Genotyping of hepatitis A virus detected in bivalve shellfish in Galicia (NW Spain)
C. F. Manso, D. Polo, M. L. Vilarin˜o and J. L. Romalde

ABSTRACT
Hepatitis A virus (HAV) represents a significant public health problem due to its high persistence in the environment and its transmission through contaminated water and food. Bivalve shellfish are filter feeders that can bioaccumulate human pathogens found in contaminated waters, their consumption being a potential cause of hepatitis A outbreaks. In this work, cultured and wild bivalve shellfish from the Rı́a de Vigo (Galicia, NW Spain) were analysed for the presence and genotyping of HAV. A total of 160 shellfish samples were collected between March 2004 and December 2006, including 68 samples from cultured mussels (Mytilus galloprovincialis), 30 from wild clams (Rupitapes decussatus), 31 from wild cockles (Cerastoderma edule) and 31 from wild mussel. HAV detection, carried out by quantitative RT-PCR, was positive for 29 (42.6%) cultured and 40 (43.5%) wild samples, with levels ranging from $3.1 \times 10^2$ and $1.4 \times 10^{10}$ RNA copies/g of shellfish digestive tissue. The phylogenetic analysis of VP1-P2A and VP3-VP1 regions, separately or as concatenated sequences, revealed that all HAV strains analysed belong to subgenotype IB. These results indicate a high prevalence of this subgenotype in the area studied.

Key words | bivalve shellfish, genotyping, hepatitis A virus, phylogeny

INTRODUCTION
Hepatitis A virus (HAV) is one of the major causative agents of acute hepatitis worldwide (Hollinger & Emerson 2001). The clinical manifestation of HAV infection in humans may vary greatly, ranging from asymptomatic infection to fulminant hepatitis. The virus is transmitted primarily via the faecal-oral route, and represents a significant public health problem in many countries due to persistent circulation of the virus in the environment and contamination of food and water (Hussain et al. 2005). Bivalve shellfish are filter feeders that can bioaccumulate human pathogens found in contaminated waters, their consumption being a potential cause of hepatitis A outbreaks (Lees 2000). Many of the commercial species are common inshore estuaries or similar shallow or drying areas where nutrient levels are high and the waters are sheltered. However, growing waters are also frequently contaminated with human sewage. In the process of filter-feeding, bivalve shellfish may also concentrate and retain human pathogens derived from such sewage contamination (Le Guyader et al. 2000; Romalde et al. 2002). The hazards posed by bioaccumulation of harmful microorganisms are compounded by the traditional consumption of certain species raw, or only lightly cooked, and the consumption of the whole animal including the viscera. These circumstances are largely unique to bivalve shellfish and they, therefore, represent a special case among the microbial hazards associated with food.

HAV is a non-enveloped 7.5 Kb positive-stranded RNA virus that belongs to the Picornaviridae family, being the only member of genus Hepatovirus (Melnick 1992). All the viral proteins are encoded by a single open reading frame (ORF). This ORF of 2,227 amino acids is organised into three functional regions termed P1, P2 and P3. The P1 region encodes the structural proteins (VP1, VP2, VP3 and the
putative VP4), P2 and P3 regions encode the non-structural proteins associated with replication. The ORF is preceded by a 5' untranslated region (UTR) and is followed by a 3' UTR with a short poly(A) tail (Hollinger & Emerson 2001).

Only one serotype of HAV has been identified worldwide (Robertson et al. 1992). Despite this low amino acidic heterogeneity, there is substantial sequence variability between HAV isolates from different part of the world. This fact allows the classification of HAV strains into different genotypes. Based on sequencing VP1-P2A region, different HAV strains have been classified into seven different genotypes designed I-VII, which are distinguishable by 15% sequence diversity (Robertson et al. 1992). Genotypes I and III have been further divided into subgenotypes A and B. Viruses from four of the genotypes (I, II, III and VII) were detected from cases of hepatitis A in humans, whereas viruses from the other three genotypes (IV, V and VI) were isolated only from animal species (Byum et al. 2001).

