Virus Research 155 (2011) 316–324

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

# First full-length genomic sequence of a hepatitis A virus isolated in Argentina shows recombination between subgenotypes IA and IB

Sebastian Aguirre[a], Viviana Malirat[a], Eduardo Scodeller[b], Nora Mattion[a,*]

[a] Centro de Virología Animal, Instituto de Ciencia y Tecnología Dr. César Milstein, CONICET, Saladillo 2468, 1440 Buenos Aires, Argentina
[b] Instituto de Medicina y Biología Experimental de Cuyo (IMBECU – CONICET), Av. Ruiz Leal s/n, 5500 Mendoza, Argentina

ARTICLE INFO

*Article history:*
Received 29 July 2010
Received in revised form 21 October 2010
Accepted 27 October 2010
Available online 4 November 2010

*Keywords:*
Hepatitis A
Full-length sequence
Recombination event
Argentine isolate

ABSTRACT

A hepatitis A virus (HAV) recovered in Argentina from a stool sample of a sick child in the year 2006 (HAV-Arg/06) was entirely sequenced. Phylogenetic analysis included the HAV-Arg/06 sequence in subgenotype IA, either considering the usual VP1-2A variable junction fragment or the full length nucleotide sequence. Interestingly, a recombination event with subgenotype IB, involving a portion of the 2C–3A nonstructural proteins coding region (nucleotides 4961–5140) was detected using specific software. Only subgenotype IA strains have been detected in Argentina or Uruguay, whereas subgenotype IA and IB strains have been reported to circulate in Brazil. Although recombination has been given an important role in the evolution of picornaviruses, there have been only a few reports of its involvement in the evolution of HAV, probably due to the limited number of complete HAV sequences available. This study constitutes the first report of a full-length HAV sequence in Argentina and the third in South America, after the sequence of the IA isolate HAV5 from Uruguay and the IB isolate HAF-203 from Brazil. The availability of new sequence data covering the complete HAV genome will help establish a more consistent genetic relatedness among HAV isolates and the role of recombination in its evolution.



## 1. Introduction

Hepatitis A virus (HAV), the causative agent of type A viral hepatitis, is a positive-strand RNA virus which has been classified within the family Picornaviridae, in the Hepatovirus genus (Hollinger and Emerson, 2007). The genome is approximately 7500 nucleotides in length and contains a single large open-reading frame (ORF) encoding a polyprotein of 2227 amino acids. The capsid proteins (VP1–VP4) represent the amino-terminal third, with the remainder of the polyprotein comprising the nonstructural proteins required for HAV RNA replication: 2B, 2C, 3A and 3B (also known as VPg, that is covalently linked to the 5′ end of the genomic RNA and that probably serves as the protein primer for RNA synthesis), 3Cpro (a cysteine protease responsible for most post-translational cleavage events within the polyprotein), and 3Dpol (the viral RNA-dependent, RNA polymerase). The 5′ non-coding region of the genomic RNA of HAV harbors an internal ribosome entry site (IRES), which directs the ribosomes to the initiation codon of the polyprotein, and the 3′ non-coding region has a short poly(A) tail (Brown et al., 1994; Glass et al., 1993; Hollinger and Emerson, 2007; Weitz et al., 1986).

As other RNA viruses, HAV exists in vivo as distributions of closely related variants referred to as quasispecies (Costa-Mattioli et al., 2006; Sanchez et al., 2003a). Quasispecies dynamics is characterized by continuous generation of variant viral genomes, competition among them, and selection of the fittest mutant distributions in any given environment. In this context, negative selection appears to be the main force of evolution of HAV, selecting against most replacement changes in the protein coding regions, which may contribute to explain, at least in part, the presence of a single conserved immunogenic neutralization site and of only one serological group of HAV worldwide (Sanchez et al., 2003b; Stapleton and Lemon, 1987).

However, despite this limited amino acid heterogeneity, a significant degree of nucleic acid variability has been observed among different isolates from different regions of the world (Costa-Mattioli et al., 2001; Robertson et al., 1992) and also in strains isolated in South America (Costa-Mattioli et al., 2002, 2001; De Paula et al., 2002; Mbayed et al., 2002).

The studies on HAV genetic variability were based on the analysis of limited fragments of the HAV genome, such as the C terminus of VP3 (Jansen et al., 1990), the N terminus of VP1 (Robertson et al., 1991) or the putative VP1-2A junction region. Using this approach, genetic analysis of 152 HAV strains recovered around the world resulted in the designation of seven genotypes of HAV (I–VII) (Robertson et al., 1992). In 2002, Costa-Mattioli et al. reported phylogenetic studies using the entire VP1 sequences (900 nucleotides),

* Corresponding author. Tel.: +54 11 4686 6225; fax: +54 11 4686 6225.
*E-mail addresses:* nmattioncevan@centromilstein.org.ar, mattion@fibertel.com.ar (N. Mattion).

0168-1702/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2010.10.030

**Table 1**
Primers used in the present study.

| Primer | Sequence 5′–3′ | Reference or source |
|---|---|---|
| 1F | TTCAAGAGGGGTCTCCGG | This study[a] |
| 240R | GGAGAGCCCTGGAAGAAAGA | Costafreda et al. (2006) |
| 68F | TCACCGCCGTTTGCCTAG | Costafreda et al. (2006) |
| 725R | AGGAATGAGGAAAAACCTAAA | García-Aguirre and Cristina (2008) |
| 661F | TTATGTGGTGTTTGCCTCTGAG | García-Aguirre and Cristina (2008) |
| 1249R | AACAAACCATGAGGATAAACT | García-Aguirre and Cristina (2008) |
| 1180F | ATTTGAGCTAGCGTGCTATGGTACCTGGTGACCA | This study |
| 2115R | GTTAGAAGGAGAGGTCAATCTG | This study |
| 2050F | GTGAGTACACTGCCATTGG | This study |
| 3225R | CATTTCCTAGGAGGTGG | García-Aguirre and Cristina (2008) |
| 3006F | CCAGAGCTCCATTGAACTC | This study |
| 3908R | AACTGAACAACCAATATCTGC | García-Aguirre and Cristina (2008) |
| 3875F | CATAATTGGTTTGTTGAGAGTC | García-Aguirre and Cristina (2008) |
| 4985R | ATGAATTCAGTCATGTTTTG | This study |
| 4977F | GACTGAATTCATGGAGTTG | This study |
| 6225R | CATATCAAGATCTAGAAAATCATC | This study |
| 5654F | AACTGTAAATGGAACCCCTAT | García-Aguirre and Cristina (2008) |
| 6636R | ATAATTCATCCCACTGTCTATC | García-Aguirre and Cristina (2008) |
| 6225F | GGGTAAGACTCAGTTAGTTGATG | This study |
| 7044R | TTGTCCAATCAAATCAAGATTATC | This study |
| 6984F | TTGTCCAATCAAATCAAGATTATC | This study |
| 15dT | TTTTTTTTTTTTTTT | This study |

F, sense polarity; R, antisense polarity.
[a] Modified from García-Aguirre and Cristina (2008).

as an alternative method. Based on those studies, a novel classification of HAV genotypes was proposed to include six different genotypes (Costa-Mattioli et al., 2003), three isolated from humans (I–III) and three from a simian origin (IV–VI). Genotypes I and III are the most prevalent genotypes isolated from humans (Cristina and Costa-Mattioli, 2007). Genotype I is most prevalent worldwide, and sub-genotype IA is more common than IB. Recently, a study carried out in France for Genotype II isolates, was based on the entire P1 and 5′ untranslated regions of IIA (Desbois et al., 2010).

