Rediscovery and genomic characterization of bovine astroviruses

Herman Tse,1,2,3† Wan-Mui Chan,1† Hoi-Wah Tsoi,1 Rachel Y. Y. Fan,1 Candy C. Y. Lau,1 Susanna K. P. Lau,1,2,3 Patrick C. Y. Woo1,2,3 and Kwok-Yung Yuen1,2,3

Correspondence
Kwok-Yung Yuen
hkumicro@hkucc.hku.hk

Received 26 January 2011
Accepted 18 April 2011

1Department of Microbiology, The University of Hong Kong, Hong Kong SAR
2Research Centre of Infection and Immunity, The University of Hong Kong, Hong Kong SAR
3State Key Laboratory of Emerging Infectious Diseases, Hong Kong Special Administration Region, Hong Kong SAR

The genus Mamastrovirus belongs to the family Astroviridae and consists of at least six members infecting different mammalian hosts, including humans, cattle and pigs. In recent years, novel astroviruses have been identified in other mammalian species like roe deer, bats and sea lions. While the bovine astrovirus was one of the earliest astroviruses to have been studied, no further research has been performed recently and its genome sequence remains uncharacterized. In this report, we describe the detection and genomic characterization of astroviruses in bovine faecal specimens obtained in Hong Kong. Five of 209 specimens were found to be positive for astrovirus by RT-PCR. Two of the positive specimens were found to contain sequences from two different astrovirus strains. Complete genome sequences of approximately 6.3 kb in length were obtained for four strains, which showed similar organization of the genome compared to other astroviruses. Phylogenetic analysis confirmed their identities as members of the genus Mamastrovirus, and showed them to be most closely related to the Capreolus capreolus astrovirus. Based on the pairwise genetic distances among their full-length ORF2 sequences, these bovine astroviruses may be assigned into at least three different genotype species. Sequence analysis revealed evidence of potential recombination in ORF2. In summary, we report the first genome sequences of bovine astroviruses and clearly establish the species status of the virus. Additionally, our study is among the first to report co-infection by different astrovirus genotypes in the same host, which is an essential step for recombination to occur.

INTRODUCTION

Astroviruses are a group of small, non-enveloped RNA viruses with an icosahedral capsid of 27–30 nm in diameter. Their name originated from the Greek word ‘astron’ for the characteristic five- or six-pointed star-shaped appearance under the electron microscope. Each virus contains a linear, non-segmented, positive-sense ssRNA genome of around 6800–7900 nt in length. Each genome contains a 5’ UTR, three ORFs, a 3’ UTR and a poly(A) tail. The first two ORFs, ORF1a and ORF1b, also form part of an extended transcript containing a translational frameshift. Unlike coronaviruses, this programmed −1 ribosomal frameshift is not mediated by an RNA pseudoknot, but by a ‘shifty heptamer’ and downstream stem–loop structure (Lewis & Matsui, 1995, 1996). ORF1a encodes a polyprotein that is cleaved to form several non-structural proteins, while ORF1b encodes an additional RNA-dependent RNA polymerase (RdRP) (Kiang & Matsui, 2002). ORF2 encodes a polyprotein that undergoes proteolytic cleavage to form mature structural capsid proteins (Bass & Qiu, 2000; Méndez et al., 2002).

Under the current International Committee on Taxonomy of Viruses system, members of the family Astroviridae are placed into one of two genera, Avastrovirus or Mamastrovirus, which infect birds and mammals, respectively. Astroviruses are known for their association with diarrhoeal disease in different animals, though selected species can cause extraintestinal illness, such as fatal hepatitis in ducklings and acute interstitial nephritis in chickens (Fu et al., 2009; Imada et al., 2000; Méndez & Arias, 2007). Human astrovirus (HAstV) infections are common and have a global distribution, with higher

†These authors contributed equally to this paper.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are HQ916313–HQ916317 and JF796126–JF796127.

Supplementary material is available with the online version of this paper.