Galicia (NW Spain), due to its particular geography presenting a number of big estuaries with high primary productivity, is a natural area specially suited to shellfish production. In fact, it is one of the main world producers of mussels, cultured on floating rafts, with annual productions of approximately 250,000 metric tonnes (mt) (FAO 2000). Although some works have been published on the prevalence of enteric viruses, including HAV, in shellfish from this area (Romalde et al. 2002; Vilarin˜o et al. 2009) or from some Galician rural human populations (González-Quintela et al. 2005), to our knowledge, no studies have been performed to determine the predominant genotypes in this geographical area.

Therefore, the aims of this study were to determine the prevalence of HAV in wild and cultured populations of bivalve shellfish from the Ría de Vigo (Galicia) and the genotypic characterisation of the strains detected by the analysis of VP1-P2A and VP3-VP1 regions of the HAV genome.

MATERIAL AND METHODS

Samples

A total of 160 bivalve shellfish samples collected in the Ría de Vigo (Galicia, NW Spain)(Figure 1) between March 2004 and December 2006 were analysed: 68 (42.5%) of this sample were cultured mussels (Mytilus galloprovincialis), 31 (19.3%) were wild mussels, 30 (18.7%) were clams (Ruditapes decussatus) and 31 (19.3%) were cockles (Cerastoderma edule).

RNA purification

Shellfish (10 individuals from mussel samples or 20 individuals from clam and cockle samples) were cleaned and separated from their valves in sterile conditions. The hepatopancreas was dissected and homogenised with a volume of 0.1% peptone water pH 7.5 using a blender during 60 sec. The homogenate was centrifuged 1,000 g during 5 min to recover the supernatant. The RNA extraction was carried out with the Total Quick RNA Cells and Tissues mini preparation Kit (Talent) (Ribao et al. 2004).

HAV detection

HAV detection was carried out by quantitative RT-PCR (qRT-PCR) with TaqMan probe (HAV150). For this purpose a fragment of 173 bp defined by primers HAV240 and HAV68 was selected. These primers are in positions 241-222 and 68-85, respectively (Costafreda et al. 2006)(Table 1).

The quantitative RT-PCR was carried with platinum Quantitative RT-PCR Thermoscript One-Step System Kit (Invitrogen), following the manufacturer’s recommendations. The reverse-transcription step was performed at 45°C for 1h. The amplification conditions were 40 cycles of 94°C for 1min, 55°C for 1 min and 72°C for 30 sec in a iCycler IQ Real-Time detection system (BioRad).

To know the extraction efficiencies a mutant non virulent infective strain VMC0 was employed as a control for the process of nucleic acid extraction (Costafreda et al. 2006). Prior to the viral RNA extraction for the shellfish homogenates, these were spiked with a known amount of vMC0. Viral RNA extracted from shellfish was tested undiluted and ten-fold diluted to evaluate the effect of RT-PCR inhibitors. Extraction efficiency values were evaluated by comparing the Ct value for the
mengovirus-positive amplification control with the Ct value for the tested virus, and was classified as poor (<1%), acceptable (1 to 10%), or good (>10%) (Da Silva et al. 2007).

To test the presence of RT-PCR inhibitors and calculate the qRT-PCR efficiency, co-amplification of 2.5 μL of each extracted RNA with 2.5 μL containing 10^3 copies of internal controls for HAV was performed (Costafreda et al. 2006). Thus, the Ct value of a sample mixed with internal controls was compared with the Ct value of internal control alone. Efficiency values were classified in the same three categories as extraction efficiency (poor, acceptable and good) (Da Silva et al. 2007).

The number of viral RNA copies present in positive samples was estimated using standard curves generated from RNA transcripts as previously described (Costafreda et al. 2006; Da Silva et al. 2007).

**RT-PCR and seminested PCR**

Purified RNA of those positive samples for HAV detection was used to generate a cDNA with M-MLV Reverse Transcriptase (Invitrogen). Complementary DNA was synthesised at 48°C for 45 min, using 5 μL of purified RNA and 100 u of M-MLV Reverse Transcriptase in 20 μL
reaction mixture containing 1.25 mM dNTP and 25 ng of reverse primer.