Sub-genotype IA has been defined as the major HAV population in North and South America (Costa-Mattioli et al., 2001; Robertson et al., 1992). However, a recent report revealed that, as more isolates were characterized in Brazil, a greater presence of subgenotype IB became evident (De Paula et al., 2002).

Analysis of more than one hundred samples recovered from different areas and sequenced at the VP1/2A junction, clearly demonstrated that only subgenotype IA is circulating in Argentina (Mbayed et al., 2002; Munne et al., 2007). Most of the Argentine isolates grouped in at least two clusters. One of them included South American strains, thus suggesting a co-circulation of related isolates among neighbor countries, such as Argentina, Uruguay and Brazil. The other cluster was composed only of Argentine specimens, while a few sequences were also found scattered along the phylogenetic trees (Mbayed et al., 2002; Munne et al., 2007).

Clinical isolates of HAV replicate inefficiently in cell culture unless mutations are acquired throughout the genome. This has constituted a barrier to study HAV antigenic diversity by serological methods, such as seroneutralization. In the present study, the full-length sequence of an Argentine isolate was characterized and compared with other HAV complete sequences.

## 2. Materials and methods

### 2.1. Virus sample

Hepatitis A virus (HAV-Arg/06) was recovered from a 4 year-old child with no history of travel abroad, who contracted sporadic acute hepatitis A in the year 2006, in Santa Fe, Argentina. The patient was diagnosed based on clinical symptoms and positive IgM-class antibodies to HAV in blood. Presence of HAV in fecal samples (kindly provided by Dr. Juan Carlos Beltramino, "Dr. O. Alasia" Hospital, Santa Fe, Argentina) was confirmed by RT-PCR using specific primers for the 5′ UTR genomic region.

### 2.2. HAV genome amplification

One gram of the stool sample was suspended in 10 ml of PBS, clarified and precipitated with 20% (w/v) PEG 8000 and 0.2% NaCl. RNA was extracted from the concentrated stool sample using total RNA extraction kit (RBC, Bioscience, Taiwan) according to manufacturer's instructions and subjected to first strand cDNA synthesis with reverse transcriptase (AMV, Promega) and random hexamers. PCR amplification of 11 overlapped genome fragments was carried out in reactions containing a mixture of 5 μl of 10× reaction buffer, 7 μl of 2.5 mM dNTPs, 3 μl of cDNA template, 300 nM positive-sense primer, 300 nM negative-sense primer and 2.5 U of Taq DNA Polymerase (PBL, Bernal, Argentina). The reaction mixture (50 μl final volume) was subjected to 95 °C for 5 min, then 35 automated cycles of denaturation at 95 °C for 30 s, annealing at 50–55 °C for 30 s, and extension at 72 °C for 1 min. The PCR products were gel purified, ligated into pGEM®-T Vector (Promega) and used to transform competent *E. coli* DH10B cells.

### 2.3. Sequence determination, editing and alignment

Selected clones were sequenced in both directions using universal primers for T7 and SP6, as well as internal primers (Table 1). For each fragment, three recombinant plasmids were sequenced in both directions, by the dideoxymediated chain-termination method, using an automatic sequencer (ABI Prism, Perkin Elmer, Boston, MA). The sequences of the 5′ and 3′ ends were obtained as follows. The primer used for the 5′ end was a shorter version of the primer described by García-Aguirre and Cristina (2008), obtained from a consensus sequence of subgenotype IA strains, therefore it may not represent the real sequence of the first 18 nucleotides. As the genomic RNA was copied at the 3′ end using an oligo(dT), a modification of the real sequence is not expected in this region.

Chromatograms obtained for each individual reaction were checked to avoid misreading of peak dyes. Nucleotide sequences were edited on an IBM compatible personal computer with the aid of the BioEdit program, version 7.0.5 (Hall, 1999), and sequence alignments were obtained using the CLUSTALW algo-

**Table 2**
HAV isolates with full length genome sequences available used in this study.

| Strain | Genotype | Source | Place of isolation | Accession number |
|---|---|---|---|---|
| AH1 | IA | Human | Japan | AB020564 |
| AH2 | IA | Human | Japan | AB020565 |
| AH3 | IA | Human | Japan | AB020566 |
| FH1 | IA | Human | Japan | AB020567 |
| FH2 | IA | Human | Japan | AB020568 |
| FH3 | IA | Human | Japan | AB020569 |
| DL3 | IA | Human | China | AF512536 |
| LU38-wt | IA | Human | China | AF357222 |
| LY6 | IA | Human | China | AF485328 |
| H2 | IA | Human | China | EF406357 |
| LP014 | IA | Human | Thailand | EF207320 |
| HAV5 | IA | Human | Uruguay | EU131373 |
| LA | IA | Human | USA | K02990 |
| GBM-wt | IA | Human | Germany | X75215 |
| FG | IA | Human | Italy | X83302 |
| HM-175-wt | IB | Human | Australia | M14707 |
| HAF 203 | IB | Human | Brazil | AF268396 |
| L-A-1 | IB | Human | China | AF314208 |
| MBB | IB | Human | North Africa | M20273 |
| CF53/Berne | IIA | Human | France | AY644676 |
| SLF88 | IIB | Human | France | AY644670 |
| HA-JNG04-90F | IIIA | Human | Japan | AB279732 |
| HA-JNG08-92F | IIIA | Human | Japan | AB279733 |
| HAJ95-8F | IIIA | Human | Japan | AB279734 |
| NOR-21 | IIIA | Human | Norway | AJ299464 |
| PN-IND | IIIA | Human | India | EU011791 |
| SIM27 | IIIA | Human | India | FJ227135 |
| CP-IND | IIIA | Human | India | DQ991029 |
| GBS-IND | IIIA | Human | India | DQ991030 |
| HA-JNG06-90F | IIIB | Human | Japan | AB258387 |
| HAJ85-1F | IIIB | Human | Japan | AB279735 |
| AGM-27 | V | Simian | Kenya | D00924 |

rithm (Thompson et al., 1994), as implemented in the referred program.

