Genetic variability of hepatitis A virus isolates in Rio de Janeiro: implications for the vaccination of school children

L.M. Villar1, E. Lampe1, A. Meyer2 and A.M.C. Gaspar1
1Departamento de Virologia, Instituto Oswaldo Cruz, and 2Laboratório de Toxicologia CESTEH/Fiocruz, Rio de Janeiro, RJ, Brasil

Abstract

The epidemiology of hepatitis A virus (HAV) infection is shifting from high to intermediate endemicity in Brazil, resulting in increased numbers of susceptible individuals and a greater potential for the emergence of outbreaks. Universal vaccination against HAV has been recommended for children, but updated sero-epidemiological data are necessary to analyze the level of natural immunity and to identify candidates for preventive measures. In addition, more molecular studies are necessary to characterize the genotypes involved in HAV infections and outbreaks. Sera from 299 school children (5-15 years old) and 25 school staff members, collected during an outbreak of HAV at a rural public school in June 2000, were tested for IgM and total anti-HAV antibodies (ELISA). Viral RNA was amplified by RT-PCR from anti-HAV IgM-positive sera and from 19 fecal samples. Direct nucleotide sequencing of the VP1/2A region was carried out on 18 PCR-positive samples. Acute HAV infection was detected by anti-HAV IgM in 93/299 children and in 3/25 adult staff members. The prevalence of total anti-HAV antibodies in IgM-negative children under 5 years of age was only 10.5%. HAV-RNA was detected in 46% IgM-positive serum samples and in 16% stool samples. Sequence analysis showed that half the isolates belonged to subgenotype IA and the other half to IB. On the basis of these data, mass vaccination against HAV is recommended without prevaccination screening, especially for children before they enter school, since nearly 90% of the children under 5 years were susceptible. Molecular characterization indicated the endemic circulation of specific HAV strains belonging to subgenotypes IA and IB.

Introduction

Hepatitis A virus (HAV) belongs to the Picornaviridae family and is the representative member of the Hepatovirus genus (1). The HAV genome is a positive strand RNA of approximately 7.5 kb with a single-open reading frame for a large polyprotein of about 2230 amino acids (2). HAV strains recovered from many parts of the world are classified within a single serotype and divided into seven genotypes on the basis of phylogenetic analysis of nucleotide sequences in the VP1/2A region (3-5). Most of the human strains studied belong to genotypes I and III, with 80% of them being genotype I.

HAV infection is a major public health problem in Brazil, accounting for the majority...
of reported cases of acute hepatitis (6,7). Transmission of HAV usually occurs by the fecal-oral route. Contaminated water and food frequently cause outbreaks (8-11), especially in closed communities such as schools (12,13) and day-care centers (14,15). A shift from high to medium endemicity of HAV infection has been observed in Brazil in the last decades due to improvements in hygiene and sanitation levels as also reported for many other Latin American countries. Paradoxically, an increase in the number of susceptible children, adolescents and adults has been observed, a situation resulting in a greater potential for the emergence of outbreaks (16). In areas of low or intermediate endemicity, children and adolescents are regarded as high risk groups since they are less likely to have natural immunity.

Hepatitis A is among the most common vaccine-preventable diseases in the world. Hepatitis A vaccination has proved to be safe, even in children under two years of age (17). Mass HAV vaccination of children is being recommended in regions of low endemicity and to interrupt outbreaks among children living in communities with the highest rates of disease (17).

In Brazil, the prevalence of HAV has been studied in various population groups (13,18,19). However, there is only one study on the Rio de Janeiro area in which the HAV genotypes were determined (15), showing that the most prevalent genotypes were IA and IB. More detailed epidemiological and molecular studies are needed for a better understanding of the risk factors and genotypes involved in the transmission of HAV.

The present study was undertaken to evaluate the seroprevalence of HAV infection (total and anti-HAV IgM) among school children during an outbreak of HAV infection and to gain more insight into the degree of genetic variability among the HAV strains isolated during this outbreak. Furthermore, an epidemiological investigation was conducted to identify possible risk factors associated with HAV infection. This information can be used to develop a strategy for HAV vaccination in this region.