1888 030817 © 2011 SGM  Printed in Great Britain
prevalence in young children, elderly and institutionalized patients. The virus is also known to be associated with food-borne outbreaks and outbreaks in various institutional settings (Belliot et al., 1997; Mitchell et al., 1999; Oishi et al., 1994). While most cases of HAstV-associated diarrhoea are mild and self-limited, the disease may be more severe in the elderly and immunocompromised patients (Liste et al., 2000; Méndez & Arias, 2007).

In recent years, numerous novel astroviruses have been discovered in humans and a diverse range of animals, including bat, cheetah, doe deer, dog, bottlenose dolphin, mink, brown rat, sea lion and turkey (Atkins et al., 2009; Blomström et al., 2010; Chu et al., 2008, 2010; Guy et al., 2004; Mittelholzer et al., 2003; Rivera et al., 2010; Smits et al., 2010; Toffan et al., 2009; Zhu et al., 2009). Phylogenetic characterization of these astroviruses has shown that a single host species may be susceptible to infection by divergent astroviruses. For example, analysis of the partial ORF1b and ORF2-encoded amino acid sequences revealed that the classic (genotypes 1–8), VA and MLB strains of HAstV are phylogenetically distant and do not cluster to form a monophyletic group. Among the known mammalian astroviruses, bovine astrovirus (BAstV) is one of the earliest to be discovered and studied extensively (Woode & Bridger, 1978; Woode et al., 1984). It has been identified in USA and England, with two BAstV serotypes established based on the results of a virus neutralization assay (Woode et al., 1985).

The discovery of novel viruses is important for the understanding of many human and animal diseases. In recent years, our research group has conducted an ongoing surveillance study to identify novel viruses in different birds and mammals, which has resulted in the discovery of diverse viruses (Lau et al., 2005, 2008, 2010; Woo et al., 2009, 2010). In this paper, we report the identification of astroviruses from surveillance of bovine faecal samples over a period of 2 years. The nearly full-length genome sequences of four strains and partial genome sequences of other strains were obtained and analysed accordingly.

RESULTS

Surveillance and identification of BAstV

A total of 209 rectal swabs from cattle in Hong Kong SAR (HKSAR) were obtained. RT-PCR using broadly reactive degenerate primers for a 418 nt fragment in the pol gene of astroviruses was positive in four specimens. Repeated screening using more sensitive specific primers was positive for one additional specimen. Two specimens (B76 and B161) produced mixed populations of PCR products, and were analysed by additional cloning of the amplicons and sequencing, which revealed the presence of two astroviruses in each specimen. All sequences from positive samples had <80% nt identity to known astrovirus sequences, and they fell into a cluster separate from the other astroviruses. These BAstV sequences were most closely related to the recently discovered Capreolus capreolus astrovirus found in the European doe deer (Smits et al., 2010). None of the animals, from which the RT-PCR-positive specimens were obtained, had any signs of gastrointestinal disease such as anorexia and diarrhoea.

Viral culture

No cytopathic effect (CPE) was observed in any of the cell lines inoculated with the specimens positive for BAstV by RT-PCR. Quantitative RT-PCR using culture supernatants and cell lysates for monitoring the presence of viral replication also showed negative results.

Genome sequencing and analysis of BAstV

The sequences of four complete BAstV genomes, including two from the same positive specimen, were obtained. Partial BAstV genome sequence was obtained for the remaining positive specimen, but complete genome sequencing proved difficult due to limited clinical materials available and possibly low viral titres. The complete BAstV genomes range from 6253 to 6317 nt in length, and their genomic G+C content are 53.3–54.3 mol%. The predicted genome organization is similar to other astroviruses, with three overlapping ORFs encoding the non-structural and viral capsid proteins (Fig. 1). Genome sequence identity comparison with other related astroviruses showed the highest degree of sequence conservation in the ORF1b region (Fig. 2 and Table 1).