### 2.4. Phylogenetic analysis

Unrooted trees were constructed according to sequence relatedness of nucleotides covering either the complete genome or the VP1-2A junction fragment (168 nucleotides corresponding to positions 3024–3192 of strain HM 175 wt, GenBank accession number M14707). Tree topology was inferred by the Neighbor-Joining (NJ) method using a matrix of Kimura's two-parameter (K-2p) distances, and bootstrap analysis with 1000 replicates, as implemented in the computer program MEGA, version 4 (Kumar et al., 2008), to facilitate the comparison with most already published phylogenetic studies performed with either partial or complete genomes. Additionally, and as the model that best represented our sequence data, established by the PALM package (Chen et al., 2009) is the GTR+I (General Time Reversible, modeling the non-uniformity of evolutionary rates among sites by assuming that a certain fraction of sites are evolutionarily invariable), an unrooted tree using maximum composite likelihood was also constructed. General topology of these trees was consistent with those derived from the NJ-K2p analysis, that is, the major clades and the relationships among sequences were maintained (data not shown).

### 2.5. Recombination analysis

Detection of putative recombination events was carried out by the use of the automated suite of algorithms contained within the Recombination Detection Program 3 (RDP3) (Heath et al., 2006) and checked using SimPlot software (Lole et al., 1999).

Recombination events were analyzed by the Recombination Detection Program (RDP) (Martin and Rybicki, 2000), GENECONV (Padidam et al., 1999), BOOTSCAN (Martin et al., 2005), MaxChi (Smith, 1992), CHIMAERA (Posada and Crandall, 2001), and SISCAN (Gibbs et al., 2000) methods embedded in the RDP3 package, where the default settings for each program suite were applied. RDP3 was set to report all hits detected by at least 2 algorithms. Common settings for all methods were to polish breakpoints and to check alignment consistency. Statistical significance was set at the $P = 0.05$ level, after considering Bonferroni correction for multiple comparisons, as implemented in RDP.

To test the phylogenetic congruence of the putative recombination breakpoints reported, a tree was constructed using the same settings as referred in Section 2.4.

## 3. Results

### 3.1. Full length sequence and phylogenetic analysis of the HAV-Arg/06 strain

The HAV-Arg/06 genome consisted of 7477 nucleotides, excluding the poly(A) tract at the 3′ end. The ORF, that begins in the 11th AUG (at position 733–735), encodes a polyprotein of 2227 amino acids, presenting the same length and functional regions as reference strains for IA, IB and IIB genotypes. The complete genomic sequence of isolate HAV-Arg/06 has been deposited in the GenBank database under the accession number HM769724.

The complete genome nucleotide sequence of HAV-Arg/06 isolate was compared to the HAV complete sequences available in the GenBank, including strains from different genotypes and geographical areas. Table 2 shows the list of the strains used in the analysis, their genotype, locations of isolation and their respective GenBank accession numbers.

From the phylogenetic trees generated, it became evident that the HAV-Arg/06 isolate belonged to the IA

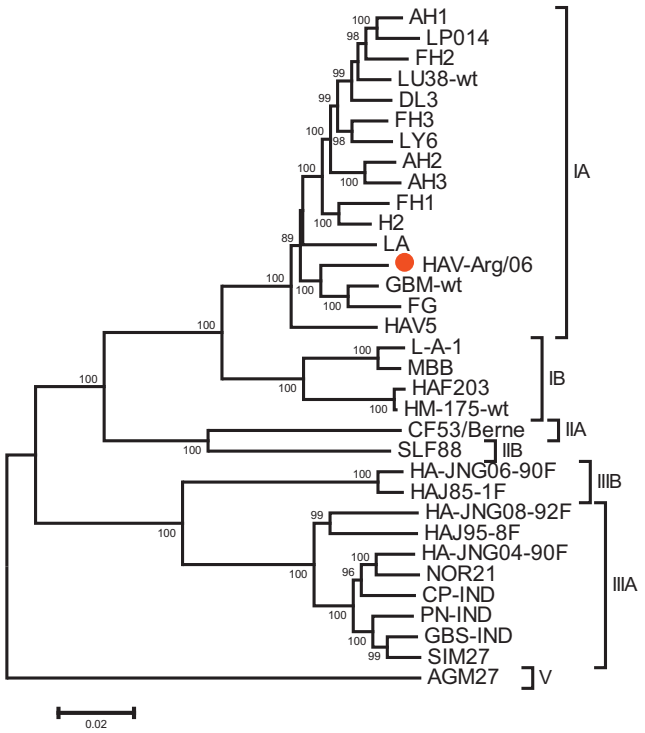


**Fig. 1.** Phylogenetic tree of HAV strains for which complete genome sequences were available. The phylogenetic relatedness between HAV-Arg/06 isolate (●) and strains listed in Table 1 was inferred using the Neighbor-Joining method. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). Bootstrap values (1000 replicates) are shown next to the branches. Phylogenetic analyses were conducted in MEGA4 (Kumar et al., 2008). Genotypes and subgenotypes are indicated.

subgenotype (Fig. 1), with a bootstrap support value of 100%.

Considering the isolates from which the complete genome sequences were available, the highest degree of nucleotide identity (97%) was recorded with the German subgenotype IA strain GBM (Graff et al., 1994).

Percent identity between HAV-Arg/06 and other selected HAV strains was calculated for both nucleotide and amino acid sequences (Table 3). As reported by others, a high identity was found at the amino acid level, particularly within genotype I sequences.

The identity of HAV-Arg/06 nucleotide sequence with the two other full-length sequences available from South America was 95.3% with Uruguayan subgenotype IA HAV5 strain (García-Aguirre and Cristina, 2008) and 91.7% with the Brazilian subgenotype IB HAF-203 strain (Baptista et al., 2006; Gaspar et al., 1992). At the amino acid level, the identity of the HAV-Arg/06 sequence was 99.0% and 98.2% with the same strains, respectively (Table 3).