Patients and Methods

Epidemiological investigation

In June 2000, 14 clinical cases of acute hepatitis among pupils in a public school situated in the rural area of Paracambi were reported to the local Health Authorities. Paracambi is a municipality located 80 km from the capital city of Rio de Janeiro with a population of 40,412 inhabitants. At the time of this outbreak, there were 299 children aged 5-15 years distributed into six classes and 25 school staff members aged 18 to 70 years. The school is in session all day and the children eat meals prepared in the school kitchen. On June 13, 2000, the Paracambi Sanitary Authorities collected water samples from the artesian well and all the taps of the school for microbiological analysis.

All persons attending the school were contacted to obtain detailed demographic and risk factor information. Hepatitis A infection was defined as an illness characterized by the acute onset of jaundice and/or a positive serologic test for anti-HAV IgM antibodies.

To investigate the risk factors for the children attending the school a case-control study was carried out, choosing as controls the healthy school pupils who were negative for anti-HAV IgM. For each exposure factor, the odds ratio (OR) and 95% confidence intervals (CI) between cases and controls were calculated using the $\chi^2$ test with Yates correction. SPSS version 8.0 software was used for univariate and multivariate analysis.

Clinical samples

Serum samples were collected from all 299 students and 25 staff members of the school. Nineteen fecal samples were also obtained from the same group. Formal consent to participate in the study was obtained...
from the parents or guardians of the children and from the school staff. Blood and fecal samples were collected on June 15, 2000, following approval by the Ethics Committee of Fiocruz (RJ) and stored at -20°C before being assayed in the Virology Department of the Oswaldo Cruz Institute in Rio de Janeiro.

**Serological tests**

IgM and total anti-HAV antibodies in serum samples were detected with commercially available ELISA kits (Organon Teknika, Boxtel, The Netherlands) according to manufacturer’s instructions.

**Detection of HAV RNA**

Viral RNA was extracted from anti-HAV IgM antibody-positive sera and from all stool specimens (10% stool suspension (w/v) in PBS). Briefly, 100 µl of serum or feces suspension was extracted by the TRIzol® method as described by Chomczynski and Sacchi (20) (Life Technologies, Rockville, MD, USA). After RT with Moloney murine leukemia virus reverse transcriptase, nested PCR was performed with VP1/2A specific primers as described previously (15). PCR products (10 µl) were submitted to 1.5% agarose gel electrophoresis and stained with ethidium bromide to visualize a band of an expected length of 244 base pairs.

To reduce the probability of sample contamination by amplified DNA molecules, standard precautions were applied in all manipulations. Separate areas were used for reagents, treatment of samples and manipulation of amplified samples. All samples were analyzed twice in independent experiments and a negative control was run with each reaction.

**Sequencing and phylogenetic analysis**

The nested RT-PCR products were purified using the QIAquick Gel extraction kit (Qiagen, Hilden, Germany) according to manufacturer instructions. Direct nucleotide sequencing reaction was carried out on both strands with nested primers using the Thermo Sequenase Cy5 Dye Terminator cycle sequencing kit (Amersham Pharmacia Biotech, Barcelona, Spain) and “ALFexpress” Amersham Biosciences (Barcelona, Spain). To construct a phylogenetic tree, sequences from 168 nucleotides (position 3024 to 3191) of the VP1/2A junction were aligned using the CLUSTAL X program (21). A matrix was then generated for the Kimura-two parameter model (22), used together with the neighbor joining method to create a phylogenetic tree and assessed for reliability by bootstrap resampling (1000 pseudo-replicates). These methods were applied using software from the MEGA program (23). Forty-one HAV strains isolated from different parts of the world were also included in this dendrogram: SLF88 (accession number, L07729) from Africa; Arg-1 (AJ306374), Arg-4 (AJ306372), Arg-23 (AJ306377), and Arg-90 (L07687) from Argentina; HM-175 (M14707) from Australia; EP-13 (L07682), RDJ (L07681), RJ-213 (AF410387), RJ-193 (AF410388), EP-49 (L07686), RJ-PMC1 (AF410386), HAF-203 (AF26896), RJ-004 (AF410380), RJ-055 (AF410384), RJ-504 (AF410385), RJ-NSG2 (AF410383), RJ-PMC3 (AF410382), RJ-005 (AF410381), and EP-35 (L07684) from Brazil; Chile5 (AJ406320) and Chile8 (AJ306379) from Chile; CF-53 (L07693) from France; GBM (X75215) from Germany; HAF-FG (X83302) from Italy; Nor5 (AF050227), Nor-8 (with a history of travel to Morocco) (AF050229), Nor-9 (with a history of travel to Sudan) (AF050230), Nor-10 (with a history of travel to Israel) (AF050231), Nor13 (with a history of travel to Brazil) (AF050234), Nor14 (with a history of travel to Chile) (AF050235), and Nor15 (with a history of travel to Turkey) (AF050236) from Norway; MBB (X14666), SA3 (U68699) from South Africa; Tunisia81 (L07727) from Tunisia; Uru1 (AJ309227), Uru2 (AJ306387) and Uru3
(AJ306386) from Uruguay; HAS-15 (X15464), LA-15 (K02990), MIB35 (AF199045), and MIB4022 (AF199046) from the US. The new sequences obtained in the present study have been deposited in the GenBank sequence database under accession numbers AF538723 to AF538727.