The complete ORF1a is 2448–2454 nt long and encodes a 90 kDa non-structural protein, nsp1a, that contains a conserved serine-like protease motif and a catalytic triad of histidine, aspartic acid and serine residues similar to other viral 3C-like proteases. A bipartite nuclear localization signal (NLS), KGKTGGKRLRMVGAKEKRQR, is predicted at position 633 for BAstV-B18. Similar conserved NLS sequences are found in the same region for the other sequenced BAstV genomes. Five transmembrane helices (aa positions 147–169, 222–244, 251–273, 293–315 and 336–358 in BAstV-B18) were predicted to be present in the N-terminal half of nsp1a, similar to other astroviruses. The conserved ‘slippery heptamer’ sequence, 5’-AAAAAAC-3’, is present near the 3’ end of the ORF1a in all sequenced BAstV genomes, and is responsible for inducing ribosomal frameshift during translation to generate the polyprotein nsp1ab.

The complete ORF1b is 1566–1611 nt long, with an overlapping region with ORF1a of between 102 and 147 nt in length. Its reading frame is −1 relative to that of ORF1a for translation to initiate after programmed ribosomal frameshift occurs at the ‘slippery heptamer’ site. The 502 aa sequence encoded by the coding region of ORF1b contains a RdRP, inclusive of the characteristic YGDG motif at position 361. When considered together, ORF1a and ORF1b encodes a polyprotein nsp1ab of

http://vir.sgmjournals.org
Fig. 1. Schematic diagram showing the genome organization of the presently identified BAstV and other astroviruses.

Fig. 2. Sequence identity plot of the consensus BAstV genome with selected astroviruses.
approximately 147 kDa in size, which would undergo proteolytic cleavage by the viral 3C-like protease to become active non-structural proteins.

ORF2 is between 2280 and 2319 nt in length, and encodes the structural proteins in the same reading frame as ORF1a. The molecular mass of entire precursor protein is around 83–84.5 kDa. A short stretch of basic amino acid residues, mainly arginine and lysine, is found near the N terminus of the protein (Fig. 3), and has similarly been described for other astroviruses like HAstV, turkey astrovirus (TAv), and duck astrovirus. In addition, there is a stretch of acidic residues near the C terminus (Fig. 3). This latter stretch of acidic residues is conserved among closely related mamastroviruses like HAstV and PAstV, but largely absent in the capsid precursors of more distantly related mamastroviruses like mink astrovirus and ovine astrovirus.

Phylogenetic analysis

The phylogenetic trees constructed using the predicted amino acid sequences of the translation products of ORF1a, ORF1b, and ORF2, and the best-fit model for the structural proteins encoded by ORF1a and ORF1b, were found to be the best-fit substitution models for the non-structural proteins encoded by ORF1a and ORF1b, and ORF2, respectively. The molecular mass of the entire precursor protein is around 83–84.5 kDa. A short stretch of basic amino acid residues, mainly arginine and lysine, is found near the N terminus (Fig. 3). This latter stretch of acidic residues is conserved among closely related mamastroviruses like HAstV and PAstV, but largely absent in the capsid precursors of more distantly related mamastroviruses like mink astrovirus and ovine astrovirus.

Table 1. Pairwise amino acid sequence identities between the predicted polypeptides of BAstV and selected AstV