To amplify the VP1-P2A junction region two rounds of PCR were subsequently performed with Immolase DNA polymerase (Bioline). For the first round, outer forward and reverse primers were used. For the second round a seminested PCR was carried out with outer forward primer and inner reverse primer. The sequences and positions of the primers are in Table 1 (Pintó et al. 2007; Yun et al. 2008). The amplification conditions for the first round of PCR were: an initial denaturation of 94°C for 5 min, 40 cycles of 94°C for 45 sec, 57°C for 1 min and 72°C for 1 min and a final elongation of 72°C 10 min. The seminested PCR were performed under the following conditions: a initial denaturation of 94°C for 5 min, 40 cycles of 94°C for 45 sec, 55°C for 1 min and 72°C for 1 min and a final elongation of 72°C 10 min.

Amplification products were analysed by electrophoresis in agarose gel 1.5% and visualised after etidium bromide staining using a Gel-Doc apparatus (Bio-Rad).

Table 1 | Primers and probes used for HAV detection and amplification of the VP1-P2A and VP3-VP1 regions

<table>
<thead>
<tr>
<th>Primer/probe</th>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAV240</td>
<td>5’- GGA GAG CCC TGG AAG AAA GA- 3’</td>
<td>222–241</td>
</tr>
<tr>
<td>HAV68</td>
<td>5’- TCA CCG CCG TTT GCC TGA-3’</td>
<td>68–85</td>
</tr>
<tr>
<td>HAV150</td>
<td>6-FAM-CCT GAA CCT GCA GGA ATT AA-MGB</td>
<td></td>
</tr>
<tr>
<td>PCR VP1-P2A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR-9</td>
<td>5’-AGT CAC ACCTCT CAA GGA AAA ACT TT-3’</td>
<td>3310–3286</td>
</tr>
<tr>
<td>BR-5</td>
<td>5’-TTG TCT GTC ACA GAA CAA TCA G-3’</td>
<td>2950–2972</td>
</tr>
<tr>
<td>SEMINESTED PCR VP1-P2A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR-6</td>
<td>5’-AGG AGG TGG AAG CAC TTC ATT TGA-3’</td>
<td>3217–3193</td>
</tr>
<tr>
<td>BR-5</td>
<td>5’-TTG TCT GTC ACA GAA CAA TCA G-3’</td>
<td>2950–2972</td>
</tr>
<tr>
<td>PCR VP3-VP1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAV2</td>
<td>5’-CAG GAA ATG TCT CAG GTA CTT TAT CT-3’</td>
<td>2415–2391</td>
</tr>
<tr>
<td>HAV1</td>
<td>5’-GCT CCT CTT TAT CAT GCT ATG GAT-3’</td>
<td>2172–2196</td>
</tr>
<tr>
<td>SEMINESTED PCR VP3-VP1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAV2</td>
<td>5’-CAG GAA ATG TCT CAG GTA CTT TCT CT-3’</td>
<td>2415–2391</td>
</tr>
<tr>
<td>HAV3</td>
<td>5’-ATG TTA CTA CAC AAG TTT GAG AT-3’</td>
<td>2195–2218</td>
</tr>
</tbody>
</table>

Sequencing of VP1-P2A and VP3-VP1 regions

Before sequencing, PCR were purified with QIAQuick Gel Extraction Kit (Qiagen). Sequencing of the fragments VP1-P2A and VP3-VP1 were carried out using primers BR6 and HAV4 respectively and the GenomeLab Dye Terminator Cycle Sequencing-Quick Start Kit. The amplification conditions for both fragments were 40 cycles of 96°C for 20 sec, 50°C for 20 sec and 60°C for 4 min. The sequence analysis was performed in an Automatic DNA Sequencer, model 371A (Beckman Coulter).

For the reference strains, sequences were retrieved from Genbank (Access Codes: DO0924, AY644676, M10035, M59286, X83302, M66695, X75215, AB258387, M14707, KK02990, M20273, AL299464, M34084, AY644670). Sequence analysis was performed with the DNASTAR Lasergene SEQMAN program. Phylogenetic trees were constructed by using the Neighbour-joining algorithm (Yun et al. 2008). Distance matrices were calculated by Kimura’s two parameters correction and stability of the groupings was estimated by bootstrap analysis (1,000 replicates) using the program MEGA version 4.0.