Interestingly, in the analysis at the individual protein level, it was found that the nucleotide and amino acid sequences of the genomic region encoding for the 3A non-structural protein of HAV-Arg/06 were closer to subgenotype IB strains (98.6% amino acid identity with HAF-203 and HM-175 strains) than to the subgenotype IA GBM strain (93.2%) (Table 3, shown in bold). The same was observed for the Uruguayan strain HAV5. The HAF-203 strain, which was obtained from a fecal specimen of an acute hepatitis A patient, was sequenced following adaptation to culture in FRhK-4 cells. The identity between HAF-203 strain and the subgenotype IB HM-175 strain was 99.7% and 99.3% at the nucleotide or amino acid level, respectively.

**Table 4**
Average *P*-value of the recombinant event analyzed by six recombination detection methods.

| Method | Average *P*-value |
|---|---|
| RDP | $4.900 \times 10^{-2}$ |
| GENECONV | $3.061 \times 10^{-3}$ |
| BOOTSCAN | $4.663 \times 10^{-4}$ |
| MaxChi | $5.488 \times 10^{-1}$ |
| CHIMAERA | $4.662 \times 10^{-1}$ |
| SISCAN | $4.779 \times 10^{-5}$ |

**Table 3**
Nucleotide and amino acid sequence identities between HA-Arg/06 and other selected HAV strains.

| | Genomic region | GBM (IA) | HAV5 (IA) | HAF-203 (IB) | HM-175 (IB) | CF-53-Berna (IIA) | SLF-88 (IIB) | HA-JNG-08-92 (IIIA) | HA-JNG06-90F (IIIB) |
|---|---|---|---|---|---|---|---|---|---|
| Nucleotide sequence | Full length | **97.0** | **95.3** | **91.7** | **91.9** | **86.3** | **86.6** | **83.1** | **83.5** |
| | 5′ UTR | 98.7 | 96.5 | 96.2 | 96.5 | 94.2 | 94.5 | 91.9 | 91.6 |
| | VP4 | 97.1 | 98.6 | 94.2 | 94.2 | 94.2 | 92.8 | 95.7 | 95.7 |
| | VP2 | 96.3 | 94.8 | 91.7 | 90.3 | 85.1 | 86.2 | 83.2 | 83.7 |
| | VP3 | 96.6 | 95.7 | 91.3 | 91.5 | 84.8 | 85.2 | 83.6 | 84.3 |
| | VP1 | 97.0 | 94.3 | 90.0 | 90.3 | 86.1 | 84.4 | 79.9 | 81.7 |
| | 2A | 97.5 | 95.2 | 89.9 | 89.9 | 84.0 | 86.1 | 79.5 | 80.8 |
| | 2B | 96.3 | 96.0 | 89.7 | 91.0 | 87.2 | 86.0 | 81.3 | 81.9 |
| | 2C | 96.8 | 94.8 | 89.3 | 89.5 | 85.9 | 85.9 | 82.2 | 81.9 |
| | 3A | **92.3** | **94.1** | **95.0** | **95.0** | 82.9 | 84.2 | 73.4 | 76.1 |
| | 3B | 95.7 | 91.3 | 91.3 | 91.3 | 82.6 | 82.6 | 85.5 | 82.6 |
| | 3C | 97.3 | 95.9 | 92.8 | 92.8 | 85.7 | 87.2 | 83.9 | 83.6 |
| | 3D | 97.1 | 95.4 | 92.0 | 92.2 | 85.7 | 85.9 | 82.9 | 82.8 |
| | 3′ UTR | 98.4 | NA | 100.0 | 100.0 | 79.9 | NA | 95.3 | 92.2 |
| Aminoacid sequence | Polyprotein | **98.8** | **99.0** | **98.2** | **98.8** | **96.5** | **97.2** | **94.2** | **94.4** |
| | VP4 | 95.7 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| | VP2 | 99.5 | 99.5 | 99.1 | 99.1 | 100.0 | 99.5 | 98.6 | 98.2 |
| | VP3 | 100.0 | 100.0 | 99.2 | 99.6 | 98.0 | 98.8 | 98.0 | 98.0 |
| | VP1 | 99.0 | 99.0 | 97.7 | 99.0 | 99.3 | 99.0 | 95.7 | 97.3 |
| | 2A | 99.5 | 98.4 | 97.3 | 97.9 | 96.8 | 96.3 | 91.5 | 93.1 |
| | 2B | 99.1 | 100.0 | 97.2 | 100.0 | 94.4 | 98.1 | 94.4 | 94.4 |
| | 2C | 98.5 | 97.7 | 97.3 | 97.9 | 95.2 | 97.0 | 93.4 | 94.0 |
| | 3A | **93.2** | **98.6** | **98.6** | **98.6** | 90.3 | 95.9 | 83.8 | 86.5 |
| | 3B | 95.5 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| | 3C | 99.5 | 100.0 | 100.0 | 100.0 | 99.1 | 99.1 | 97.2 | 97.2 |
| | 3D | 98.6 | 98.6 | 98.2 | 98.4 | 92.7 | 93.9 | 90.6 | 90.2 |