<table>
<thead>
<tr>
<th>GenBank accession no.</th>
<th>BAstV-B18</th>
<th>BAstV-B34</th>
<th>BAstV-B76</th>
<th>BAstV-B76-2</th>
<th>BAstV-B170</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF1a</td>
<td>ORF1a</td>
<td>ORF1a</td>
<td>ORF1a</td>
<td>ORF1a</td>
<td>ORF1a</td>
</tr>
<tr>
<td>Avastrovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANV-1</td>
<td>NC_003790</td>
<td>18.2</td>
<td>40.5</td>
<td>17.5</td>
<td>NA</td>
</tr>
<tr>
<td>Duck AstV</td>
<td>NC_012437</td>
<td>15.8</td>
<td>39.6</td>
<td>16.5</td>
<td>NA</td>
</tr>
<tr>
<td>Turkey AstV</td>
<td>NC_002470</td>
<td>16.6</td>
<td>39.0</td>
<td>17.8</td>
<td>NA</td>
</tr>
<tr>
<td>Mamastrovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAstV-1</td>
<td>NC_001943</td>
<td>24.3</td>
<td>57.6</td>
<td>26.5</td>
<td>NA</td>
</tr>
<tr>
<td>HAstV-6</td>
<td>GQ495608</td>
<td>26.0</td>
<td>57.9</td>
<td>27.3</td>
<td>NA</td>
</tr>
<tr>
<td>AstV-MLB1</td>
<td>NC_011400</td>
<td>27.4</td>
<td>51.5</td>
<td>25.3</td>
<td>NA</td>
</tr>
<tr>
<td>AstV HMO-A</td>
<td>NC_013443</td>
<td>24.9</td>
<td>50.6</td>
<td>23.7</td>
<td>NA</td>
</tr>
<tr>
<td>Mink AstV</td>
<td>NC_004579</td>
<td>24.9</td>
<td>52.2</td>
<td>21.2</td>
<td>NA</td>
</tr>
<tr>
<td>Ovine AstV</td>
<td>NC_013443</td>
<td>26.5</td>
<td>50.6</td>
<td>23.7</td>
<td>NA</td>
</tr>
</tbody>
</table>

Phylogenetic analysis

The phylogenetic trees constructed using the predicted amino acid sequences of the translation products of ORF1a, ORF1b, and ORF2, and the best-fit model for the structural proteins encoded by ORF1a and ORF1b, were found to be the best-fit substitution models for the non-structural proteins encoded by ORF1a and ORF1b, and ORF2, respectively. The molecular mass of the entire precursor protein is around 83–84.5 kDa. A short stretch of basic amino acid residues, mainly arginine and lysine, is found near the N terminus (Fig. 3). This latter stretch of acidic residues is conserved among closely related mamastroviruses like HAstV and PAstV, but largely absent in the capsid precursors of more distantly related mamastroviruses like mink astrovirus and ovine astrovirus.
Recombination analysis

Recombination analysis was performed on the capsid precursor amino acid sequences of BAstV and CCAstV, with PAstV-2 as an outgroup. The presence of recombination is strongly supported by the Single Breakpoint Recombination analysis, and recombination breakpoints are most strongly supported at the positions 637 and 660 (BAstV-B76 numbering) as identified by GARD. As these two positions are proximal to each other and only separated by a gap-rich region in the multiple alignment, we consider them to be the likely result of a single recombination event. Additional breakpoints were identified at positions 65 and 430, but were not supported by the Kishino–Hasegawa (KH) test (adjusted P>0.05). To illustrate the phylogenies of the different partitions, neighbour-joining trees were constructed for the separate partitions of BAstV and CCAstV ORF2 (Fig. 6). Recombination analysis on amino acid sequences encoded by ORF1a and ORF1b did not reveal any additional recombination breakpoints supported by GARD and the KH test.

Additional analysis was performed on the complete genome sequences of BAstV by bootscanning in SimPlot, which does not strongly support the presence of recombination. However, the choice of parental sequences for the bootscan analysis was suboptimal, as there are no complete genome sequences for CCAstV, which is the astrovirus most closely related to the presently characterized BAstV strains.

Fig. 3. Sliding window analysis of basic and acidic amino acid residues in the viral capsid precursor of BAstV-B76, human astrovirus 6 (HAstV-6) and astrovirus HMO-A. The window and step sizes were set at 15 and 3, respectively.
Fig. 4. Maximum-likelihood trees constructed from the amino acid sequences of BAstV proteins. (a) Non-structural protein nsp1a encoded by ORF1a, (b) portion of non-structural protein nsp1ab encoded by ORF1b, (c) viral capsid precursor encoded by ORF2. (AstV, astrovirus; ANV, avian nephritis virus; BAstV, bovine astrovirus; CCAstV, *Capreolus capreolus* astrovirus; H AstV, human astrovirus; PAstV, porcine astrovirus). Bars, number of inferred substitutions per site.

