Characterization of phylogenetically diverse astroviruses of marine mammals

Rebecca Rivera,[1] Hendrik H. Nollens,[1,2] Stephanie Venn-Watson,[3] Frances M. D. Gulland[4] and James F. X. Wellehan, Jr[2]

Correspondence
Rebecca Rivera
Rrivera@hswri.org

[1]Hubbs–SeaWorld Research Institute, San Diego, CA 92109, USA

[2]Marine Mammal Health Program, College of Veterinary Medicine, University of Florida, Gainesville, FL 32610, USA

[3]Navy Marine Mammal Program Foundation, San Diego, CA 92106, USA

[4]The Marine Mammal Center, 2000 Bunker Road, Sausalito, CA 94965, USA

Astroviruses are small, non-enveloped, positive-stranded RNA viruses. Previously studied mammalian astroviruses have been associated with diarrhoeal disease. Knowledge of astrovirus diversity is very limited, with only six officially recognized astrovirus species from mammalian hosts and, in addition, one human and some bat astroviruses were recently described. We used consensus PCR techniques for initial identification of five astroviruses of marine mammals: three from California sea lions (*Zalophus californianus*), one from a Steller sea lion (*Eumetopias jubatus*) and one from a bottlenose dolphin (*Tursiops truncatus*). Bayesian and maximum-likelihood phylogenetic analysis found that these viruses showed significant diversity at a level consistent with novel species. Astroviruses that we identified from marine mammals were found across the mamastrovirus tree and did not form a monophyletic group. Recombination analysis found that a recombination event may have occurred between a human and a California sea lion astrovirus, suggesting that both lineages may have been capable of infecting the same host at one point. The diversity found amongst marine mammal astroviruses and their similarity to terrestrial astroviruses suggests that the marine environment plays an important role in astrovirus ecology.

INTRODUCTION

Astroviruses are small, non-enveloped viruses with a positive-stranded RNA genome and a distinct star-like surface morphology. They were discovered relatively recently and were first reported by Madeley & Cosgrove (1975). Based mainly on the host of the virus and the genome structure, the family *Astroviridae* is divided into two genera. Members of the genus *Avastrovirus* are found in avian hosts, whereas members of the genus *Mamastrovirus* are found in mammalian hosts (Monroe *et al.*, 2005). Known astrovirus diversity is very limited, with only three astrovirus species from avian hosts recognized by the International Committee on Taxonomy of Viruses (ICTV) and six recognized astrovirus species from mammalian hosts [bovine astrovirus, feline astrovirus, human astrovirus (serotypes 1–8), mink astrovirus,

ovine astrovirus and porcine astrovirus] (Monroe *et al.*, 2005). Recently, a divergent human astrovirus from a child with diarrhoea (Finkbeiner *et al.*, 2008), and a number of astroviruses from vespertillonid and rhinolophid bats have been described (Chu *et al.*, 2008). The mamastroviruses have a small, positive-sense, single-stranded RNA genome of <7000 nt that encodes three open reading frames (ORFs): ORF1a, ORF1b and ORF2. A frameshift between ORF1a and ORF1b allows ORF1 to encode both a protease and an RNA-dependent RNA polymerase (RdRp). ORF2 encodes the viral capsid protein. Cloning of ORF2 in expression vectors has allowed the *in vitro* assembly of virus-like particles (Caballero *et al.*, 2004).

Whilst avastroviruses can cause intestinal or renal disease, mamastroviruses predominantly establish infections in the gastrointestinal tract of their hosts. Human astroviruses are a frequent cause of enteric disease in neonatal, elderly and immunocompromised humans (Dennehy *et al.*, 2001; Gallimore *et al.*, 2005; Marshall *et al.*, 1987), and one study of faecal electron microscopy of cats found that astroviruses were the most common virus particles observed in cats with diarrhoea (Marshall *et al.*, 1987). The astrovirus capsid protein plays a unique role in the

pathogenesis of diarrhoea: the capsid protein interacts with apical enterocyte membranes, increasing permeability independent of virus replication. Much like *Vibrio cholerae*, astroviruses cause a secretory diarrhoea without much of a histological footprint (Moser *et al.*, 2007). Histopathology is therefore an insensitive test for diagnosis of astroviral diarrhoea, potentially leading to underdiagnosis of astrovirus infections.

Only very recently has the presence of astroviruses in wildlife hosts been reported (Atkins *et al.*, 2009; Chu *et al.*, 2008). Here, we report the first detection of five genetically distinct astroviruses from three marine mammal host species.

RESULTS

Negative-staining electron microscopy

Virus particles were detected in all five faecal samples. The five isolates are henceforth referred to as California sea lion astroviruses 1–3 (CslAstV-1, -2 and -3), Steller sea lion astrovirus 1 (SslAstV-1) and bottlenose dolphin astrovirus 1 (BdAstV-1). Individual capsids were 30–35 nm wide. In all samples, the non-enveloped icosahedral virus particles had distinct star-like surface projections and were consistent in size and morphology with members of the family *Astroviridae* (Fig. 1).

Degenerate PCR

The primer combination Astr4380F/Astr4811R (see Methods) yielded a band of 431 bp on isolates CslAstV-1, CslAstV-3, SslAstV-1 and BdAstV-1. The primers Astr4574F/Astr4722R yielded a band of 148 bp on isolates CslAstV-1, CslAstV-2, CslAstV-3 and SslAstV-1. The primer combination 5159F/5819R yielded a band of 660 bp on all five isolates.