RESULTS AND DISCUSSION

HAV detection in bivalve shellfish samples

HAV was detected in 69 (43.12%) of the 160 bivalve shellfish samples collected between March 2004 and December 2006 from the Ría de Vigo (NW Spain).
Sixty-eight of the analysed samples were cultured mussels from an area designated as B (230-4,600 E. coli/100g shellfish tissue) and 29 (42.6\%) of them were positive for HAV. The remaining 92 samples were clams, cockles and wild mussels from an area designated as C (>4,600 E. coli/100g shellfish tissue). HAV was detected in 40 (43.5\%) of these samples. Twelve (38.7\%) of the 31 clams samples analysed were positive for HAV. In the case of cockles, HAV was detected in 11 (35.4\%) of the 31 samples analysed, whereas in the 30 wild mussels samples HAV was detected in 17 (56.6\%) (Table 2). The percentages of positive samples for HAV were similar to those obtained in other studies employing molecular detection procedures (Le Guyader et al. 2000; Romalde et al. 2002). In addition, the similar HAV prevalence obtained for cultured and wild samples reinforces the idea that bacterial organisms are not good indicators of viral contamination (Le Guyader et al. 2000; Lees 2000; Romalde et al. 2002; Vilarino et al. 2009).

HAV levels ranged from $3.1 \times 10^2$ to $1.4 \times 10^{10}$ RNA copies/g of shellfish digestive tissue (Table 2). In general, cultured samples showed lower HAV levels, since 60\% of them presented values lower than $1 \times 10^3$ RNA copies/g of digestive tissue, whereas 96.9\% of wild samples showed values higher than $1.9 \times 10^5$ RNA copies/g of digestive tissue (data not shown).

**Table 2** | Results obtained for HAV detection and quantification in the bivalve shellfish samples from the Ría de Vigo. Classification of harvesting areas according to the current EU legislation (Anon. 2004). Samples from area classified as B are cultured samples, while those from area C are wild samples

<table>
<thead>
<tr>
<th>Total samples</th>
<th>Area classification</th>
<th>Species</th>
<th>No. of samples</th>
<th>Samples positive for HAV (%)</th>
<th>RNA copies/g of digestive tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
<td>B</td>
<td>Mussel (Mytilus galloprovincialis)</td>
<td>68</td>
<td>29 (42.6)</td>
<td>$3.1 \times 10^2$–1.4 $\times 10^{10}$</td>
</tr>
<tr>
<td>92</td>
<td>C</td>
<td>Clam (Ruditapes decussatus)</td>
<td>31</td>
<td>12 (38.7)</td>
<td>$2.8 \times 10^1$–3.9 $\times 10^8$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cockle (Cerastoderma edule)</td>
<td>31</td>
<td>11 (35.4)</td>
<td>$2.4 \times 10^3$–1.2 $\times 10^9$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mussel (Mytilus galloprovincialis)</td>
<td>30</td>
<td>17 (56.6)</td>
<td>$8.2 \times 10^2$–3.8 $\times 10^8$</td>
</tr>
</tbody>
</table>

Genetic analysis of HAV strains

Sequence analysis of the VP1-P2A and VP3-VP1 regions enables HAV strains to be distributed into seven genotypes (Robertson et al. 1992). Despite the limited amino acid heterogeneity of HAV, a significant degree of nucleotide variability has been observed among strains from different parts of the world. Therefore, the molecular characterisation of HAV strains allows the elucidation of their geographical origin and their transmission patterns.

Phylogenetic analysis of the VP1-P2A junction sequences revealed that both the 30 strains detected in cultured mussels and the 41 strains from wild mussels, clams and cockles belong to genotype IB. The strains were analysed according to the type of sample in which they were isolated. Although these strains are the same subgenotype, there are nucleotide differences among them. The similarity percentage within the VP1-P2A sequences ranges from 85.9\% to 100\%. The phylogenetic trees showed this divergence, enabling the establishment of different clusters although with low bootstrap values (Figure 2). According to the bootstrap values, the analysis of the VP1-P2A junction seems to show better the differences between genotypes and subgenotypes than the differences between clusters. These results, in agreement with other studies, suggest a high prevalence of genotype I (Rodrigues et al. 2007).