Fig. 5. Neighbour-joining tree constructed from the nucleotide sequences in the 414 nt region used for BAstV screening in the present study. Bars, number of inferred substitutions per site.
Phylogenetic analysis of the amino acid sequences encoded by the BAstV genomes confirmed that they are members of the genus Mamastrovirus. Analysis of the ORF2 sequence showed their close relationship to CCAstV, with average sequence identities of 41.6–68.2% between the BAstV and CCAstV sequences. Given their positions in a monophyletic group and the strong branch support with an approximate Likelihood-Ratio Test (aLRT) statistic of 1.00, it may be appropriate to consider the presently identified BAstV and CCAstV as different strains of the same virus species despite their different hosts, similar to the situation for FAstV and cheetah astrovirus. To better resolve this uncertainty in taxonomic classification, we performed additional analysis of the genetic distances among the various BAstV and CCAstV in the ORF2 region (see Supplementary Results, available in JGV Online). The analysis results were interpreted according to the preliminary criteria set within the latest draft proposal for revision of Mamastrovirus taxonomy. Overall, the results supported classification of these astroviruses into three to four species under the proposed ‘genocluster GI’ of the genus Mamastrovirus, with two CCAstV being clustered together with BAstV-B18 and BAstV-B76-2. A more definitive analysis of its phylogenetic position would require the complete CCAstV genome sequence, or at least the full-length sequences of its non-structural proteins.

It is not uncommon for astroviruses to be detected alongside other enteric viruses, such as noroviruses, in alimentary specimens (Bhattacharya et al., 2006; Koh et al., 2008; Taylor et al., 1997). However, to our knowledge, a single case of co-infection by two different genotypes of astroviruses has been reported in pigs only recently (Luo et al., 2011). Hence, the two instances of natural co-infection in cattle by BAstV-B76/B76-2 and BAstV-B161/B161-2 represent the second report of this phenomenon, which may be more prevalent than previously thought. The occurrence of co-infection is particularly significant for the genomic evolution of these viruses, as it is a pre-requisite for recombination between different strains. Naturally recombinant human and animal astroviruses have been reported previously (Pantin-Jackwood et al., 2006; Rivera et al., 2010; Strain et al., 2008; Ulloa & Gutiérrez, 2010; Walter et al., 2001; Wolfaardt et al., 2011) and a model to explain the mechanism of genomic recombination in ssRNA viruses is available (Lai, 1992). Results of recombination analysis on BAstV and CCAstV confirmed that at least one recombination event has occurred during the evolution of these astroviruses. The detection of recombination in ORF2 is perhaps not surprising, as recombination in this region can generate diversity rapidly in the viral capsid proteins and enable the virus to escape host immunity. Given the likely role of the capsid proteins in defining the cell tropism and host range of the virus, it may also be possible that the recombination in this region can...
aid the virus in jumping host species. Further research on the maturation and structure of the capsid proteins would be needed to explore the positional significance of the recombination breakpoint.

Previous studies have suggested that BAstV may be excreted by up to 60–100 % of calves on farms (Bridger et al., 1984). However, the low prevalence of BAstV identified in our study is not comparable to their figures because of significant differences in methodology. Firstly, our sampling population was limited to adult animals, while the earlier results were based on sampling of calves within the first 5 weeks of life. Presumably, the less mature immune system may predispose to a higher BAstV carriage rate. Other factors that may contribute to the different BAstV carriage rates include strain-specific differences in virulence, immunization status of the animals and host genetic resistance to infection. Additionally, we employed RT-PCR detection with sequence confirmation for our present study, whereas the method of detection used in the other study was not clearly specified. In any case, available methods for detection of astroviruses at that time were largely limited to electron microscopy, viral culture and immunofluorescence. As the specificities of these methods are usually much lower than RT-PCR, it is likely that the presence of false-positive results inflated the actual prevalence of the virus in their animal population. In our opinion, a large-scale surveillance study of BAstV should be performed in commercially farmed cattle populations to reassess the epidemiology of this enteric virus.