Sequence extension via specific PCR yielded final contiguous molecules of 1340 bp (CslAstV-3) and 1348 bp (SslAstV-1). Additional 3′ GeneRacer (Invitrogen) analysis yielded final contiguous molecules of 3174 bp (CslAstV-1), 3505 bp (CslAstV-2) and 3985 bp (BdAstV-1). The contiguous molecules corresponded to the partial capsid gene (ORF1b) and the full-length RdRp gene of reference astroviruses (Fig. 2). The contiguous sequences were submitted to GenBank under accession numbers FJ890351 (CslAstV-1), FJ890352 (CslAstV-2), FJ890353 (CslAstV-3), FJ890354 (SslAstV-1) and FJ890355 (BdAstV-1).

Comparison with other sequences in GenBank revealed that all five contiguous molecules represented novel astroviruses. TBLASTX results for BdAstV-1 showed the highest amino acid identity score with human astrovirus 3 (GenBank accession no. AF141381) for the RdRp and with human astrovirus 7 (GenBank accession no. Y08632) for the capsid precursor. TBLASTX results for CslAstV-1 showed the highest score with ovine astrovirus (GenBank accession no. AF141381) for the RdRp and with mink astrovirus
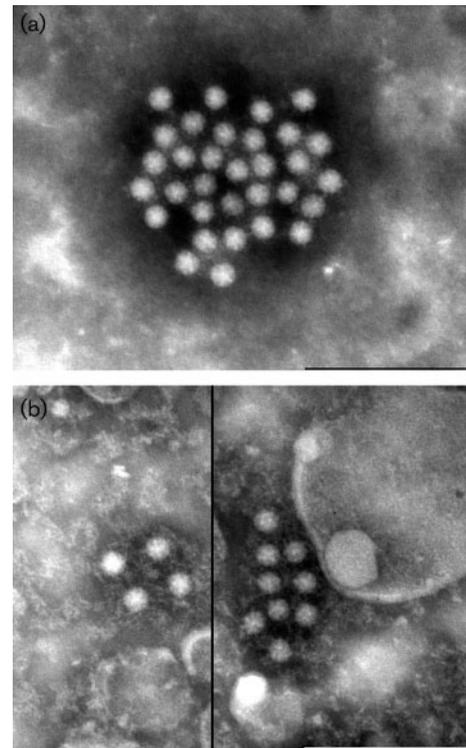

Fig. 1. Electron micrographs of faecal samples. Negative staining revealed clusters of virus particles in faecal samples from three California sea lions (*Zalophus californianus*), one Steller sea lion (*Eumetopias jubatus*) and one bottlenose dolphin (*Tursiops truncatus*). Individual capsids were 30−35 nm wide. The non-enveloped, icosahedral virus particles had distinct star-like surface projections consistent with astroviruses. Isolates CslAstV-1 (a) and BdAstV-1 (b) are shown. The image in (b) is a composite of two images taken from the same sample at the same magnification.

(GenBank accession no. AY179509) for the capsid precursor. TBLASTX results for CslAstV-2 showed the highest score with human astrovirus 2 (GenBank accession no. L13745) for the RdRp and with porcine astrovirus (GenBank accession no. AB037272) for the capsid precursor. TBLASTX results for CslAstV-3 showed the highest score with human astrovirus 3 (GenBank accession no. AF141381) for both the RdRp and the capsid precursor. TBLASTX results for SslAstV-1 showed the highest score with mink astrovirus (GenBank accession no. AY179509) for both the RdRp and the capsid precursor.

Phylogenetic analysis

Bayesian phylogenetic analysis showed the greatest harmonic mean of estimated marginal likelihoods using the MUSCLE alignment for the RdRp (see Supplementary Fig. S1, available in JGV Online) and the T-Coffee alignment for the capsid (see Supplementary Fig. S2, available in JGV Online). For the RdRp, the Wag model of amino acid substitution was found to be most probable with a
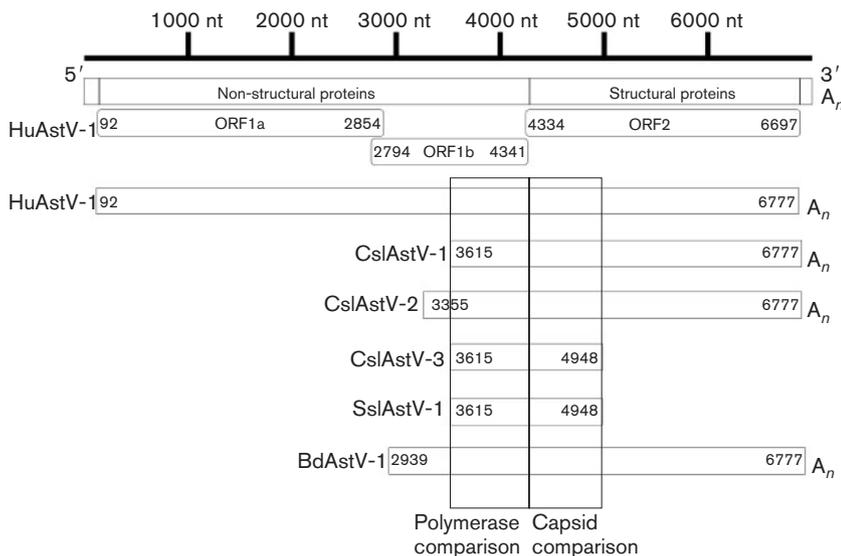
Fig. 2. Alignment of the 3174 bp CslAstV-1 genome segment, 3505 bp CslAstV-2 genome segment, 1340 bp CslAstV-3 genome segment, 1348 bp SslAstV-1 genome segment and 3985 bp BdAstV-1 genome segment, compared with the full-length genome sequence of human astrovirus 1 (6697 bp; GenBank accession no. AY720892). The relative positions of the ORFs encoding the non-structural and structural proteins are also represented. Partial gene segments of the polymerase and capsid genes used for phylogenetic analysis are indicated.

posterior probability of 1.000 (Whelan & Goldman, 2001). For the capsid precursor protein, the Wag model was also most probable with a posterior probability of 0.993, and a posterior probability of 0.007 for the JTT model (Jones *et al.*, 1992). Bayesian trees using the MUSCLE alignment for the RdRp and the T-Coffee alignment for the capsid gene are shown in Figs 3 and 4, respectively.

