera and stool suspensions were stored at −70°C.

Extraction of viral RNA and amplification of HAV RNA. Viral RNA was extracted from anti-HAV IgM-positive sera and from the 10% stool suspensions. We used QiAamp viral RNA minikits (Qiagen, Hilden, Germany) to extract 140 μl from both serum and stool suspensions, according to the manufacturer’s instructions. The viral RNA was eluted with a total volume of 50 μl elution buffer from the spin column.

To amplify HAV for the detection of the VP3/VP1 and VP1/P2A junction regions, we performed a single-step reverse transcription–PCR (RT-PCR) assay with Ready2Use RT-PCR One-step Redplus Mastermix (GenDEPOT, CA). The RT-PCR was performed with a total volume of 50 μl containing 5 μl purified RNA, 10 μM (each) forward (HAV1 and BR-5) and reverse (HAV2 and BR-9) outer primers (Table 1), 25 μl of 2× Satoshi
reaction mixture, 1 μl of RNase inhibitor (40 U/μl; Promega), and 1 μl of avian myeloblastosis virus (AMV) reverse transcriptase (1 U/μl). Reaction for reverse transcription was performed for 40 min at 48°C, followed by denaturation for 4 min at 95°C; 30 cycles of denaturation for 30 s at 95°C, annealing for 30 s at 55°C, and extension for 30 s at 72°C; and a final extension step for 7 min at 72°C.

Nested PCR amplification of cDNA was performed with a total volume of 50 μl containing 5 μl of cDNA template, 10 μM (each) forward (HAV3 and RJ-3) and reverse (HAV4 and BR-6) inner primers (Table 1), 5 μl of 10× Ex Taq buffer, 4 μl of deoxynucleoside triphosphate (dNTP) (2.5 mM each dNTP), and 0.5 μl of TaKaRa Ex Taq (5 units/μl). The second round of PCR was performed with an initial denaturation step for 4 min at 95°C; 30 cycles of denaturation for 30 s at 95°C, annealing for 30 s at 55°C, and extension for 30 s at 72°C; and a final extension step for 7 min at 72°C. PCR products (6 μl) were loaded onto a 1% agarose gel, electrophoresed, and stained with ethidium bromide for band visualization (expected lengths, 186 and 234 bp).

Cloning of PCR products and sequencing. After the amplified products were excised from the agarose gel, they were purified with a QIAquick gel extraction kit (Qiagen). The purified DNA was cloned into the pCR 2.1-TOPO cloning vector (Invitrogen, Carlsbad, CA) using the following protocol: 96°C for 1 min and 25 cycles at 96°C for 10 s, 50°C for 10 s, and 60°C for 4 min, with both the forward (M13F) and reverse (M13R) primers for the vector.

Analysis of nucleotides. We used a multiple-alignment algorithm (the Clustal method) to compare the nucleotide sequences of the Korean HAV isolates with HAV sequences previously deposited in the GenBank database ([http://www.ncbi.nlm.nih.gov/GenBank/index.html](http://www.ncbi.nlm.nih.gov/GenBank/index.html)). Analyses were performed by using the MegAlign package (for Windows, version 3.12e; DNASTAR, Madison, WI) and the MEGA program (version 4.0). We used the CLUSTAL W program ([32](http://jcm.asm.org/)) to align strains from various geographical regions. We visualized the relationships between sequences using a dendrogram, in which the length of each pair of branches represents the distance between the sequences. Phylogenetic trees were constructed by using the neighboring method.

RESULTS

**Epidemiology.** Between June 2007 and May 2009, 5 HAV outbreaks were reported to the Center for Infectious Diseases, National Institute of Health, Korea Centers for Disease Control and Prevention. A total of 64 patients from the Kaesong Industrial Region, Jeonnam, Daegu, Seoul, and Incheon were found to be anti-HAV IgM positive (Fig. 1 and Table 2). The patients ranged in age from 16 to 43 years (mean, 28 years). Only 1 infection (in outbreak 3) was fatal; the patient had previously suffered from chronic hepatitis B virus (HBV) infection.

Three of the outbreaks (outbreaks 1, 2, and 5) occurred in companies, 1 (outbreak 3) occurred in a school, and 1 (outbreak 4) occurred in 2 high schools. We did not detect any HAV RNA in samples taken from food handlers in the cafeterias associated with the outbreaks, indicating that these were not the sources of the HAV infections.

**Detection of HAV RNA in serum and stool samples.** Among serum samples taken from the 64 HAV IgM-positive patients, 60.9% (n = 39) were positive for HAV RNA (Table 2). We were able to amplify HAV RNA from the VP3/VP1 and VP1/P2A junction regions of all 39 samples.