**Results**

**Epidemiology**

The index case hepatitis A patient, with onset of symptoms on May 22, 2000, was a 5-year-old girl who lived in the rural area of Paracambi (RJ). Between May 23 and June 14, 13 additional persons became ill. All of the acute cases notified to the Health Authorities of Paracambi were children who attended the same public school as the index case. The cases were distributed among all classes and only one case was a teacher. All had fever, nausea and fatigue, and 11/14 were icteric. In thirteen of the 14 patients, HAV infection was confirmed serologically by the detection of anti-HAV IgM. None of the patients was hospitalized and all recovered without sequelae. None received a vaccine or immunoglobulin. The epidemic curve showed a sharp peak with most of the clinical cases occurring within a two-week period (Figure 1).

Of the 299 serum samples from children, 47.5% (142/299) were positive for total anti-HAV antibodies. Among them, 93 (31.1%) were seropositive for anti-HAV IgM, indicating acute infection and 49 (16.4%) were seronegative for anti-HAV IgM, indicating past infection. The prevalence of anti-HAV IgM was higher in 7-year-old children (43.54%) compared to 26.31% for 5-year-old children (P = 0.026; OR = 2.0; 95% CI = 1.1-3.7). Furthermore, anti-HAV antibodies were not detected in 152 (52.5%) of the 299 children and 63.15% of the 5-year-old children showed no anti-HAV antibodies (Figure 2). Of the 25 adult staff members, 3 (12%) presented IgM and one of them was HAV-RNA positive.

For the 5- to 15-year-old children involved in this outbreak, multiple plausible risk factors for hepatitis A infection were examined first by univariate logistic regression analysis and then by stepwise multiple logistic regression (Table 1). No statistically significant association was found between the different factors investigated and anti-HAV IgM reactivity among the school children. However, the OR for the lack of a sewage system and the contact with suspected cases of hepatitis A were 1.13 and 1.00, respectively, in the univariate analysis. In the multivariate model, we chose variables associated with the sanitary facilities (toilet without piped water, artesian well, sewage disposal, piped water), plus two other variables, i.e., exposure to a person with hepatitis and bathing in the river. No statisti-
Genetic variability of hepatitis A in Rio de Janeiro

cally significant results were obtained for any of these variables, suggesting that the risk factor for HAV infection was located in the school.

The bacteriological analysis of water samples obtained on June 13, 2000 yielded a population density of >300 cfu/0.1 ml total coliform bacteria and 300 cfu/0.1 ml fecal coliform bacteria in all water samples obtained at the school. Furthermore, in the water tank for the cafeteria, the density of fecal coliform bacteria was >300 cfu/0.1 ml.

HAV-RNA detection

A total of 96 serum samples (93 from children and 3 from adults) positive for anti-HAV IgM and 19 stool samples were tested by RT-PCR. HAV-RNA was detected in 44/96 (46%) IgM-positive serum samples and in 3/19 (16%) stool samples. Among the PCR-positive stool samples, one was obtained from an 8-year-old child who tested negative for anti-HAV IgM in her serum sample and the other two PCR-positive stool samples were from the school staff.

Phylogenetic analysis

Seventeen serum samples and one stool sample collected during the epidemic were analyzed by direct nucleotide sequencing of PCR products. To determine the degree of nucleotide homology, the product of 18 nucleotide sequences (218 bp excluding both primers) was compared with the corresponding sequences of the prototype strain of subgenotype IB (HM-175). The percentage of nucleotide identity ranged from 89.4 to 100% in the VP1/2A region.