Maximum-likelihood (ML) analysis found the most likely tree from the MUSCLE alignment and the PMB model of amino acid substitution for the RdRp, and from the T-Coffee alignment and the JTT model of amino acid substitution for the capsid precursor. These parameters were used for bootstrap analysis. Bootstrap values from ML analysis are shown on the trees (Figs 3 and 4).

Recombination analysis

The MUSCLE nucleotide alignment was 1479 nt in length. Recombination analysis identified a probable recombination event in CslAstV-3, from parents human astrovirus 4 and CslAstV-2, with a *P*-value of 4.911×10^{-6} after corrections for multiple comparisons. It was supported by RDP ($P=2.197 \times 10^{-4}$), Recscan ($P=4.359 \times 10^{-3}$), MaxChi ($P=7.943 \times 10^{-4}$), Chimaera ($P=4.911 \times 10^{-6}$), SiScan ($P=3.089 \times 10^{-5}$), LARD ($P=2.339 \times 10^{-6}$) and 3Seq ($P=2.634 \times 10^{-3}$). GENECONV did not support this event. This event started at nt 832 of the alignment, with a *P*-value for the beginning breakpoint of 8.637×10^{-3} (Fig. 5). This initial breakpoint falls early in the coding region for ORF2, at aa 15. Prior to the breakpoint, CslAstV-3 showed greater similarity to human astrovirus 4, and greater similarity to CslAstV-2 after the breakpoint. The end point was not clear – although predicted for nt 1062 of the alignment, the *P*-value was only 0.600. A bootscanning diagram for this event is shown in Fig. 5.

Differences in branching patterns between the capsid and polymerase in the mink astrovirus/CslAstV-1/SslAstV-1 clade were not reconciled by identification of a recombination event in the initial analysis, so other sequences in the alignment were masked for additional analysis. Weak support was found for a recombination event between nt 727 and 847, without correction for multiple comparisons. This region covers the end of ORF1b and the start of ORF2. This event was supported by RDP ($P=0.02957$), Recscan ($P=0.04462$), SiScan ($P=7.198 \times 10^{-4}$), LARD ($P=2.405 \times 10^{-4}$) and 3Seq ($P=0.01088$), but not by GENECONV ($P=0.074$), MaxChi ($P=0.1178$) or Chimaera ($P=0.2422$).

DISCUSSION

This report documents the first identification of astroviruses in marine mammals. Surprising diversity was identified in these hosts. With the exception of MLB1 (Finkbeiner *et al.*, 2008), the astroviruses of humans, the host species that has been investigated most heavily, constitute a single species with eight serotypes. In contrast, the evolutionary distance between the first three astroviruses found in California sea lions is comparable to that seen between recognized species within the family *Astroviridae*. BdAstV-1 and SslAstV-1 also appear to be distinct from other astroviruses at a distance consistent with species differentiation.

