VIROLOGY JOURNAL

Complete genome sequence of an astrovirus identified in a domestic rabbit (*Oryctolagus cuniculus*) with gastroenteritis

Mark D Stenglein[1], Eric Velazquez[1], Cheryl Greenacre[2], Rebecca P Wilkes[3], J Graham Ruby[1], Julia S Lankton[3], Donald Ganem[1,4], Melissa A Kennedy[3] and Joseph L DeRisi[1,4*]

Abstract

A colony of domestic rabbits in Tennessee, USA, experienced a high-mortality (~90%) outbreak of enterocolitis. The clinical characteristics were one to six days of lethargy, bloating, and diarrhea, followed by death. Heavy intestinal coccidial load was a consistent finding as was mucoid enteropathy with cecal impaction. Preliminary analysis by electron microscopy revealed the presence of virus-like particles in the stool of one of the affected rabbits. Analysis using the Virochip, a viral detection microarray, suggested the presence of an astrovirus, and follow-up PCR and sequence determination revealed a previously uncharacterized member of that family. Metagenomic sequencing enabled the recovery of the complete viral genome, which contains the characteristic attributes of astrovirus genomes. Attempts to propagate the virus in tissue culture have yet to succeed. Although astroviruses cause gastroenteric disease in other mammals, the pathogenicity of this virus and the relationship to this outbreak remains to be determined. This study therefore defines a viral species and a potential rabbit pathogen.

Keywords: Rabbit astrovirus, Rabbit virus, Virus discovery, Virochip, Gastroenteritis, Diarrhea virus, Mucoid enteropathy, Enterocolitis

Background

The astroviruses form a family of small, non-enveloped, positive strand RNA viruses that infect a variety of mammalian and avian hosts [1-22]. First identified in human stool samples in 1975, these viruses were named after their star-shaped appearance in some electron micrographs [2,3]. Astroviruses have been shown to replicate in cells of the intestinal tracts of infected organisms, and in some hosts, infection has been demonstrated to cause gastroenteritis [6,13,23-31]. In birds, astroviruses have been linked to intestinal and extra-intestinal pathology [10,32-35].

High throughput, unbiased molecular methods can be used to identify candidate etiologic agents in diseases of unknown origin [36-38]. The Virochip is a DNA microarray that has been used to identify known and to discover novel viruses [39,40]. Next generation sequencing provides a complementary virus discovery method, and is being increasingly used as its cost decreases. The sensitivity of these methods decreases as divergence from known viruses increases and as abundance of the viral nucleic acid in the sample decreases.

In this study, the Virochip microarray was used to screen samples from an outbreak of gastroenteritis in a commercial rabbit colony in Tennessee, USA. A virus was suspected as one possible cause because of virus-like particles in electron microscopic examination of stool from an affected animal, but traditional diagnostic approaches, including virus isolation, failed to identify a candidate viral pathogen. Virochip screening suggested that an astrovirus was present, and subsequently the complete genome was recovered and characterized. Recently, the partial genome sequence of a related virus was recovered from rabbits in Italy [22]. The astroviruses form a diverse family of viruses that infect many animals, and metagenomic investigations such as the

* Correspondence: joe@derisilab.ucsf.edu
[1]Departments of Medicine, Biochemistry and Biophysics, and Microbiology, University of California San Francisco, San Francisco, CA, USA
[4]Howard Hughes Medical Institute, Chevy Chase, MD, USA
Full list of author information is available at the end of the article

one described here will continue to reveal the full extent of their diversity and host range.

Results

Outbreak and clinical diagnostics

A domestic rabbit colony of about 100 animals of various breeds near Johnson City, TN, USA, experienced an outbreak of enterocolitis. The predominant rabbit breeds were angora, satin and mini rex. No changes in diet had occurred in 1.5 years. The rabbits were housed outdoors, on the ground, in an enclosure measuring 8 x 20 feet with an attached hutch. The owner reported voles in the area. Characteristics of the outbreak included a 1–6 day history of lethargy, bloating and diarrhea, followed by death. Initially, animals under 5 months old were affected (~15 in the first 2 weeks), then the adults were affected (~75 over the next 3 weeks). Consistent clinical signs included lethargy, dehydration, increased intestinal sounds with intestinal gas (bloating) and cecal impaction with mucoid diarrhea. Severe intestinal coccidiosis (*Eimeria* spp.) was identified and the entire colony was placed on sulfamethazine in the water (Sulmet, 12.5%, Fort Dodge). Necropsy findings on seven rabbits (2 dead on arrival, 5 euthanized) were all consistent with mucoid enteropathy and cecal impaction. (Table 1) At necropsy, common gross findings within the intestinal tract included mucoid intestinal contents (5/7), cecal impaction (4/7), serosal hemorrhage (4/7) and gastrointestinal gas distention (4/7). Microscopically, common findings included lymphoplasmacytic enteritis (5/7), necrotizing heterophilic enterotyphlocolitis (2/7), intestinal coccidiosis (5/7), and gastric or cecal candidiasis (2/7). In addition to the intestinal coccidiosis identified in all animals tested, one rabbit had hepatic coccidiosis (*Eimeria stediae*), two had *Trichuris* spp., and 4 of 5 small intestinal cultures grew *E. coli* that was found not to be an attaching effacing *E. coli*. 30 nm virus-like particles were observed by electron microscopy in the stool of 1 of 2 animals tested (Figure 1 and Table 1). Follow-up PCR-based testing determined that this virus was not Rabbit Hemorrhagic Disease Virus. Initial attempts to isolate virus from 2 samples in tissue culture were unsuccessful.