Unlike many other astroviruses, BAstV has not been found to cause diarrhoeal disease in cattle despite asymptomatic viral shedding in stool samples (Woode & Bridger, 1978; Woode et al., 1984), which was consistent with the findings of our present study. The closely related CCAstV was found in diarrhoeal stool samples of roe deer, though the authors of the study did not infer any causative links between the astroviruses and gastrointestinal disease. Indeed, whether BAstV and CCAstV could cross the host species barrier to infect both cattle and roe deer and possibly exhibit differential virulence are important questions that deserve further study. The exploration of genetic diversity and evolution of astroviruses in wild and domestic animals would be key to our knowledge and further understanding of this diverse group of important gastrointestinal viruses.

**METHODS**

**Surveillance and sample collection.** The Food and Environmental Hygiene Department (FEHD), HKSAR provided access to samples collected from various locations in HKSAR over a 24-month-period (September 2008–August 2010). All cattle were adult animals and approximately 24 months in age at the time of sampling. None of the cattle showed signs of diarrhoeal disease at the time of collection. Rectal swabs were collected by the staff of the FEHD.

**RNA extraction.** Viral RNA was extracted from the rectal swabs using a QIAamp Viral RNA mini kit (Qiagen). The RNA was eluted into 50 µl RNase-free water and was used as the template for RT-PCR.

**RT-PCR of the pol gene of astroviruses using conserved primers, and sequencing.** Screening was performed by amplifying a 432 nt fragment of the pol gene of astroviruses using conserved degenerate primers (5'-GAYTGGACBCGHTWTGATG-3' and 5'-KTTYRATCCATATNNCCAA-3') that were designed from the multiple alignment of RdRp gene sequences of different astrovirus species. Reverse transcription was performed using a SuperScript III kit (Invitrogen) according to manufacturer’s instructions for first-strand synthesis. The PCR mixture (25 µl) contained cDNA, PCR buffer (10 mM Tris/HCl pH 8.3, 50 mM KCl, 2 mM MgCl2 and 0.01 % gelatin), 200 µM of each dNTP and 1.0 U AmpliTaq Gold polymerase (Applied Biosystems). PCR cycling conditions were as follows: hot start at 94°C for 7 min, followed by 50 cycles of 94°C for 1 min, 48°C for 1 min and 72°C for 1 min with a final extension at 72°C for 10 min in an automated thermal cycler (Applied Biosystems). Standard precautions were taken to avoid PCR contamination and no false-positive signal was observed in the negative controls. The PCR products were gel-purified using a QIAquick gel extraction kit (Qiagen). Both strands of the PCR products were sequenced twice with an ABI Prism 3730xl DNA Analyser (Applied Biosystems), using the two PCR primers.

**RT-PCR screening of BAstV using specific primers** Additional RT-PCR screening was performed on the same samples using specific primers designed from the pol gene sequences of BAstV obtained from previous rounds of RT-PCR and sequencing, as RT-PCR screening with specific primers usually offer higher sensitivity than a comparable screening with consensus degenerate primers. The primer sequences are as follows: set A forward primer LPW13996 5'-AACAAACAACCGTGAGAGG-3', set A reverse primer LPW13997 5'-GGTGTGCTGTAGTCTACCCAGA-3', set B forward primer LPW13998 5'-GGTGGAGAACGAATCAAGGCTG-3', set B reverse primer LPW13999 5'-GGTGTGATGCTGTCCTCAGAGG-3'. PCR screening conditions were as follows: hot start at 94°C for 7 min, followed by 50 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 1 min with a final extension at 72°C for 10 min in an automated thermal cycler (Applied Biosystems). Standard precautions were taken to avoid PCR contamination and no false-positive signal was observed in the negative controls. PCR product purification and sequencing were performed as above.