In order to gain insight into the genetic relatedness and to identify the genotypes that circulated during this outbreak, phylogenetic analysis of the (168-) nucleotide sequences in the VP1/2A region was carried out (most of the isolates from GenBank show 168 bp in this region). Figure 3 shows the
A phylogenetic tree constructed with samples collected in the present study and with previously published sequences of isolates from different geographical areas, in addition to sequences representative of genotypes IA, IB, II and VII (3). All isolates sequenced in this outbreak clustered in genotype I, with 50% (9/18) clustering in subgenotype IA and the other 50% in subgenotype IB. The resulting bootstrap values were higher than 80%, confirming the branching order of the phylogenetic tree.

Strains belonging to genotype IA from the outbreak under study were divided into three subclusters and most of them (7/9) showed in the 168-base segment 100% identity with the strain (RJ-PMC1-6X) previously isolated in Rio de Janeiro in 1999 (15). The strains belonging to subgenotype IB were identical to each other and to strains RJ-004 (from 16 patients with sporadic cases of acute hepatitis in 1999); (15) and to strain HAF-203 which were isolated in Brazil in 1992 (24). Strains isolated from this out-

Table 1. Demographic and environmental characteristics of children attending school during the hepatitis A outbreak in Paracambi, Rio de Janeiro, 2000.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Anti-HAV IgM</th>
<th>Risk for HAV infection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive N = 93 (31.10%)</td>
<td>Negative N = 206 (68.9%)</td>
</tr>
<tr>
<td>Number of children living in the same household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2</td>
<td>69 (33.33)</td>
<td>138 (66.66)</td>
</tr>
<tr>
<td>&gt;2</td>
<td>24 (26.66)</td>
<td>66 (73.33)</td>
</tr>
<tr>
<td>Number of persons per room</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤2</td>
<td>51 (36.17)</td>
<td>90 (63.82)</td>
</tr>
<tr>
<td>&gt;2</td>
<td>42 (27.27)</td>
<td>112 (72.72)</td>
</tr>
<tr>
<td>Sewage disposal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>76 (30.52)</td>
<td>173 (69.47)</td>
</tr>
<tr>
<td>No</td>
<td>15 (33.33)</td>
<td>30 (66.66)</td>
</tr>
<tr>
<td>Swimming in rivers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>12 (27.27)</td>
<td>32 (72.72)</td>
</tr>
<tr>
<td>No</td>
<td>75 (31.91)</td>
<td>160 (68.08)</td>
</tr>
<tr>
<td>Toilet without piped water</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>38 (44.18)</td>
<td>75 (37.87)</td>
</tr>
<tr>
<td>No</td>
<td>48 (55.81)</td>
<td>123 (62.12)</td>
</tr>
<tr>
<td>Piped water</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>78 (31.07)</td>
<td>173 (68.92)</td>
</tr>
<tr>
<td>No</td>
<td>12 (30.76)</td>
<td>27 (69.23)</td>
</tr>
<tr>
<td>Artesian well</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>17 (27.41)</td>
<td>45 (72.58)</td>
</tr>
<tr>
<td>No</td>
<td>71 (31.98)</td>
<td>151 (68.01)</td>
</tr>
<tr>
<td>Class attendance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-school</td>
<td>20 (28.57)</td>
<td>50 (71.42)</td>
</tr>
<tr>
<td>1st grade</td>
<td>32 (44.44)</td>
<td>40 (55.55)</td>
</tr>
<tr>
<td>2nd grade</td>
<td>12 (26.08)</td>
<td>34 (73.91)</td>
</tr>
<tr>
<td>3rd grade</td>
<td>10 (20.00)</td>
<td>40 (80.00)</td>
</tr>
<tr>
<td>4th grade</td>
<td>11 (31.42)</td>
<td>24 (68.57)</td>
</tr>
<tr>
<td>Supplementary class</td>
<td>4 (28.57)</td>
<td>10 (71.42)</td>
</tr>
<tr>
<td>Exposure to a person with hepatitis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25 (35.21)</td>
<td>46 (64.78)</td>
</tr>
<tr>
<td>No</td>
<td>37 (28.46)</td>
<td>93 (71.53)</td>
</tr>
</tbody>
</table>

Numbers within parentheses are in percent. HAV = hepatitis A virus; OR = odds ratio; CI = confidence interval; ND = not done. Totals do not add up to 299 because of missing values.
break as well as other strains from Brazilian isolates tended to form separate subclusters in the phylogenetic tree constructed with different strains isolated from other parts of the world.