The genetic distances between these novel viruses were generally slightly greater in the capsid region than in the RdRp region (Figs 3 and 4), even though we examined the region of capsid that is expected to be most conserved. This is consistent with other studies. Capsids are typically under strong positive selective pressure from the host immune system, and one analysis found that most positively
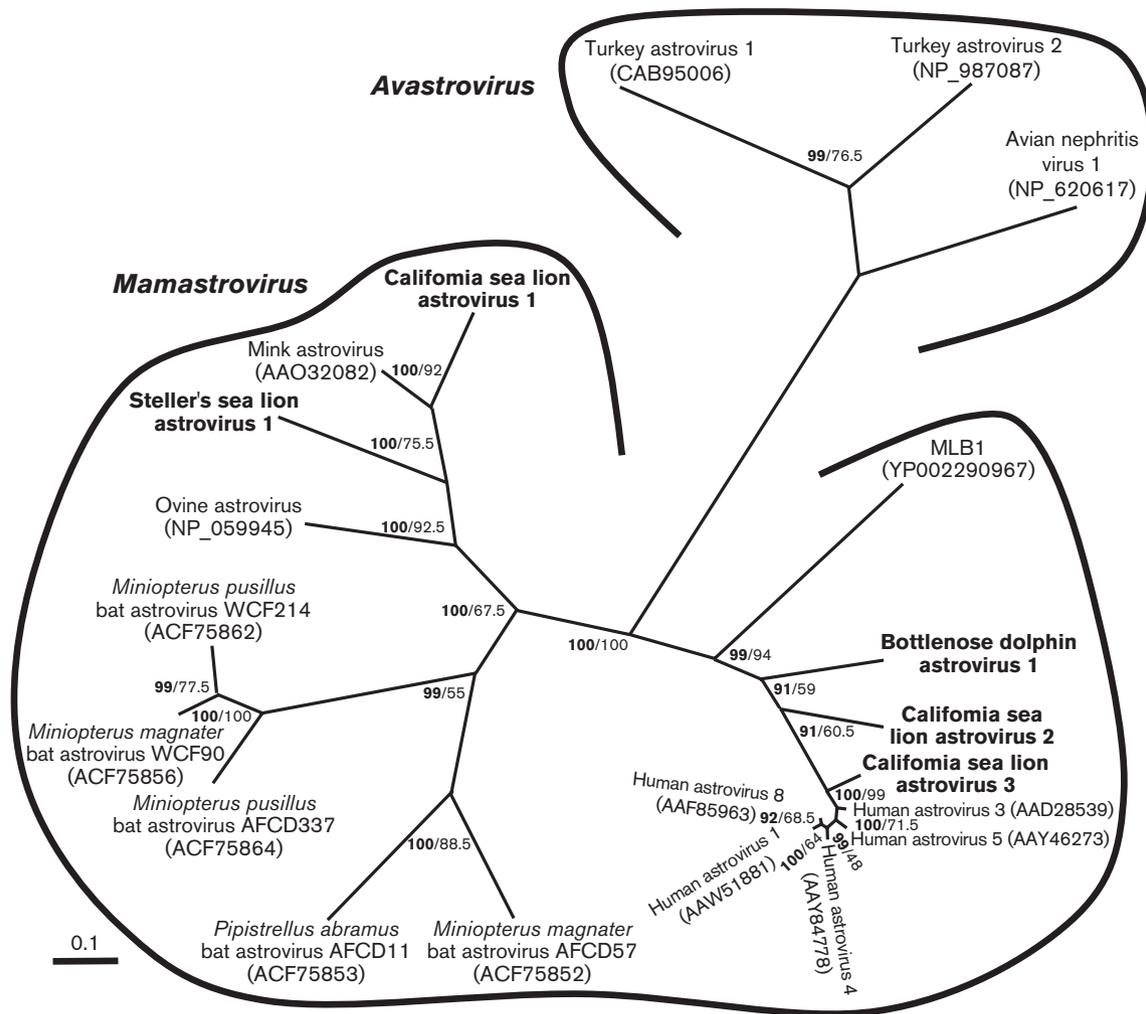
Fig. 3. Bayesian phylogenetic tree of predicted 237–254 aa partial astrovirus RdRp sequences, based on MUSCLE alignment. Bayesian posterior probabilities of branchings as percentages are shown in bold, and ML bootstrap values for branchings based on 200 resamplings are given to the right. Avian nephritis virus 1 was designated the outgroup. Virus genera are delineated by brackets. Marine mammal astroviruses are shown in bold. Bar, 0.1 substitutions per site.

selected sites in astroviruses are present in the capsid (Van Hemert *et al.*, 2007).

The phylogenetic topology determined in this study is largely in agreement with previous analyses of astrovirus phylogeny (Jonassen *et al.*, 2001; Lukashov & Goudsmit, 2002), although we do not find support for the findings of Chu *et al.* (2008) that bat astroviruses AFCD11 and AFCD57 are not monophyletic with the other chiropteran mamastroviruses.

Of the six mamastrovirus species recognized by the ICTV, five are from hosts from one mammalian superorder, Laurasiatheria. The other superorder of placental mammals, Euarchontoglires, has only one host from which an astrovirus is recognized, humans. The greater diversity of astroviruses within laurasiatherian hosts may imply a longer host–virus relationship. Bats, from whom additional

astroviruses have recently been described, are also members of the superorder Laurasiatheria. The astroviruses of bats mostly appear to form a distinct monophyletic group, unlike those of marine mammals, which are distributed across the tree of the known mamastroviruses. As this manuscript was being prepared, evidence of bat astroviruses outside the clade of the viruses found by Chu *et al.* (2008) was published (Zhu *et al.*, 2009).