**Genetic analysis of HAV strains from outbreak cases.** Of the 39 isolated strains, a total of 13 (11 strains from outbreak 1 and 2 strains from outbreak 2) were classified as genotype IA strains, with 100% homology among samples from the same outbreak. A further 17 strains (15 strains from outbreak 4 and 2 strains from outbreak 5) were classified as genotype IIA strains, with 98.9 to 100% homology among strains from the same outbreak. The remaining 9 strains (from outbreak 3) were classified as genotype

![](http://jcm.asm.org/)

**FIG 1** Geographical distribution of reported and confirmed hepatitis A cases in Korea from 2007 to 2009. Asterisks indicate the settlement of Kaesong in North Korea. Kaesong Industrial Region is a special administrative industrial region of North Korea. The rectangle box indicates Jeju Island, which is in the southwest of South Korea. 

**TABLE 1 Nucleotide sequences of primers used for the amplification of HAV RNA**

<table>
<thead>
<tr>
<th>Region</th>
<th>Primer</th>
<th>Sequence (5′→3′)</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP3/VP1 junction</td>
<td><strong>External</strong></td>
<td>HAV1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>GCTCCTCTTTATCATGCTATGGAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HAV2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>CAGGAAATGTCTCAAGTTTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HAV3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>ATGTTACTACAACTGTTTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HAV4&lt;sup&gt;b&lt;/sup&gt;</td>
<td>GATCCTCAATTGTTGATAGC</td>
</tr>
<tr>
<td>VPI/P2A junction</td>
<td><strong>External</strong></td>
<td>BR-3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>TTTCTCGTACTACAGTCAGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR-9&lt;sup&gt;c&lt;/sup&gt;</td>
<td>AAGTYACCTCAAGCCCAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RJ-3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>TCCCCAGAGCWCWTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR-6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>AGGAGTTGGMARCATTTT</td>
</tr>
</tbody>
</table>

<sup>a</sup> See reference 1.

<sup>b</sup> See reference 27.

<sup>c</sup> See reference 13.

<sup>d</sup> We modified nucleotide sequences of primers for the detection of the VP1/P2A junction region.
IIIA strains, with 95.7 to 100% homology among samples from the same outbreak (Fig. 2 and Table 2).

HAV strains from outbreaks 1, 2, 4, and 5 each belonged to a single cluster. However, the HAV strains from outbreak 4 formed 2 distinct clusters, indicating that the outbreak may have originated from more than one source.

**Amino acid analysis of the VP3/VP1 and VP1/P2A junction region sequences.** We obtained 39 samples for which we could compare VP3/VP1 junction region amino acid sequences with those from a reference sample. We observed 3 substitutions among isolates from outbreak 1 (Kaesong Industrial Region) \((n = 11)\): 499F→S, 504P→L, and 525K→E (Fig. 3A and Table 3). There were also 3 substitutions among isolates from outbreaks 3 to 5 (Daegu, Seoul, and Incheon, respectively) \((n = 26)\): 512L→P, 527I→T, and 534R→Q (Fig. 3B and Table 3).

However, when we compared VP1/P2A junction region amino acid sequences with those of a reference strain, we confirmed only 1 substitution among 11 isolates sampled: a Q→S amino acid change at position 810, corresponding to the C→T and A→C nucleotide substitutions at nucleotides 2428 and 2429, respectively (Table 3).

**DISCUSSION**

Of 39 HAV strains sampled in Korea during the 2007 to 2009 outbreaks, 13 (from outbreaks 1 and 2) were classified as genotype IA strains, with 100% homology among strains from the same outbreak, and 26 (from outbreaks 3 to 5) were classified as genotype IIIA strains, with 95.7 to 100% homology among strains from the same outbreak. The genotype IIIA strain was particularly dominant in outbreak 3 (Daegu), with 95.7 to 100% homology among strains.

HAV strains from outbreaks 1, 2, 4, and 5 each belonged to a single cluster. However, the strains from outbreak 3 could be grouped into 2 clusters, suggesting that there were multiple outbreak sources (Fig. 1). Interestingly, genotype IA strains, which were responsible for outbreaks 1 and 2, were isolated only between June 2007 and April 2008. Genotype IIIA strains were isolated from the remaining 3 outbreaks (outbreaks 3 to 5), indicating that there are 2 genotypes of HAV strains cocirculating in Korea and that genotype IIIA is responsible for the recent surge in outbreaks.

Previous studies reported similar results for amino acid analyses of the genotype IA VP1/P2A junction region: an Arg-801-to-Lys (R→K) substitution and a Gln-810-to-Ser (Q→S) substitution (3, 35). These features were observed only for strains isolated in Korea. Since the VP1/P2A region is more conserved and exhibits fewer amino acid sequence variations, these substitutions could be very important (16, 24).