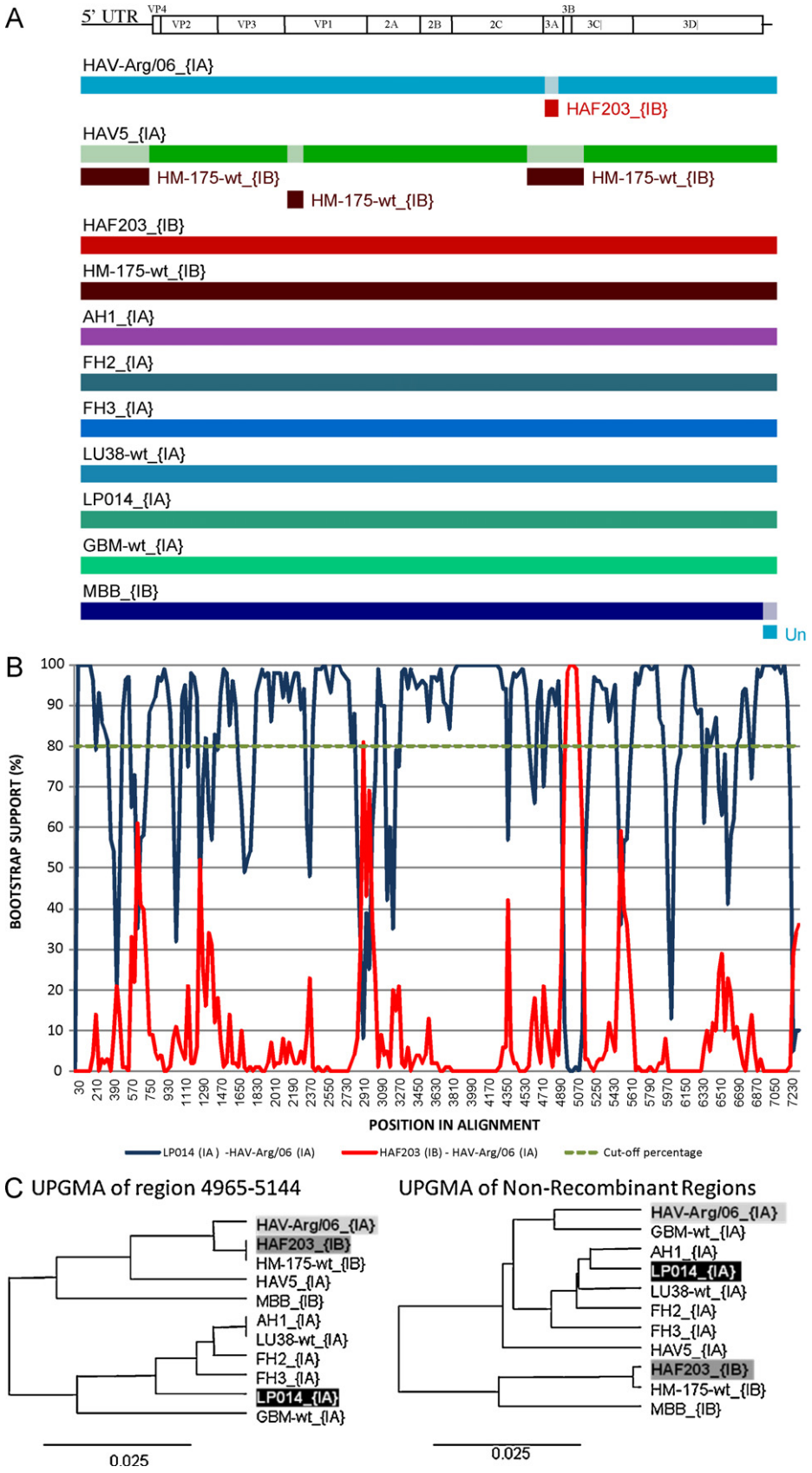


**Fig. 2.** Potential recombination event in the 3A genomic region HAV-Arg/06. (A) Identification of recombination events by the RDP3 package. (B) BOOTSCAN evidence for the recombination events and breakpoint determination calculated with a window size of 160, step size of 30, and 200 bootstrap replicates. (C) UPGMA tree derived from the RDP3 analysis, constructed using different portions of the sequences: to the right, region 4965–5144 in the alignment (corresponds to positions 4961–5140 in the HAV-Arg/06 sequence); to the left, non recombinant regions.

HAV-Arg/06 is the first complete HAV genome reported from Argentina. In order to compare the HAV-Arg/06 isolate with other local HAV strains, we used the partial sequences corresponding to the VP1-2A junction region as have been published previously (Mbayed et al., 2002; Munne et al., 2007). A Blast search was carried out with the commonly used 168 nt fragment of the VP1-2A region and a phylogenetic tree was generated (Supplementary Fig. 1A) including the 100 most closely related sequences retrieved by the Blast search (identity values of 97–99.4%), 42 of which corresponded to Argentinean isolates. HAV-Arg/06 showed the closest relationship with strains isolated previously (years 1999–2005) in Argentina, although isolates from other regions were also recorded within the cluster (HAV 073 (EU416241.1) and HAJ99-2 (AB258596.1)). Additionally another phylogenetic reconstruction was carried out adding the remaining sequences from all HAV Argentinean isolates available. HAV-Arg/06 clustered within one of the two clades (Supplementary Fig. 1B) described previously for Argentine HAV strains (Mbayed et al., 2002; Munne et al., 2007).

### 3.2. Recombination between subgenotypes IA and IB

Although the phylogenetic analysis of the complete genome of the HAV-Arg/06 isolate placed this strain within the IA subgenotype, results obtained from the partial percent identity calculated for each individual gene pointed to a potential recombination with subgenotype IB, in the genomic region coding for 3A nonstructural protein (Table 3). The Tajima's *D* statistical test performed using MEGA software for identification of a potential positive selection resulted negative (data not shown).

To further explore this possibility, and considering that the Uruguayan HAV5 strain was reported to undergo recombination with a subgenotype IB strain (Liu et al., 2010), detection of potential recombinant events was studied using the RDP3 package, as described in Section 2.5. One potentially significant recombination event was found between subgenotypes IA and IB in the genomic region encoding the 3A protein of HAV-Arg/06 (Fig. 2A). The event was pointed out by all six methods tested in the RDP3 program (Table 4), and the putative breakpoints were identified as positions 4961 (beginning) and 5140 (ending), involving also a short C-terminal portion of the 2C protein. The recombination event was described to have occurred between HAV lineages represented by the strain LP014 (subgenotype IA, Table 2) and the Brazilian isolate HAF 203 (subgenotype IB) giving rise to the recombinant HAVArg/06 isolate. To confirm the inter-subgenotypic recombination, the relevant strains were further analyzed by BOOTSCAN plot (Fig. 2B), and by UPGM tree construction (Fig. 2C) of the recombinant and of non-recombinant regions, as implemented in the RDP3 program, and also confirmed by the SimPlot software (data not shown).

Moreover a phylogenetic analysis carried out considering the complete coding region for 3A protein (positions 5001–5222) (Hollinger and Emerson, 2007) through the construction of a Neighbor-Joining unrooted tree (bootstrap analysis with 1000 replicates) as implemented in the MEGA4 software, supported the existence of the recombination event between both subgenotypes in the 3A genomic region (Fig. 3). The alignment of nucleotide sequences in the 2C–3A region also showed closer identity of HAV-Arg/06 strain with respect to HAF-203 (IB) strain than to GBM (IA) strain within the putative recombination region (Fig. 4).

In addition to the recombination event described above, the RDP3 package identified three recombination events between subgenotypes IA and IB in the Uruguayan HAV5 strain (Fig. 2A), two of which have been described previously (Liu et al., 2010).