Genome recovery and phylogenetic analysis

Because virus-like particles were observed but diagnostic testing failed to identify known viruses, several samples were subjected to Virochip analysis. Total nucleic acid was extracted from the stool of four sick animals and one unaffected control, fluorescently labeled, and hybridized to the Virochip (other case samples lacked sufficient remaining material). To enrich for viral nucleic acid, the stool sample from the EM virus-positive sample was filtered and treated with micrococcal nuclease, to digest non-protected nucleic acid. Virochip analysis of the nuclease-treated sample suggested the presence of astrovirus-related sequence (Table 1 and Figure 2).

To validate this finding, several sets of degenerate primers were designed based on the positive Virochip probe sequences and on conserved regions of astrovirus genomes (Figure 2 and Table 2). PCR using one of these primer pairs produced an amplicon of the expected size, which was cloned and sequenced, confirming that nucleic acid from an apparently divergent astrovirus was present in the sample. Additional degenerate primers and 3′ RACE were used to recover a 5 kbp contiguous sequence. But, repeated attempts using 5′ RACE and other PCR-based strategies failed to recover the full predicted ~7 kb genome.

To recover the remainder of the genome, metagenomic sequencing using the Ion Torrent PGM instrument was performed [41]. This produced 3×10^6 sequences of median length 125 nucleotides (standard deviation 17.5). The average Sanger quality score of these sequences was 19.5 at base 1, 17.3 at base 100, and

Table 1 Summary of diagnostic findings

| Animal #	Enteritis	Died / Euthanized	Parasites	Culture	EM VLPs	Astrovirus Virochip	Rabbit astrovirus RT-PCR
1	+	Died	Intestinal *Eimeria* spp.	No growth	ND	(–)	(–)
2	+	Euthanized	*Trichuris* spp.	*E. coli*	ND	(–)	(–)
				C. perfringens			
3	+	Died	Intestinal *Eimeria* spp.	ND	ND	ND	ND
			Hepatic *Eimeria* spp.				
4	+	Euthanized	Intestinal *Eimeria* spp.	*E. coli*	(–)	(–)	(–)
5	+	Euthanized	Intestinal *Eimeria* spp.	*E. coli*	ND	ND	ND
6	+	Euthanized	Intestinal *Eimeria* spp.	*E. coli*	(+)	(+)	(+)
7	+	Euthanized	Intestinal *Eimeria* spp.	ND	ND	ND	ND
			Trichuris spp.				
8 (Control)	–	ND	ND	ND	ND	(–)	(–)

EM VLPs, virus-like particles evident by electron microscopic examination. ND, not determined. *E. coli, Escherichia coli. C. perfringens, Clostridium perfringens.*
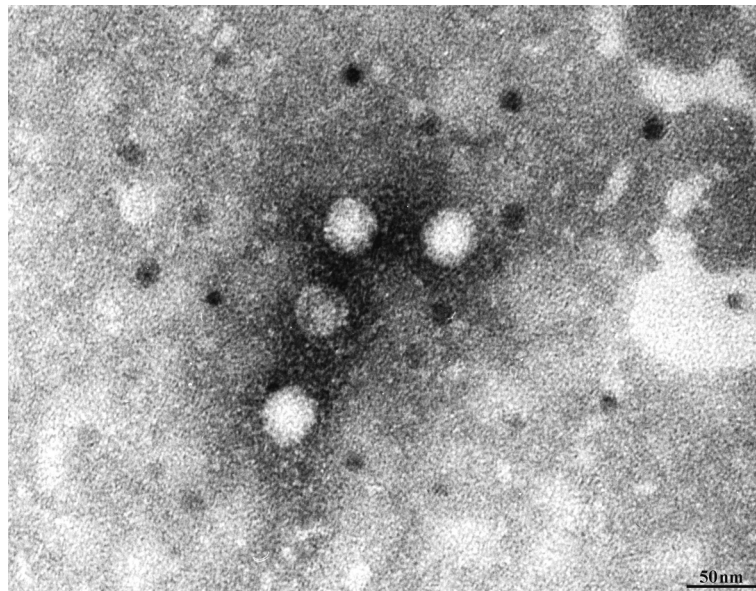
Figure 1 Electron micrograph of virus like particles in the stool of one animal (Table 1). Scale bar indicates 50 nm.