Astroviruses are very stable in aquatic environments (Espinosa *et al.*, 2008). Surveys have found a human astrovirus prevalence of up to 61 % in some marine shellfish populations, which are good particle concentrators (Elamri *et al.*, 2006). The wide diversity seen in astroviruses identified from marine mammals implies that the marine environment may play a significant role in astrovirus ecology. Similarly, the marine environment is
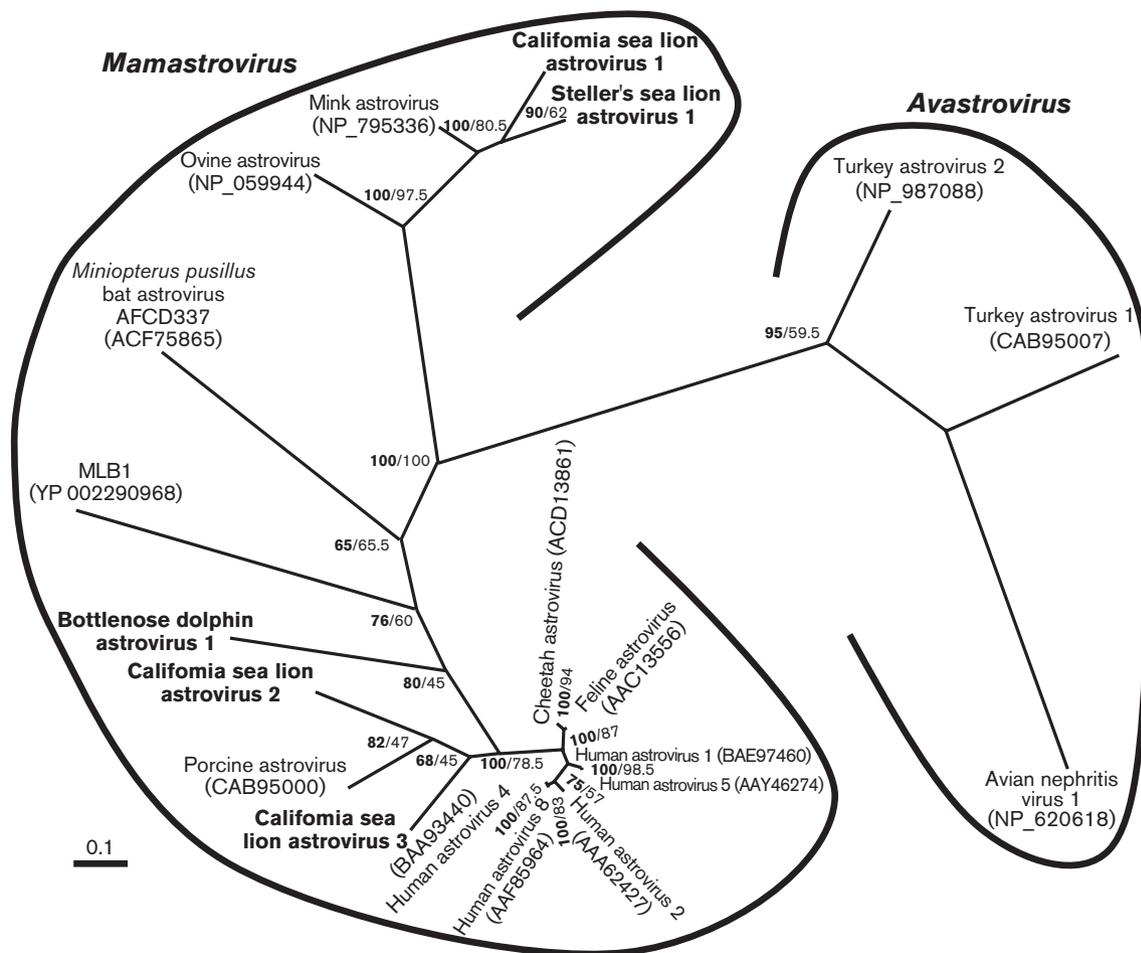
Fig. 4. Bayesian phylogenetic tree of predicted 186–207 aa partial astrovirus capsid sequences, based on T-Coffee alignment. Bayesian posterior probabilities of branchings as percentages are shown in bold, and ML bootstrap values for branchings based on 200 resamplings are given to the right. Avian nephritis virus 1 (GenBank accession no. NP_620618) was designated the outgroup. Virus genera are delineated by brackets. Marine mammal astroviruses are shown in bold. Bar, 0.1 substitutions per site.

central in the ecology of caliciviruses, a better-studied group of small, non-enveloped, positive-stranded RNA viruses (Smith *et al.*, 1998). An understanding of diverse astroviruses in wildlife may enable more appropriate epidemiological responses to novel astrovirus infections in humans.

Our data suggest that a recombination event may have occurred between a human and a marine mammal astrovirus isolate, resulting in CslAstV-3. Recombination is common in other non-enveloped, positive-stranded RNA viruses such as picornaviruses and caliciviruses and, as a result, evolution of structural and non-structural regions of the genome may appear semi-independent (Simmonds, 2006). Previous studies have found evidence for recombination amongst human astrovirus serotypes, as well as turkey astroviruses (Pantin-Jackwood *et al.*, 2006; Simmonds, 2006). The lineage of the region of CslAstV-3

after the area of strong similarity with CslAstV-2 is not clear. This may represent a separate recombination event with an as-yet-unidentified clade of astroviruses. An analysis of the capsid region of a cheetah astrovirus found that it clustered relatively more closely with human astroviruses than was found in the analysis of the polymerase (Atkins *et al.*, 2009). If due to a recombination event, this is an apparently opposite event from that seen in CslAstV-3, which has a human astrovirus-like polymerase and a CslAstV-2-like capsid, at least in the 5′ end of ORF2. The evidence for recombination in the mink astrovirus/CslAstV-1/SslAstV-1 clade is less clear. Data from additional viruses in this clade would be needed to clarify whether a recombination event in these viruses is probable.

Recombination is an important mechanism for rapid evolution of a virus, allowing rapid acquisition of sequence that is less likely to be deleterious than random mutations.
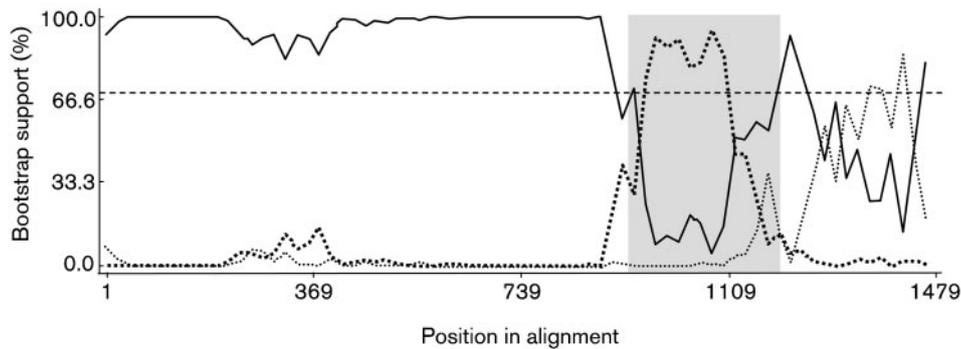
Fig. 5. Bootscanning analysis of MUSCLE alignment of astrovirus sequences, produced by using the RDP3 suite. The solid line indicates bootstrap support for monophyly of human astrovirus 4 and CslAstV-3, with a 200 bp window size and a step size of 20, with 200 bootstrap replicates. The thick dotted line indicates support for monophyly of CslAstV-2 and CslAstV-3. The thin dotted line indicates bootstrap support for monophyly of human astrovirus 4 and CslAstV-2. The grey area is the identified possible recombinant region; the dashed horizontal line indicates 70 % bootstrap support. In addition to the marine mammal viruses found in this study, sequences used in the MUSCLE nucleotide alignment included human astrovirus 1 (GenBank accession no. AY720892), human astrovirus 3 (AF141381), human astrovirus 4 (DQ070852), human astrovirus 5 (DQ028633), human astrovirus MLB1 (FJ222451), mink astrovirus (AY179509), ovine astrovirus (NC_002469) and *Miniopterus pusillus* bat astrovirus AFCD337 (EU847155).