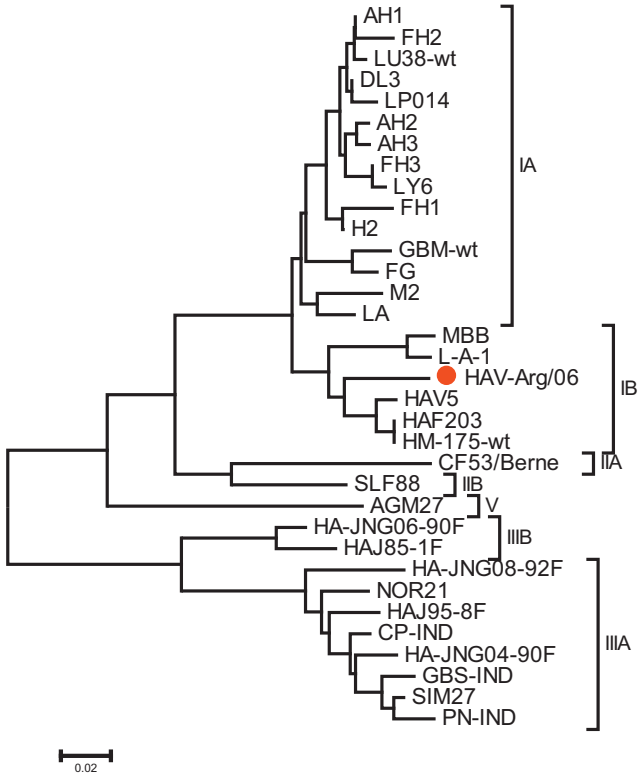


**Fig. 3.** Phylogenetic tree of the 3A region of HAV strains. Phylogenetic tree of the 3A region showing the relationships between the HAV isolate ARG2006 and other strains of different genotypes. When only 3A region is considered, isolate HAV-Arg/06 (●) shows the closest relationship with strains HAF 203 and HM175, both belonging to genotype IB, giving support to the observation of a putative recombination spot within this genomic region for isolate HAV-Arg/06. In addition to strains included in Fig. 1, the sequence of the subgenotype IA M2 strain (AY974170) of Cuban origin was considered for this analysis.

## 4. Discussion

Argentina is considered a country of intermediate endemicity (Munne et al., 2007) and universal vaccination in 1-year-old children has been implemented since the year 2005. The present study reports the first full length HAV strain (HAV-Arg/06) sequenced in Argentina. In South America there is only one other complete genome sequence of HAV subgenotype IA strain (HAV5) (García-Aguirre and Cristina, 2008). The other full length sequence corresponds to a tissue culture adapted Brazilian strain HAF-203 belonging to IB subgenotype, reported in the year 2006 (Baptista et al., 2006; Gaspar et al., 1992).

Compared with the 32 full-length sequences reported worldwide, it was clear that HAV-Arg/06 belonged to the IA subgenotype, with close identity in the VP1-2A region with other Argentine strains partially sequenced before (Mbayed et al., 2002; Munne et al., 2007).

An interesting feature reported in this study is the recombination event found in the subgenotype IA HAV-Arg/06 strain with the Brazilian subgenotype IB HAF-203 strain, with breakpoints located at nucleotides 4961–5140, involving a short C-terminal portion of 2C protein (4%) and a significant N-terminal portion of 3A protein (63%). A significant teschovirus breakpoint cluster was also reported at the 2C–3AB ORF interface, which seems to correspond with breakpoint clusters at or near the 2C–3AB ORF interface in the A, B, and C enteroviruses (Heath et al., 2006).

Genetic exchange among HAV strains in nature has been reported only for a few cases. A recombination event in the VP1 capsid protein has been reported for strain 9F94 (Costa-Mattioli

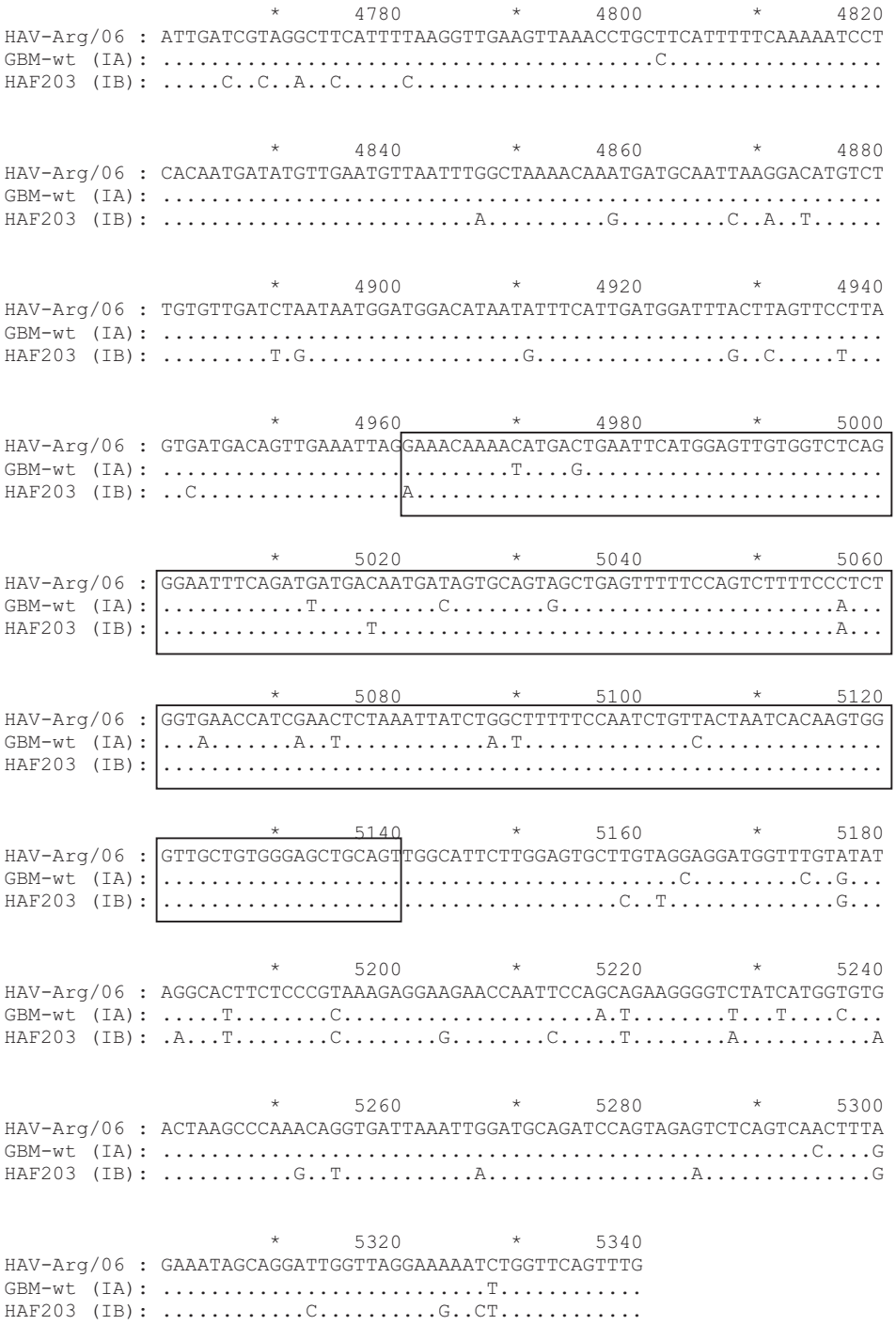


**Fig. 4.** Partial nucleotide sequence of strains HAV-Arg/06, GBM-wt (IA) and HAF-203 (IB) aligned in the region 2C/3A, showing the putative recombination portion (inside boxes) defined by the beginning and ending breakpoints (nt 4961–5140). The sequence identity of HAV-Arg/06 in the recombination region is 98.3% with IB and 93.3% with IA subgenotypes, while outside the box, in the flanking regions, the identity is 89.7% and 96.1%, respectively.

