Molecular Characterization of Hepatitis A Virus Isolates from Nigeria

Joseph C. Forbi    Mathew D. Esona    Simon M. Agwale

Clinical Virology Laboratory, Innovative Biotech, Keffi/Abuja, Nigeria

Key Words
HAV isolates · Nigeria · Genotype · Subgenotype

Abstract
Objective: Despite the endemicity of hepatitis A virus (HAV) in Nigeria, genetic information on the HAV genotypes/subgenotypes circulating in the country remains unknown. The objective of this study was to investigate HAV strains using molecular epidemiological and genetic analyses among apparently healthy adult Nigerian subjects. Methods: Testing for HAV-RNA was performed on 114 serum samples by the reverse transcription-polymerase chain reaction and sequenced with primers encompassing the VP1/P2A junction. Results: Twelve serum samples tested were found to be HAV-RNA positive. Phylogenetic analysis revealed that all 12 HAV isolates were classified as subgenotype IA exhibiting 98.3% nucleotide identity. Interestingly, the Nigerian HAV/IA subgenotype consisted of two distinct genomic sublineages with a unique majority (n = 11) corresponding to strains endemic in Cameroon and the other (n = 1) shows a probable link with European sequences. Predicted conserved amino acid sequences and the few deduced substitution in the VP1/P2A junction might play a role in the development of a novel Nigerian-Cameroon sublineage within the HAV/IA subgenotype and might explain the stability of HAV/IA in this subregion. Conclusion: This study reveals the development of a new HAV/IA sublineage in the Nigerian-Cameroon subregion. The presence of a single subgenotype indicates that this HAV strain has been predominantly circulating in Nigeria.

Introduction

Hepatitis A virus (HAV) is a member of the genus Hepatovirus within the family Picornaviridae and contains a positive-sense, single-stranded RNA genome of approximately 7.5 kb in length [1]. Although HAV displays only a single serotype, HAV strains isolated from different parts of the world have been classified into six genotypes (I–VI); genotypes I–III are found in humans, and each of these is further divided into subgenotypes A and B. HAV genotypes have unique geographic distributions.
tions; however, cocirculation of multiple genotypes or subgenotypes has been reported in some regions of the world [2]. HAV is transmitted primarily via the fecal-oral route by contaminated food or water [3], but has also been associated with outbreaks in injecting drug users and men who have sex with men [4, 5]. HAV infection has been a major public health problem in many countries worldwide. In Nigeria, HAV infection is endemic as serological tests revealed that over 82% of the healthy unvaccinated Nigerians between the ages of 5–70 years had antibody to HAV [6].

Despite the endemicity of HAV in Nigeria, information is not available about the genotypes/subgenotypes circulating in the country and as a consequence nothing is known regarding the relatedness of HAV isolates from this country and those from other parts of the world. This knowledge is important for establishing evolutionary relationships and transmission events. This study presents the results of the molecular characterization of HAV cases in a group of randomly screened adult individuals with no clinical signs of HAV infection and analyzes the phylogenetic relatedness of the HAVs circulating in North-Central Nigeria. As far as we know, this is the first report on the identification of HAV genotypes and subgenotypes in Nigeria.

**Materials and Methods**

**Samples, Extraction, Reverse Transcription-Polymerase Chain Reaction**

Serum samples (random survey between January and February 2006) from 114 apparently healthy adult Nigerians between the ages of 20 and 52 years living in a locality with suboptimal hygienic standards were used in the present study. This research was carried out in accordance with the Helsinki Declaration and informed consent to participate in the study was obtained from all participants. Viral RNA was extracted from serum samples using microspin columns (QIAamp Viral RNA mini kit; Qiagen, Valencia, Calif., USA) and amplified by nested reverse transcription-polymerase chain reaction (RT-PCR) of the 371-bp region of the VP1/P2A junction of the HAV genome as described elsewhere [2, 7, 8]. The primers used in the first round of PCR were: sense 2870, 5’-GACAGATTCTACATTGGATTGGT-3’; and antisense 3381, 5’-CCATTTCAAGAGTCCACACACT-3’. Inner primers were: sense, 2897 5’-CTATTCAGATTGCAAATTACAAT-3’; and antisense 3288, 5’-AACTTCATTATTTCATGCTCCT-3’.