**Discussion**

The outbreak described in the present study is an example of the changing patterns of HAV infection occurring in recent decades in Brazil. In the past, when the infection was highly endemic, infection occurred most frequently in the first years of life, with a low disease/infection ratio. More recently, reduced endemicity has led to the accumulation of a large number of susceptible individuals (18,19). Consequently, the introduction of HAV in a community may easily lead to outbreaks of disease, with the occurrence of a large number of cases and the involvement of different age groups (13).

During the present outbreak, 96 individuals tested positive for anti-HAV IgM and 14 acute clinical cases of hepatitis A were diagnosed. The initial symptoms of the acute clinical cases occurred in a unimodal curve with a peak from May 28 to June 2. The epidemic curve, with a sharp peak where most of the clinical cases occurred within the two-week period, is characteristic of a common source outbreak (9). The bacteriological analysis showed heavy fecal contamination in the school water supply, mainly in the watering trough of the cafeteria. Further investigation showed that there was a break in the pipe causing sewage to mix with piped water. These findings implicated the school water as the most probable vehicle of HAV transmission. However, water samples were not tested by RT-PCR, and a definitive connection between water consumption and the viruses found was not established.

Overall, 47.5% (142/299) of the school children had total anti-HAV antibodies but only 49 children were immune. The seroprevalence of total anti-HAV antibody in younger children, particularly at the age of 5 years (the age when children are exposed to the outer environment in addition to their home and are therefore at risk of contact with HAV infection), would be important to assess the magnitude of the actual target population at risk for HAV infection. In fact, a low level of HAV seropositivity was found in 5-year-old children, indicating that mass HAV vaccination in this population might be cost-effective. These prevalences are lower than that found by Vitral et al. (18), demonstrating that the level of immunity has been declining in the Brazilian population, so that outbreaks will probably occur frequently.

Genotype I comprises 80% of the HAV strains isolated worldwide, as described by Robertson et al. (3). Subgenotype IA represents most of the human strains found worldwide (3). Early studies demonstrated that HAV strains isolated in Latin America belong only to subgenotype IA (25,26). In this report we observed the co-circulation of subgenotypes IA and IB. In previous studies both genotypes have been detected in samples collected in Rio de Janeiro (15,27), confirming that both genotypes are present in this region. Similarly, in an outbreak in Spain, genotypes IA and IB were found in clinical samples at a proportion of 50% each. Co-circulation of two subgenotypes was also reported in France (28) and South Africa (29).

Sequences from this outbreak as well as other sequences from Brazilian isolates tended to form a separate subcluster in the phylogenetic tree constructs, particularly in the subgenomic IA grouped sequences, in which all isolates were unrelated to strains obtained in other parts of the world. Furthermore, HAV strains from this outbreak were closely related to strains previously isolated in Rio de Janeiro. In addition, strains belonging to genotype IA isolated in this outbreak agree with the strain isolated from a patient in Norway (30) who had a history of travel to
Brazil in 1996. Also, all strains classified as genotype IB were clustered with the strain HAF-203, which was isolated in Brazil in 1992 (24). These findings indicate that some strains have been circulating endemically in Brazil for more than 12 years.

Hepatitis A vaccination can potentially serve as an effective intervention, particularly if a group of susceptible subjects can be identified and targeted. Children should be targeted in any treatment program due to their critical role in hepatitis A transmission. In countries with intermediate endemicity, the risk of outbreaks is high, with a large and growing population of susceptible individuals. In the present study, we demonstrated that vaccination should be recommended without prevaccination screening, mainly before children start to frequent school since more than 60% of 5-year-old children are susceptible. The prevalence of disease in these locations probably reflects the prevalence within the community (17). The future routine use of a vaccine in children in intermediate endemicity communities, as well as in other susceptible groups, may result in benefits such as decreased morbidity and lower rates of fulminant hepatitis.

Acknowledgments

The authors thank the Health Authorities of Paracambi for help with data collection and the Sanitary Authorities of Paracambi for the bacteriological analyses of the water. They also wish to thank Luiz Guilherme Bauzer for assistance with the manuscript, and the technicians of the Brazilian Reference Center for Viral Hepatitis for technical assistance.

References


17. Centers for Disease Control (1999). Prevention of hepatitis A through active or passive immunization; recommendations of the