et al., 2003). Recently, recombination analyses performed on 31 complete HAV genomes from infected humans and simians found three potentially significant intra-genotypic recombination events (Liu et al., 2010). One of them corresponded to the FG Italian isolate (X83302) which might be a recombinant of subgenotypes IA and IB, with a crossover point within the 3D region. The other two corresponded to the Uruguayan strain HAV5, with two crossover points with MBB North African strain (Paul et al., 1987), one of them in the 5168–5388 nucleotides region, involving the 3′ end of the 3A protein coding region. In this study three recombinant events between HAV5 and HM-175 were detected. A possible explanation of the differences between the results found in this study and the report of Liu et al. (2010) is that they had not included in the recombination analysis the sequences of subgenotype IB strains HAF-203 and HM-175. HAV IB strains HAF203 and HM-175 have a high identity (99.7%) and both have been adapted to tissue culture.

Since young children usually have asymptomatic forms of hepatitis A, it is tempting to speculate if a recombination event in the 2C–3A genomic region might be related to the symptomatic nature of the HAV infection, considering that the Uruguayan HAV5 strain, possessing a recombination in the same region, was also isolated from a symptomatic patient (García-Aguirre and Cristina, 2008).

There are no consistent data on correlation between genotypes and symptomatic disease. However, a role of a stable intermediate precursor product of the HAV 3CPro cysteine protease, 3ABC, in the evasion of the innate immune response has been described (Yang et al., 2007). The 3ABC precursor targets to mitochondria, where it colocalizes with and cleaves mitochondrial antiviral signaling protein (MAVS), thereby disrupting activation of the interferon regulatory factor 3 (IRF3) through the MDA5 pathway. This cleavage has a major impact on the production type I interferon. The subcellular localization of this HAV polypeptide depends on the transmembrane domain of 3A, since it acts as an anchor of the 3ABC precursor in the mitochondria. It has been proposed that 3ABC-mediated proteolysis of MAVS may also contribute to the initial clinically quiescent phase of the infection and may also facilitate the ability of the virus to establish persistent noncytopathic infections in cultured cells (Yang et al., 2007). Interestingly, a deletion of three amino acids in the N-terminal portion of 3A has been involved in the induction of cytopathic effect in cell culture of an Italian HAV strain (Beneduce et al., 1995). Although not as well studied, HAV 2C has been shown to induce rearrangement of intracellular membranes that support RNA replication (Teterina et al., 1997). More data is needed for a better understanding of the impact of this genomic region on the pathogenesis of HAV infection.

At present, there are no reports of the presence of subgenotype IB strains circulating in Argentina or Uruguay. However the report of IA/IB recombination events in HAV strains from both countries (Liu et al., 2010; this report), the discovery of circulation of both subgenotypes in Brazil (De Paula et al., 2003), and the clustering of other South American strains together with subgenotype IA Argentine strains (Mbayed et al., 2002; Munne et al., 2007) are suggestive of increased co-circulation or introduction of recombinant strains from Brazil to neighbor countries.

The present study was limited by the absence of more full-length HAV sequences in the South American region and also worldwide. However, it supports previous reports, indicating that, as has been reported for other picornaviruses, recombination may act as an important factor to generate HAV diversity, with important implications for its evolution. A consistent genetic relatedness will only be inferred on the basis of a more extensive sequencing of HAV isolates.

## Acknowledgements

We thank Miss. Silvia Rojana for her technical assistance. This work was partially supported by FONCyT PID2005-35996.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.virusres.2010.10.030.

## References

Baptista, M.L., Silva, M., de Lima, M.A., Yoshida, C.F.T., Gaspar, A.M., Galler, R., 2006. Genetic variability of hepatitis A virus strain HAF-203 isolated in Brazil and expression of the VP1 gene in *Escherichia coli*. Mem. Inst. Oswaldo Cruz 101, 759–766.

Brown, E.A., Zajac, A.J., Lemon, S.M., 1994. In vitro characterization of an internal ribosomal entry site (IRES) present within the 5′ nontranslated region of hepatitis A virus RNA: comparison with the IRES of encephalomyocarditis virus. J. Virol. 68, 1066–1074.

Beneduce, F., Pisani, G., Divizia, M., Panh, A., Morace, G., 1995. Complete nucleotide sequence of a cytopathic hepatitis A virus strain isolated in Italy. Virus Res. 36, 299–309.

Chen, S.-H., Su, S.-Y., Lo, C.-Z., Chen, K.-H., Huang, T.-J., Kuo, B.-H., Lin, C.-Y., 2009. PALM: a paralleled and integrated framework for phylogenetic inference with automatic likelihood model selectors. PLoS ONE 4 (12), e8116.

Costafreda, I.M., Bosch, A., Pinto, R., 2006. Development, evaluation, and standardization of a real-time TaqMan reverse transcription-PCR assay for quantification of hepatitis A virus in clinical and shellfish samples. Appl. Environ. Microbiol. 72, 3846–3855.

Costa-Mattioli, M., Cristina, J., Romero, H., Perez-Bercof, R., Casane, D., Colina, R., Garcia, L., Vega, I., Glikman, G., Romanowsky, V., Castello, A., Nicand, E., Gassin, M., Billaudel, S., Ferre, V., 2002. Molecular evolution of hepatitis A virus: a new classification based on the complete VP1 protein. J. Virol. 76, 9516–9525.

Costa-Mattioli, M., Domingo, E., Cristina, J., 2006. Analysis of sequential hepatitis A virus strains reveals coexistence of distinct viral subpopulations. J. Gen. Virol. 87, 115–118.

Costa-Mattioli, M., Ferre, V., Casane, D., Perez-Bercoff, R., Coste-Burel, M., Imbert-Marcille, B.M., Andre, E.C., Bressollette-Bodin, C., Billaudel, S., Cristina, J., 2003. Evidence of recombination in natural populations of hepatitis A virus. Virology 311, 51–59.