14.4 at base 150, quality values consistent with previously reported metrics for this sequencing platform [42]. Insertions and deletions, which occurred at a frequency of 1.3% (1.3 indels per 100 bases aligned), were more frequent than mismatches, which occurred at a frequency of 0.28%. Retrospective mapping of the sequences revealed that 1.9×10^5 unique sequences (6% of sequences) derived from the astrovirus, corresponding
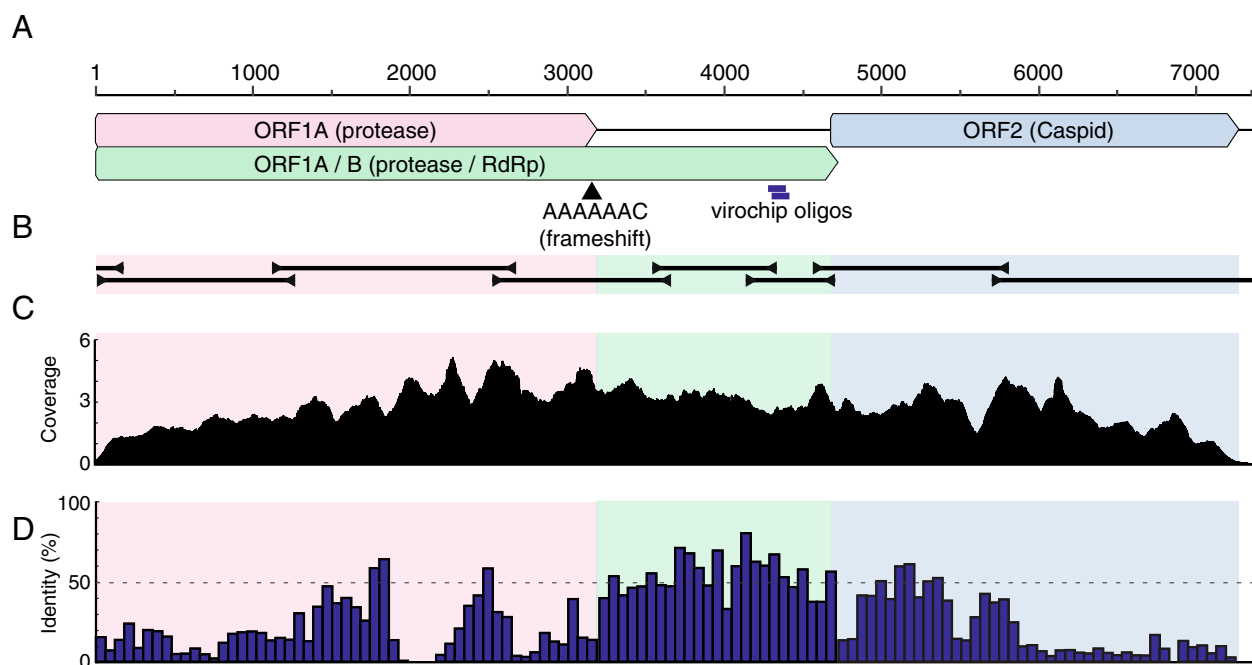
Figure 2 The rabbit astrovirus genome. (**A**) A cartoon showing the organization of the viral genome. Scale bar indicates nucleotide position. The location of a putative ribosomal frameshift motif is indicated by an arrowhead. The position of the two overlapping Virochip probes that were originally positive for an astrovirus are indicated. (**B**) Location of overlapping PCR, 5′, and 3′ RACE amplicons used to determine and confirm the sequence of the viral genome. (**C**) A histogram showing the number of Ion Torrent reads (x1000 reads) aligning at each nucleotide position in the genome. (**D**) The average pairwise percent identity between the predicted rabbit astrovirus protein sequences and the sequences of the proteins from the astroviruses shown in Figure 3. The histogram is binned into 20 amino acid windows.