The most common mechanism of recombination among non-enveloped, positive-stranded RNA viruses is switching of the polymerase complex from copying one template to another (Jarvis & Kirkegaard, 1992). Non-replicative RNA recombination has also been shown in picornaviruses, but also requires unencapsulated RNA cotransfected into the same cell (Gmyl *et al.*, 1999). Therefore, a recombinant of a human astrovirus and a sea lion astrovirus implies that a sea lion, a human or a third host species was infected by both viruses at the same point. Emerging disease is frequently associated with host switches. One recent meta-analysis of human diseases found that 816 of 1407 (58 %) are zoonotic; of human diseases, zoonotic diseases are significantly more likely to be emerging (Woolhouse & Gowtage-Sequeria, 2005). Most recent emerging human diseases have been associated with host switches, including SARS, Hendra virus, Nipah virus and AIDS. The aforementioned study also found that viral diseases were much more likely to be emerging, especially RNA viruses (Woolhouse & Gowtage-Sequeria, 2005). The apparent ability of some of these viruses to infect disparate hosts suggests that further study of the ecology and host range of astroviruses may be relevant to human health.

We were unable to clarify the clinical significance of astroviruses for marine mammals from our dataset. Only one sea lion case (CslAstV-3 from a stranded, free-ranging pup) had clinical diarrhoea. However, bacterial culture of the diarrhoeal sample yielded a *Salmonella* sp. isolate. Given the clinical signs associated with other mamastroviruses, it would be reasonable to hypothesize that these viruses may cause a secretory diarrhoea that is most likely to present clinically in neonates and weanlings. Further

studies are under way to determine the clinical relevance of these viruses to their marine mammal hosts.

In conclusion, we have identified five novel astroviruses from marine mammal hosts. These viruses are diverse and all appear to be consistent with novel species. These viruses are situated across the mamastrovirus tree and do not form a monophyletic group. There is evidence of recombination between human and marine mammal astroviruses. Further study of these viruses and their clinical significance in marine mammal populations is indicated.

METHODS

Animals and samples. Faecal samples were collected as part of routine health surveillance from two clinically healthy California sea lions (*Zalophus californianus*) and one bottlenose dolphin (*Tursiops truncatus*) housed in open ocean enclosures at the US Navy Marine Mammal Program in San Diego, CA, USA. In addition, faecal samples were collected from one stranded, free-ranging California sea lion pup with diarrhoea and one stranded, free-ranging Steller sea lion (*Eumetopias jubatus*) pup without outward signs of diarrhoea, both housed at the Marine Mammal Center in Sausalito, CA, USA. All samples were collected between December 2006 and October 2008. Faecal samples were stored in sterile vials and frozen at −80 °C until laboratory analysis.

Negative-staining electron microscopy and sample processing. Upon arrival at the laboratory, each faecal sample was divided into three fractions. The first fraction was sent to the Florida State Diagnostic Laboratory for negative electron microscopy. The second fraction was stored at −80 °C for future analysis. The third fraction was suspended at a 1 : 10 ratio in 0.89 % (w/v) NaCl and centrifuged at 4000 g for 20 min at 4 °C. The clarified supernatant was collected by using a sterile syringe and passed consecutively through 0.8, 0.45 and 0.22 μm syringe filters to eliminate cellular and bacterial particles.

The final filtrate was transferred to a Microsep concentrator column (Pall Life Sciences) and centrifuged at 1500 **g** for 25–45 min at 4 °C. A 140 µl aliquot of the concentrated filtrate was used for RNA extraction using a Viral RNA Mini kit (Qiagen), following the manufacturer's instructions.

Degenerate PCR. Degenerate primers designed based on conserved astroviral sequences (Atkins *et al.*, 2009) were used in a nested or a semi-nested format to amplify conserved regions of ORF1b (RdRp) and ORF2 (capsid). For amplification of the partial RdRp gene, primers Astr4380F (5′-GAYTGGRCNCGNTWYGATGGNACIAT-3′) and Astr4811R (5′-GGYTTNACCCACATNCCAAA-3′) (round 1) and primers Astr4574F (5′-GGNAAYCCMTCWGGICA-3′) and Astr4722R (5′-ARNCKRTCATCNCCATA-3′) (round 2) were used on all five isolates. For individual viruses, additional degenerate primer combinations were used to obtain more astrovirus RdRp gene sequence (see Supplementary Table S1, available in JGV Online). For amplification of the partial capsid gene, primers Astr4811F (5′-TTTGGNATGTGGGTNAARCC-3′) and Astr5819R (5′-TCATTN-GTGTYNGTNANCCACCA-3′) (round 1) and primers Astr5159F (5′-TGGAGGGGMGGACCAAAG-3′) and Astr5819R (round 2) were used on all five isolates.