We also found amino acid substitutions in the VP3/VP1 junction region sequences. Interestingly, 11 isolates from the first outbreak (Kaesong Industrial Region) contained 2 different substitutions (Phe-499 to Ser and Lys-525 to Glu) than those previously reported for both overseas and Korean strains. Amino acid substitutions among strains from the Kaesong Industrial Region were similar to those observed for a Korean strain isolated in 2006 (35), although we were unable to confirm whether they matched substitutions documented for HAV strains \((n = 244)\) detected during a 2007-2008 outbreak (data not shown). Regardless, it appears that there has been an increase in amino acid variation among genotype IA strains isolated in Korea.

We observed 3 amino acid substitutions (512L→P, 527I→T, 534R→Q) among the HAV strains from outbreak 4 (Fig. 3B and Table 3). However, when we compared VP1/P2A junction region amino acid sequences with those from a reference strain, we confirmed only 1 substitution among 11 isolates sampled: a Q→S amino acid change at position 810, corresponding to the C→T and A→C nucleotide substitutions at nucleotides 2428 and 2429, respectively.
and 534R→Q) (Fig. 3B) in the VP3/VP1 junction region among amino acid sequences obtained from outbreaks 3, 4, and 5 (genotype IIIA) (n = 26). HAV strains isolated from outbreak cases 4 and 5 in 2009 did not cluster with other genotype IIIA strains isolated from Korean outbreaks between 2005 and 2008 (data not shown). Thus, it appears that the prevalence of genotype IIIA has changed over time. Specifically, the nucleotide variation of genotype IIIA strains abruptly increased in Korea during 2009.

FIG 2 Neighbor-joining phylogenetic trees of the nucleotide sequences of HAV strains isolated from Korean outbreaks between 2007 and 2009. (A) Phylogenetic tree of the VP3/VP1 region of 39 strains from this study and reference HAV strains. (B) Phylogenetic tree of the VP1/P2A region of 39 strains from this study and reference HAV strains. We analyzed 186-nt and 234-nt sequence lengths for the VP3/VP1 and VP1/P2A regions, respectively. The phylogenetic tree incorporates Korean strain data from the current study as well as data from reference strains derived from GenBank (accession numbers AB020568 for FH1, AF453466 for Arg081, L20541 for M-118, M14707 for HM-175, AF268396 for HAF-203, L07693 for CF-53, AY032861 for SLF-88, M34084 for PA21, AT574076 for B8V15, L07668 for GA-76, L07688 for H-122, L20532 for A-229, L07732 for CY-145, D00924 for AGM-27, and L07731 for JM-55). Sequence analysis was conducted using MEGA 4.0 software. Bootstrap values are indicated for the major nodes as a percentage of the data obtained from 1,000 resamplings.
Historical data suggest that acute hepatitis A infection can result in significantly increased mortality in patients with underlying hepatitis B infection and probably other chronic liver diseases as well. HBV, whether manifesting as an acute, chronic, or occult infection, has the potential to significantly worsen the course of underlying liver disease. Strong epidemiological evidence suggests an increased occurrence of fulminant liver failure, cirrhosis, and hepatocellular carcinoma in patients with HBV superinfection (15, 26). In this study, 1 fatal case was reported for outbreak 3. The patient already suffered from chronic hepatitis B infection prior to being infected with HAV (Table 2). The patient had a high HBV titer (1.8 × 10^5 IU/ml) and subsequently died from fulminant liver failure. Thus, both this current case and previous research strongly suggest that patients with chronic hepatitis should be given HAV vaccination.

We focused our efforts on a 168-nucleotide (nt) fragment containing the VP1/P2A junction (28); recently, this fragment has been used for analyses of many sequences, especially for comparing sequences of isolates obtained from several countries (5, 8–10, 12, 20, 34). Since the 168-nucleotide fragment is insufficient to identify the genotype or the genetic relationship among the HAV isolates, many studies have extended their focus to a larger fragment of the VP1/P2A junction (5, 9, 10, 20, 21, 23). At first, our study was limited by the fact that we could not detect the VP1/P2A junction region in genotype IIIA strains. Therefore, we confirmed the primer target region of Korean HAV strains to detect the VP1/P2A junction region. We modified primer sequences that cover this larger fragment (Table 1).

In a previous study of HAV infections in Korea during 2005 to 2006, we found that genotype IA was the major strain, while the prevalence of genotype IIIA strains was gradually increasing. Our current work confirms that these 2 genotypes are still cocirculating in Korea. However, genotype IIIA is now more prevalent, which may be the reason for the current HAV outbreaks.

**REFERENCES**