**Sequencing and Sequencing Analysis**

All HAV-RNA-positive samples were sequenced using the Big Dye Terminator v.3.1 cycle sequencing kit (Applied Biosystems, Nieuwerkerk, The Netherlands) following the manufacturer’s instructions and an ABI 3130 automatic sequencer (Applied Biosystems). Genotyping was performed by sequence analysis after

---

**Fig. 1.** Phylogenetic tree based on 371 nucleotides of the VP1/P2A junction of sequences obtained during the present study (indicated by ▲) as well as reference sequences of genotypes I, II and III (maximum likelihood algorithm using the Kimura two-parameter model). Cameroonian strains are indicated by (●). Reference strains are identified by their GenBank reference numbers. Bootstrap values lower than 70% are hidden.
alignment (ClustalW) of the sequences with reference strains using the Bioedit program version 7.1.3 [9]. The maximum likelihood with Kimura two-parameter methods implemented using MEGA software version 5 [10, 11] with 1,000 bootstrap replications was applied for the phylogenetic analysis. Only bootstrap values of 70 or above are shown in the phylogenetic trees. Genetic distances were calculated with MEGA using pairwise overall mean [10]. All sequences generated in this study were searched against the NCBI database using a Web-based nucleotide BLAST program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) optimized for the blastn algorithm in order to compare our sequences to those found in the sequence database.

**Amino Acid Prediction**

The GEM/1993/Germany (X75215), LU38/2001/China (AF357222), DL3/2002/China (AF512536) GenBank reference strain of subgenotype HAV/IA and HM175/1976/Australia (M14707) representing HAV/IB were chosen for phylogenetic comparative studies. The VP1/P2A genome sequences obtained during the present study were submitted to GenBank under accession numbers JQ696954 to JQ696965.

**Results**

Molecular testing (RT-PCR) showed that HAV-RNA was detected in 12 of the 114 serum samples tested. Sequencing and phylogenetic analysis of 371 nucleotides of the VP1/P2A junction of the 12 strains attributed all of them (100%) to subgenotype HAV/IA (fig. 1). The HAV-PCR-positive cases were considered to be viremic, probably indicating a silent ongoing HAV epidemic in the country. The subgenotype IA isolates could be sorted into two sublineages. One predominant sublineage clustered with several isolates that caused widespread HAV epidemic in the neighboring Republic of Cameroon (fig. 1). The sequences from Nigeria were almost identical to one another and with the Cameroonian strains, suggesting that these two neighboring countries could share a similar source or mode of infection resulting in the development of a novel HAV/IA sublineage.

The second sublineage (n = 1) clustered with isolates from Europe, suggesting that this lineage might share a common ancestral relationship with IA strains from Europe. BLAST searches of the Nig-11 sequence revealed that it was most closely related to strains circulating in Southern Italy in 2002 and also to stains responsible for outbreaks in Germany in 2000 and 2007/2008 (AJ505576, AJ505573, AY028976, EU416242; 99% maximal identity).

The genotype IA strains from Nigeria showed an overall mean genetic distance of 98.3%. The genetic identity of sequences from Nigeria (98.3%; diversity = 1.7%) was slightly lower than among those from Cameroon (99.5%; diversity = 0.5%). The increased diversity among sequences from Nigeria could be attributable to the two ancestral sublineages as described above. Genetic diversity suggests that HAV/IA was first introduced in Nigeria before spreading to Cameroon. The source(s) of infection among the study participants was not identified.

The predicted amino acid sequences of the VP1/P2A junction region (123 amino acids) of the 12 Nigerian HAV strains were compared with each other and with previously published reference strains in the GenBank, representing subgenotypes HAV/IA and HAV/IB. All the Nigerian isolates in this study showed an overall 0.7% variability in the VP1/P2A junction region at the amino acid level when compared with each other. The data indicates that the amino acids in this region were much conserved. However, a few unique point amino acid changes could be deduced (fig. 2): all sequences except Nig-11 showed an arginine (R) → lysine (K) substitution at position VP1–297, which is a typical IB substitution [12]. Additionally, isolate Nig-11 and Nig-4 also showed serine (S) → alanine.
(A) substitution and serine (S) → proline (P) at position VP1–286, respectively. All isolates except Nig-2 showed an alanine (A) → serine (S) substitution at position VP1–232. The data also showed a valine (V) → leucine (L) substitution at position P2A-42 in all sequences (fig. 2). The presence of S286A and the absence of a substitution at position 297 of the VP1 in Nig-11 probably resulted in the segregation of the HAV/IA into two sublineages in this study.