For the first round of the PCR assays, faecal RNA was reverse-transcribed by using a OneStep RT-PCR kit (Qiagen) at 50 °C for 30 min and then denatured at 94 °C for 5 min, followed by 36 cycles of denaturation at 94 °C for 30 s, annealing at 45 °C for 30 s and extension at 72 °C for 60 s, with a final elongation step at 72 °C for 10 min. Three microlitres of product from the first round was used as template in a 20 µl nested or semi-nested second round. The second-round amplifications conditions, using Platinum *Taq* DNA polymerase (Invitrogen), were as follows: 5 min denaturation at 94 °C, followed by 36 cycles of denaturation at 94 °C (30 s), annealing at 45 °C (30 s) and extension at 72 °C (60 s), with a final elongation step at 72 °C for 10 min. PCR products from both rounds were run in 1 % agarose gels and the DNA bands were visualized under UV light after ethidium bromide staining. Bands of interest were cut from the gel and their DNA was extracted by using a QIAquick gel extraction kit (Qiagen). Direct sequencing was performed using a Big-Dye Terminator kit (Perkin-Elmer) and ABI automated sequencers. All amplicons were sequenced at least twice in both directions, and primer sequences were edited out prior to constructing contiguous sequences.

Sequence extension. For each isolate, the gap between the upstream ORF1b and downstream ORF2 sequence segment was amplified by using specific forward and reverse primers (see Supplementary Table S1) that were designed based on the sequences obtained via degenerate PCR. Again, all amplicons were sequenced at least twice in both directions. Attempts were made to sequence the remaining 5′ and 3′ sections of the virus genome by using GeneRacer (Invitrogen). Specific GeneRacer primers were designed, and virus genomic RNA was amplified by using the manufacturer's instructions. Briefly, for 3′ GeneRacer analysis, RNA was reverse-transcribed using avian myeloblastosis virus reverse transcriptase and amplified with a forward gene-specific primer and the GeneRacer 3′ Primer. For 5′ GeneRacer analysis, RNA was treated with calf intestinal phosphatase, followed by tobacco acid pyrophosphatase to remove the 5′ cap structure, and ligated to the GeneRacer RNA oligo. The dephosphorylated, uncapped and ligated RNA was then reverse-transcribed and subsequently amplified by using a specific reverse gene-specific primer (see Supplementary Table S1) and the GeneRacer 3′ Primer. PCR products were run in a 0.7 % agarose gel and bands of interest were sequenced as described above.

Phylogenetic analysis. Sequences were compared with those in GenBank/EMBL/DDBJ by using TBLASTX (Altschul *et al.*, 1997). The

predicted homologous 237–254 aa sequences of astrovirus RdRp and 186–207 aa sequences of astrovirus capsid protein were aligned by using CLUSTAL W version 2 (Larkin *et al.*, 2007), T-Coffee (Notredame *et al.*, 2000) and MUSCLE (Edgar, 2004).

Bayesian analyses of each alignment were performed by using MrBayes 3.1 (Ronquist & Huelsenbeck, 2003) with gamma-distributed rate variation and a proportion of invariant sites, and mixed amino acid substitution models. The first 10 % of 1 000 000 iterations were discarded as a burn-in.

ML analyses of each alignment were performed by using PHYLIP (Phylogeny Inference Package, version 3.66; Felsenstein, 1989), running each alignment using the program PROML with amino acid substitution models JTT (Jones *et al.*, 1992), PMB (Veerassamy *et al.*, 2003) and PAM (Kosiol & Goldman, 2005), further set with global rearrangements, five replications of random input order, gamma plus invariant rate distributions and unrooted. The values for the gamma distribution were taken from the Bayesian analysis. Avian nephritis virus 1 (GenBank accession no. AB033998) was designated the outgroup. The alignment producing the most likely tree was then used to create data subsets for bootstrap analysis to test the strength of the tree topology (200 resamplings) (Felsenstein, 1985), which were analysed using the amino acid substitution model producing the most likely tree in that alignment.

Recombination analysis. A nucleotide alignment was created by using MUSCLE on the sequence between primers Astr4380F and Astr5819R of 13 mamastroviruses: the five marine mammal astroviruses from this study, ovine astrovirus (GenBank accession no. NC002469), bat astrovirus AFCD337 (EU847155), mink astrovirus (AY179509), MLB1 astrovirus (FJ222451), human astrovirus 1 (AY720892), human astrovirus 3 (AF141381), human astrovirus 4 (DQ070852) and human astrovirus 5 (DQ028633). Potential recombination patterns were screened by using RDP (Martin & Rybicki, 2000), GENECONV (Padidam *et al.*, 1999), MaxChi (Maynard Smith, 1992), Chimaera (Posada & Crandall, 2001) and 3Seq (Boni *et al.*, 2007) in the RDP3 suite (Martin *et al.*, 2005b), using the step-down correction for multiple comparisons and a *P*-value cutoff of 0.05. Regions of potential recombinant interest were also checked with LARD (Holmes *et al.*, 1999), Recscan (Martin *et al.*, 2005a) and SiScan (Gibbs *et al.*, 2000).

ACKNOWLEDGEMENTS

REFERENCES

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, 3389–3402.

Atkins, A., Wellehan, J. F. X., Jr, Childress, A. L., Archer, L. L., Fraser, W. A. & Citino, S. B. (2009). Characterization of an outbreak of

astroviral diarrhea in a group of cheetahs (*Acinonyx jubatus*). *Vet Microbiol* **136**, 160–165.

Boni, M. F., Posada, D. & Feldman, M. W. (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* **176**, 1035–1047.

Caballero, S., Guix, S., Ribes, E., Bosch, A. & Pinto, R. M. (2004). Structural requirements of astrovirus virus-like particles assembled in insect cells. *J Virol* **78**, 13285–13292.