Costa-Mattioli, M., Ferre, V., Monpoeho, S., Garcia, L., Colina, R., Billaudel, S., Vega, I., Perez-Bercoff, R., Cristina, J., 2001. Genetic variability of hepatitis A virus in South America reveals heterogeneity and co-circulation during epidemic outbreaks. J. Gen. Virol. 82, 2647–2652.

Cristina, J., Costa-Mattioli, M., 2007. Genetic variability and molecular evolution of Hepatitis A virus. Virus Res. 127, 151–157.

De Paula, V.S., Baptista, M.L., Lampe, L., Niel, C., Gaspar, A.M., 2002. Characterization of hepatitis A virus isolates from sub-genotypes IA and IB in Rio de Janeiro, Brazil. J. Med. Virol. 66, 22–27.

De Paula, V.S., Saback, F.L., Gaspar, A.M., Niel, C., 2003. Mixed infection of a child care provider with hepatitis A virus isolates from subgenotypes IA and IB revealed by heteroduplex mobility assay. J. Virol. Methods 107, 223–228.

Desbois, D., Couturier, E., Mackiewicz, V., Graube, A., Letort, M.-J., Dussaix, E., Roque-Afonso, A.-M., 2010. Epidemiology and genetic characterisation of Hepatitis A virus genotype IIA. J. Clin. Microbiol., doi:10.1128/JCM.00667-10.

García-Aguirre, L., Cristina, J., 2008. Analysis of the full-length genome of hepatitis A virus isolated in South America: heterogeneity and evolutionary constraints. Arch. Virol. 153, 1473–1478.

Gaspar, A.M.C., Vitral, C.L., Yoshida, C.F.T., Schatzmayr, H.G., 1992. Primary isolation of a Brazilian strain of hepatitis A virus (HAF-203) and growth in a primate cell line (FRhK-4). Braz. J. Med. Biol. Res. 25, 697–705.

Gibbs, M.J., Armstrong, J.S., Gibbs, A.J., 2000. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics 16, 573–582.

Glass, M.J., Jia, X.Y., Summers, D.F., 1993. Identification of the hepatitis A virus internal ribosome entry site: in vivo and in vitro analysis of bicistronic RNAs containing the HAV 5 noncoding region. Virology 193, 842–852.

Graff, J., Normann, A., Feinstone, S.M., Flehmig, B., 1994. Nucleotide sequence of wild-type hepatitis A virus GBM in comparison with two cell culture-adapted variants. J. Virol. 68, 548–554.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Heath, L., van der Walt, E., Varsani, A., Martin, D.P., 2006. Recombination patterns in aphthoviruses mirror those found in other picornaviruses. J. Virol. 80, 11827–11832.

Hollinger, F.B., Emerson, S.U., 2007. Hepatitis A virus. In: Knipe, D.M., Howley, P.M. (Eds.), Field's Virology, 5th ed. Lippincott/Williams & Wilkins, Philadelphia, pp. 911–947.

Jansen, R.W., Siegl, G., Lemon, S.M., 1990. Molecular epidemiology of human hepatitis A virus defined by an antigen-capture polymerase chain reaction method. Proc. Natl. Acad. Sci. U. S. A. 87, 2867–2871.

Kumar, S., Nei, M., Dudley, J., Tamur, K., 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinformatics 9, 299–306.

Liu, W., Zhai, J., Liu, J., Xie, J., 2010. Identification of recombination between subgenotypes IA and IB of hepatitis A virus. Virus Genes 40, 222–224.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. 73, 152–160.

Martin, D.P., Posada, D., Crandall, K.A., Williamson, C., 2005. A modified BOOTSCAN algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Res. Hum. Retroviruses 21, 98–102.

Martin, D., Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16, 562–563.

Mbayed, V.A., Sookoian, S., Alfonso, V., Campos, R.H., 2002. Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina. J. Med. Virol. 68, 168–174.

Munne, M.S., Vladimirsky, S., Moreiro, R., Ciocca, M., Cuarterolo, M., Otegui, L., Soto, S., Brajterman, L., Castro, R., Sasbon, J., Gianivelli, S., Buamscha, D., Quarleri Gonzalez, J.E., 2007. Molecular characterization of hepatitis A virus in children with fulminant hepatic failure in Argentina. Liver Int. 28, 47–53.

Padidam, M., Sawyer, S., Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265, 218–225.

Paul, A.V., Tada, H., von der Helm, K., Wissel, T., Kiehn, R., Wimmer, E., Deinhardt, F., 1987. The entire nucleotide sequence of the genome of human hepatitis A virus (isolate MBB). Virus Res. 8, 153–171.

Posada, D., Crandall, K.A., 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. U. S. A. 98, 13757–13762.

Robertson, B.H., Jansen, R.W., Khanna, B., Totsuka, A., Nainan, O.V., Siegl, G., Widell, A., Margolis, H.S., Isomura, S.A., Ito, K.I., et al., 1992. Genetic relatedness of hepatitis

A virus strains recovered from different geographical regions. J. Gen. Virol. 73, 1365–1377.
Robertson, B.H., Khanna, B., Nainan, O.V., Margolis, H.S., 1991. Epidemiologic patterns of wild-type hepatitis A virus determined by genetic variation. J. Infect. Dis. 163, 286–292.
Sanchez, G., Bosch, A., Gomez-Mariano, G., Domingo, E., Pinto, R.M., 2003a. Evidence for quasispecies distributions in the human hepatitis A virus genome. Virology 315, 34–42.
Sanchez, G., Bosch, A., Pinto, R.M., 2003b. Genome variability and capsid structural constraints of hepatitis A virus. J. Virol. 77, 452–459.
Smith, J.M., 1992. Analyzing the mosaic structure of genes. J. Mol. Evol. 34, 126–129.
Stapleton, J.T., Lemon, S.M., 1987. Neutralization escape mutants define a dominant immunogenic neutralization site on hepatitis A virus. J. Virol. 61, 491–498.
Teterina, N.L., Bienz, K., Egger, D., Gorbalenya, A.E., Ehrenfeld, E., 1997. Induction of intracellular membrane rearrangements by HAV proteins 2C and 2BC. Virology 237, 66–77.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Weitz, M., Baroudy, M., Maloy, W.L., Ticehurst, J.R., Purcell, R.H., 1986. Detection of a genome-linked protein (VPg) of hepatitis A virus and its comparison with other picornaviral VPgs. J. Virol. 60, 124–130.
Yang, Y., Liang, Y., Qu, L., Chen, Z., Yi, M.-K., Li, K., Lemon, S.M., 2007. Disruption of innate immunity due to mitochondrial targeting of a picornaviral protease precursor. Proc. Natl. Acad. Sci. U. S. A. 104, 7253–7258.