The VP3-VP1 region was also sequenced to carry out its phylogenetic analysis. The result of this analysis confirmed that all the strains detected in bivalve shellfish from the Ría de Vigo (NW Spain) belong to subgenotype IB. In this case, similarity percentages among the VP3-VP1 sequences range from 85.7\% to 100\%. Although these percentages are very similar to those obtained for the VP1-P2A junction, the phylogenetic analysis of the VP3-VP1 region reflects better the differences between our strains. Although the bootstrap values that define the branches of genotype and subgenotype are less significant than those of the region VP1-P2A, the clusters of the HAV strains established on the basis of the VP3-VP1 region showed higher bootstrap values (Figure 3).

To complete the analysis, the phylogenetic reconstruction based on concatenated VP1-P2A and VP3-VP1
Figure 2 | Phylogenetic tree of the detected HAV strains analysing the VP1-P2A region using the NJ algorithm. A) samples from cultured mussels. B) Samples from wild shellfish.

(Sample code: [A] cultured mussel site A; [B] cultured mussel site B; [C] wild mussel; [D] clam; [E] cockle). Only bootstrap values higher than 50% are shown.
Figure 3 | Phylogenetic tree of the detected HAV strains analysing the VP3-VP1 region using the NJ algorithm. A) samples from cultured mussels. B) Samples from wild shellfish. (Sample codes as in Figure 2). Only bootstrap values higher than 50% are shown.
Figure 4 | Phylogenetic tree based on concatenated VP1-P2A and VP3-VP1 sequences NJ algorithm. A) samples from cultured mussels. B) Samples from wild shellfish. (Sample codes as in Figure 2). Only bootstrap values higher than 50% are shown.
sequences was performed (Figure 4). As expected, results confirmed that all of them belong to subgenotype IB. Although the similarity percentages among the sequences range from 85.7% to 100%, the same values as for the region VP3-VP1, and although the analysed sequences have a greater length, phylogenetic trees obtained are not robust, showing very low bootstrap values.

Within this genotype, the genetic variability is higher in subgenotype IB (Cristina & Costa-Mattioli 2007). A co-circulation of subgenotypes IA and IB exists in Europe, the former being more common (Cristina & Costa-Mattioli 2007). However, all HAV strains analysed from Galicia belonged to subgenotype IB, which may suggest an endemic circulation of HAV IB population in this region. Further studies including clinical and sewage samples would help to determine the epidemiological pattern of HAV in the studied area.

On the other hand, the exact mechanism for virus uptake and concentration by shellfish is unknown, but there is the possibility that certain genotypes of a virus have more affinity for bivalve shellfish than others. This is the case of genotype GI of Norovirus (NoV) (Tian et al. 2006). HBGAS (Histo-Blood Group Antigen) have been recognized as receptors for NoV in humans (Hutson et al. 2004). Molluscs such as oysters present type A-like HBGA in gastrointestinal cells (Tian et al. 2006) and co-localised with the binding site for GI NoV. In the case of HAV, a similar situation might exist for the genotype IB. Further studies are needed to confirm this hypothesis.

CONCLUSIONS

All the 69 strains detected in shellfish from the Ría de Vigo belong to subgenotype IB, in agreement with other studies, that suggest a high prevalence of genotype I in Spain.

Some genetic variability among the analysed strains within the subgenotype IB was observed, which was better established by analysis of the region VP3-VP1. The phylogenetic analysis of VP1-P2A junction shows the differences between genotypes and subgenotypes better than the differences between clusters.

Another work suggested a co-circulation of both IA and IB subgenotypes, but HAV strains analysed from Galicia belonged to subgenotype IB. Our results therefore may suggest an endemic circulation of HAV IB population in this region.

On the other hand, the exact mechanism for virus uptake and concentration by shellfish is unknown, but there is the possibility that certain genotypes of a virus have more affinity for bivalve shellfish than others.

ACKNOWLEDGEMENTS

This work was supported in part by contracts 2007/CP746 and 2008/CP776, from the Instituto Tecnolóxico para o Control do Medio Maríño (INTECMAR), Consellería do Mar, Xunta de Galicia (Spain).

REFERENCES