Chu, D. K., Poon, L. L., Guan, Y. & Peiris, J. S. (2008). Novel astroviruses in insectivorous bats. *J Virol* **82**, 9107–9114.

Dennehy, P. H., Nelson, S. M., Spangenberger, S., Noel, J. S., Monroe, S. S. & Glass, R. I. (2001). A prospective case–control study of the role of astrovirus in acute diarrhea among hospitalized young children. *J Infect Dis* **184**, 10–15.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.

Elamri, D. E., Aouni, M., Parnaudeau, S. & Le Guyader, F. S. (2006). Detection of human enteric viruses in shellfish collected in Tunisia. *Lett Appl Microbiol* **43**, 399–404.

Espinosa, A. C., Mazari-Hiriart, M., Espinosa, R., Maruri-Avidal, L., Méndez, E. & Arias, C. F. (2008). Infectivity and genome persistence of rotavirus and astrovirus in groundwater and surface water. *Water Res* **42**, 2618–2628.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Felsenstein, J. (1989). PHYLIP – phylogeny inference package (version 3.2). *Cladistics* **5**, 164–166.

Finkbeiner, S. R., Kirkwood, C. D. & Wang, D. (2008). Complete genome sequence of a highly divergent astrovirus isolated from a child with acute diarrhea. *Virol J* **5**, 117.

Gallimore, C. I., Taylor, C., Gennery, A. R., Cant, A. J., Galloway, A., Lewis, D. & Gray, J. J. (2005). Use of a heminested reverse transcriptase PCR assay for detection of astrovirus in environmental swabs from an outbreak of gastroenteritis in a pediatric primary immunodeficiency unit. *J Clin Microbiol* **43**, 3890–3894.

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.

Gmyl, A. P., Belousov, E. V., Maslova, S. V., Khitrina, E. V., Chetverin, A. B. & Agol, V. I. (1999). Nonreplicative RNA recombination in poliovirus. *J Virol* **73**, 8958–8965.

Holmes, E. C., Worobey, M. & Rambaut, A. (1999). Phylogenetic evidence for recombination in dengue virus. *Mol Biol Evol* **16**, 405–409.

Jarvis, T. C. & Kirkegaard, K. (1992). Poliovirus RNA recombination: mechanistic studies in the absence of selection. *EMBO J* **11**, 3135–3145.

Jonassen, C. M., Jonassen, T., Saif, Y. M., Snodgrass, D. R., Ushijima, H., Shimizu, M. & Grinde, B. (2001). Comparison of capsid sequences from human and animal astroviruses. *J Gen Virol* **82**, 1061–1067.

Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275–282.

Kosiol, C. & Goldman, N. (2005). Different versions of the Dayhoff rate matrix. *Mol Biol Evol* **22**, 193–199.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A. & other authors (2007). CLUSTAL W and CLUSTAL_X version 2. *Bioinformatics* **23**, 2947–2948.

Lukashov, V. V. & Goudsmit, J. (2002). Evolutionary relationships among *Astroviridae*. *J Gen Virol* **83**, 1397–1405.

Madeley, C. R. & Cosgrove, B. P. (1975). 28 nm particles in faeces in infantile gastroenteritis. *Lancet* **2**, 451–452.

Marshall, J. A., Kennett, M. L., Rodger, S. M., Studdert, M. J., Thompson, W. L. & Gust, I. D. (1987). Virus and virus-like particles in the faeces of cats with and without diarrhoea. *Aust Vet J* **64**, 100–105.

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

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

Martin, D. P., Williamson, C. & Posada, D. (2005b). RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* **21**, 260–262.

Maynard Smith, J. (1992). Analyzing the mosaic structure of genes. *J Mol Evol* **34**, 126–129.

Monroe, S. S., Carter, M. J., Hermann, J., Mitchell, D. K. & Sanchez-Fauquier, A. (2005). *Astroviridae*. In *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*, pp. 859–864. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. San Diego: Academic Press.

Moser, L. A., Carter, M. & Schultz-Cherry, S. (2007). Astrovirus increases epithelial barrier permeability independently of viral replication. *J Virol* **81**, 11937–11945.

Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205–217.

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

Pantin-Jackwood, M. J., Spackman, E. & Woolcock, P. R. (2006). Phylogenetic analysis of turkey astroviruses reveals evidence of recombination. *Virus Genes* **32**, 187–192.

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.

Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.

Simmonds, P. (2006). Recombination and selection in the evolution of picornaviruses and other mammalian positive-stranded RNA viruses. *J Virol* **80**, 11124–11140.

Smith, A. W., Skilling, D. E., Cherry, N., Mead, J. H. & Matson, D. O. (1998). Calicivirus emergence from ocean reservoirs: zoonotic and interspecies movements. *Emerg Infect Dis* **4**, 13–20.

Van Hemert, F. J., Lukashov, V. V. & Berkhout, B. (2007). Different rates of (non-)synonymous mutations in astrovirus genes; correlation with gene function. *Virol J* **4**, 25.

Veerassamy, S., Smith, A. & Tillier, E. R. (2003). A transition probability model for amino acid substitutions from blocks. *J Comput Biol* **10**, 997–1010.

Whelan, S. & Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* **18**, 691–699.

Woolhouse, M. E. & Gowtage-Sequeria, S. (2005). Host range and emerging and reemerging pathogens. *Emerg Infect Dis* **11**, 1842–1847.

Zhu, H. C., Chu, D. K., Liu, W., Dong, B. Q., Zhang, S. Y., Zhang, J. X., Li, L. F., Vijaykrishna, D., Smith, G. J. & other authors (2009). Detection of diverse astroviruses from bats in China. *J Gen Virol* **90**, 883–887.