**Discussion**

At present, there are no reports of the presence of HAV subgenotypes circulating in Nigeria. Molecular characterization of HAV genotypes and subgenotypes is important for establishing evolutionary relationships and for understanding its origin or patterns of transmission. Nigeria is a developing country and represents the largest population in West Africa and, as far as we know, knowledge of the specific HAV strains circulating in this population remains unknown. However, the few serological data reported suggest that the virus is endemic in Nigeria [6, 13, 14].

In this study, the molecular epidemiology of HAV viral strains were determined by RT-PCR, and sequencing in the VP1/P2A junction of HAV isolates from 12 HAV sporadic cases of adults with no overt symptoms of hepatitis were sequenced at the VP1/P2A junction of the HAV genome. Phylogenetic analysis demonstrated that all the isolates belonged to subgenotype HAV/IA (fig. 1). The presence of a single subgenotype probably indicates that this HAV strain has been predominantly circulating in Nigeria and confirms that Nigeria is part of the worldwide reservoir of HAV/IA [15]. The circulation of a single HAV subgenotype is remarkable given the high endemicity of the virus in the country. It is probable that HAV/IA is adapted to this population or a founder effect might have been responsible for the predominance of this subgenotype in this region as occurred for HIV-1 in West and Central Africa [16]. The study suggests that HAV infections are subclinical and do not necessarily only occur early in life, but viremic cases could be found among adults resulting in an ongoing silent and unrecognized HAV epidemic in this region.

Nucleotide analysis showed that the sequences from Nigeria exhibited a 98.3% identity. Interestingly, although these sequences were genetically closely related and belonged to the same subgenotype, they could be clearly distinguished into two sublineages (fig. 1). The predominant sublineage constituted a unique cluster of 11 isolates from Nigeria and isolates from Cameroon. This unique Nigeria-Cameroon cluster did not intermix with isolates from other parts of the world, but showed an extensive intermixing with isolates from these two neighboring countries. It appears that a new variant (Nigeria-Cameroon variant) is developing within the HAV/IA subgenotype and this is supported by a high bootstrap value of 99% (fig. 1). It is probable that these two countries probably share common mechanisms that enable the transmission of HAV, or it may be a reflection of the fact that HAV forms highly conserved clusters of rare codons that restrict antigenic variability that could result in similar VP1/P2A sequences from unrelated cases [17]. Full genome analysis of HAV sequences from this region would provide additional information on this. Genetic diversity estimates suggests an east to westward direction of transmission from Nigeria (genetic diversity = 1.7%) to Cameroon (genetic diversity = 0.5%). It is possible that the development of this unique lineage is a reflection of a particular mode of transmission. In this study, risk factors of HAV infection were not determined; as such, further studies to identify the source of HAV in these two countries are warranted. The second sublineage consisted of an isolate (Nig-11) that was genetically similar with Italian and German sequences, suggesting that this isolate might have been imported from Europe with an endemic pattern of transmission of closely related stains. This reflects some form of evolutionary and epidemiological link between Nigerian and European HAV sequences. The presence of two HAV/IA sublineages in this study confirms that this virus was probably introduced in this country through two independent routes.

It seems obvious that there is an endemic transmission of subgenotype HAV/IA in Nigeria. Few major point amino acid changes could be deduced (fig. 2), and these changes could account for the predominance and stability of a unique HAV/IA variant in the Nigeria-Cameroon subregion. The predominantly R → K substitution at VP1–297, A → S at VP1–323 and V → L at P2A–42 appear to play a role in the unique presentation of HAV/IA in this subregion (fig. 2). These substitutions appear to be a characteristic of this subregion. These mutations in the VP1/P2A region probably confer an advantage on the HAV/IA presence in HAV-infected individuals in this subregion. When compared to each other, all HAV/IA isolates in this study showed a variability of only 0.7% at the amino acid level in the VP1/P2A junction region. This demonstrates that although there was some nucleotide heterogeneity, there was limited amino acid heterogeneity [18]. This
provides additional support to the adaptation of this virus in the population and might explain why only one HAV serotype has been identified to date. This study reports the epidemiology of HAV in Nigeria at the molecular level. It provides essential baseline data as a reference for genetic analysis of future HAV stains in Nigeria. Further analysis using HAV isolates from the different geopolitical zones of the country may ascertain the development of a new variant reservoir and the continued circulation of HAV/IA in the Nigeria-Cameroon corridor. Parallel improvement in hygiene, sanitation, living standards, drinking water quality and adoption of regulatory measures of food handling would go a long way to interrupt HAV transmission in this sub-region.

References