**Cloning of PCR product and sequencing.** Purified PCR products were cloned into a pCR2.1-TOPO vector (Invitrogen) according to manufacturer’s instructions. The vector was then used to transform the competent Escherichia coli strain DH5α by electroporation. Positive transformants were identified by blue-white screening, and eight colonies were selected for DNA sequencing of the construct using the M13 forward and reverse primers according to the manufacturer’s instructions. Sequencing reactions were performed as described above.

**Viral culture.** Original swab materials positive for BAstV by RT-PCR were provided by TLVL for attempted virus isolation by using Caco-2 (human epithelial colorectal adenocarcinoma; ATCC HTB-37) and MDBK (bovine kidney; ATCC CCL-21) cells. For each of the positive samples, 200 µl sample in viral transport medium stored at −80°C was retrieved. After thawing, they were immediately clarified by centrifugation. The clarified inoculum was added to culture tubes by adsorption inoculation, which involved decanting off the culture medium and direct application of the inoculum to the cell monolayer. After 1 h of adsorption at 37°C in a horizontal position, excess inoculum was discarded and replaced by 1.5 ml minimum essential...
medium (MEM; Invitrogen) supplemented with either 1 or 50 μg bovine pancreatic trypsin (Sigma-Aldrich) ml⁻¹ (Aroonprasert et al., 1989). The cultures were incubated at 37 °C with 5% carbon dioxide in stationary slanted racks. They were inspected daily for CPE such as rounding or syncytium formation by inverted microscopy. At the higher concentration of trypsin, CPE features could not be observed clearly due to the detachment of cells and RT-PCR was used to detect viral replication. After 7 days of incubation, subculturing to a fresh cell line was performed. Subculturing was performed twice even if there was no CPE. When no CPE was observed at the end of incubation, the cells were frozen and thawed once and the culture lysates collected for RNA extraction and RT-PCR as described above. If RT-PCR was only positive for the first inoculated culture but negative on subsequent passages, the result was interpreted as carryover from the specimen instead of being due to viral replication in cell lines.

**Genome sequencing.** Viral genomes were amplified and sequenced using strategies we had previously used for complete genome sequencing of other RNA viruses (Woo et al., 2009, 2010), with the RNA extracted from the rectal swabs as templates. Reverse transcription was performed as described above. PCR primers were designed based on the multiple alignment of genomes of related astroviruses, including HAstV, bat astrovirus and mink astrovirus. Additional primers for subsequent rounds of PCR were designed based on the results of earlier rounds of genome sequencing. The complete set of primer sequences is available upon request. Sequences from the 5’ and 3’ ends of the viral genomes were obtained by random amplification of cDNA ends using a SMARTer RACE cDNA Amplification kit (Clontech).

**Phylogenetic and genome analysis.** The nucleotide sequences of the genomes and the deduced amino acid sequences of the ORFs were compared with those of other astroviruses. Multiple alignments of nucleotide and amino acid sequences were constructed using MUSCLE (Edgar, 2004). Best-fit substitution models for phylogenetic analysis of nucleotide and amino acid sequences were selected using jModelTest and ProtTest, respectively (Abascal et al., 2006, 2010), with the KH test was implemented in SimPlot version 3.5.1 was used for recombination analysis of the complete BAstV genomes.

**Recombination analysis.** Recombination analysis for protein sequences was performed according to the recommended protocol implemented on the DataMonkey server (Delport et al., 2010). Sequences were first analysed using the Single Breakpoint Recombination method to look for evidence of any recombination. If the results were supportive of the presence of recombination, then the GARD method was employed to identify all possible recombination breakpoints (Kosakovsky Pond et al., 2006). The KH test was performed on the inferred partitions to examine the significance of breakpoints (Kishino & Hasegawa, 1989). The bootscanning procedure as implemented in SimPlot version 3.5.1 was used for recombination analysis of the complete BAstV genomes.

**REFERENCES**