**Table 2 Oligonucleotides used in this study**

Oligo name	Note	Sequence (5′ – 3′)
MDS-4	Reverse transcription	CGCTCTTCCGATCTNNNNNN
MDS-189	Library amplification	CTGTCTGGCTCTTCCGATCT
MDS-115	Degenerate astrovirus consensus	CCATCAGGTCARWWNTCAACAAC
MDS-116	Degenerate astrovirus consensus	CTCGCTAATHTAYGGDGATGA
MDS-117	Degenerate astrovirus consensus	GGTTTNACCCACATNCCAAA
MDS-118	Degenerate astrovirus consensus	GTCCAACTGTGADNCCACARAA
MDS-119	Rabbit astrovirus diagnostic	ATAATAACATGGTCAACTATTGGCTTC
MDS-120	Rabbit astrovirus diagnostic	GACATCCTTATACATTTTCACAACTTT
MDS-156	Vertebrate rRNA (L2513)	GCCTGTTTACCAAAAACATCAC
MDS-157	Vertebrate rRNA (H2714)	CTCCATAGGGTCTTCTCGTCTT
MDS-298	Genome recovery and sequencing (27–46 forward)	CGGCCAGAGAGCCTATTACC
MDS-296	5′ RACE (149–168 reverse)	GGCTAGAGGAATGGGGTCAG
MDS-299	Genome recovery and sequencing (1146–1165 forward)	GCTCGTTCCGATAATTGCTC
MDS-300	Genome recovery and sequencing (1229–1248 reverse)	GCAACTAACCACGCACAATG
MDS-144	Genome recovery and sequencing (2554–2564 forward)	TCTGGACYGARGARGARTA
MDS-300	Genome recovery and sequencing (2639–2659 reverse)	CACTCCATTCAGGGTAACCAA
MDS-135	Genome recovery and sequencing (3573–3592 forward)	TCCACTCCCGCCTACCCCAA
MDS-145	Genome recovery and sequencing (3622–3641 reverse)	CCACCCACAATCAGAGAGGT
MDS-119	Genome recovery and sequencing (4165–4191 forward)	ATAATAACATGGTCAACTATTGGCTTC
MDS-125	Genome recovery and sequencing (4287–4308 reverse)	CATAATCAGATGGGAGGACAGG
MDS-147	Genome recovery and sequencing (4602–4620 forward)	CTTCGCAGCCACTCTCTTG
MDS-137	Genome recovery and sequencing (4668–4683 reverse)	TTGGTCCTCCCCTCCA
MDS-138	3′ RACE (5719–5729 forward)	TGGTGGTTTGT
MDS-150	Genome recovery and sequencing (5759–5779 reverse)	TCTGAAATTGCACTGTGTTGG
MDS-121	RACE oligo-dT primer	CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGCT$_{17}$
MDS-122	RACE outer adapter primer	CCAGTGAGCAGAGTGACG
MDS-123	RACE inner adapter primer	GAGGACTCGAGCTCAAGC

to >1000x average coverage of its genome (Figure 2). The PRICE de novo targeted assembly software (Graham Ruby, http://derisilab.ucsf.edu/software/price/) was used to assemble the entire viral genome, the sequence of which was validated by RT-PCR and Sanger sequencing (Figure 2). 5′ RACE was used to confirm that the Ion Torrent-based assembly extended to the 5′ terminus of the genome (Figure 2). The remainder of the reads in the metagenomic dataset consisted mainly of sequences derived from bacteria (~76% of taxonomically assigned reads), with the *Escherichia, Akkermansia, Clostridium, Salmonella, Enterobacter, and Shigella* genera being most prevalent. 56 sequences (0.002% of all sequences) aligned best to Eimeria species, consistent with the clinical diagnosis of coccidiosis (Table 1). A small percentage (0.4%) of human sequences were also evident in the data. These were likely laboratory contaminants, as the majority of them aligned to the breast cancer 1 and 2, early onset genes (BRCA1 and BRCA2), and targeted sequencing of

these genes was conducted in the laboratory where the sequencing libraries were prepared. Library contamination was also the likely source of 27 reads with nucleotide sequences nearly identical to human immunodeficiency virus-1 (HIV-1) database sequences.

In an attempt to propagate and isolate the virus, stool from the positive sample was filtered and used to inoculate a rabbit kidney cell line (RK) culture and cultures of several other mammalian cell lines previously shown to support the replication of human astroviruses: Vero, HT-29, and Caco-2 [43-45]. These cultures were maintained for 15 days, and culture supernatants were collected on days 1, 2, 4, 6, 10, and 15 post innoculation. Rabbit astrovirus RNA was detected in the day 1 supernatant (in the innoculum) by RT-PCR but not in later time points, suggesting that the virus did not replicate in these cultures.

A phylogenetic analysis of the predicted rabbit astrovirus protein sequences was performed. The rabbit astrovirus
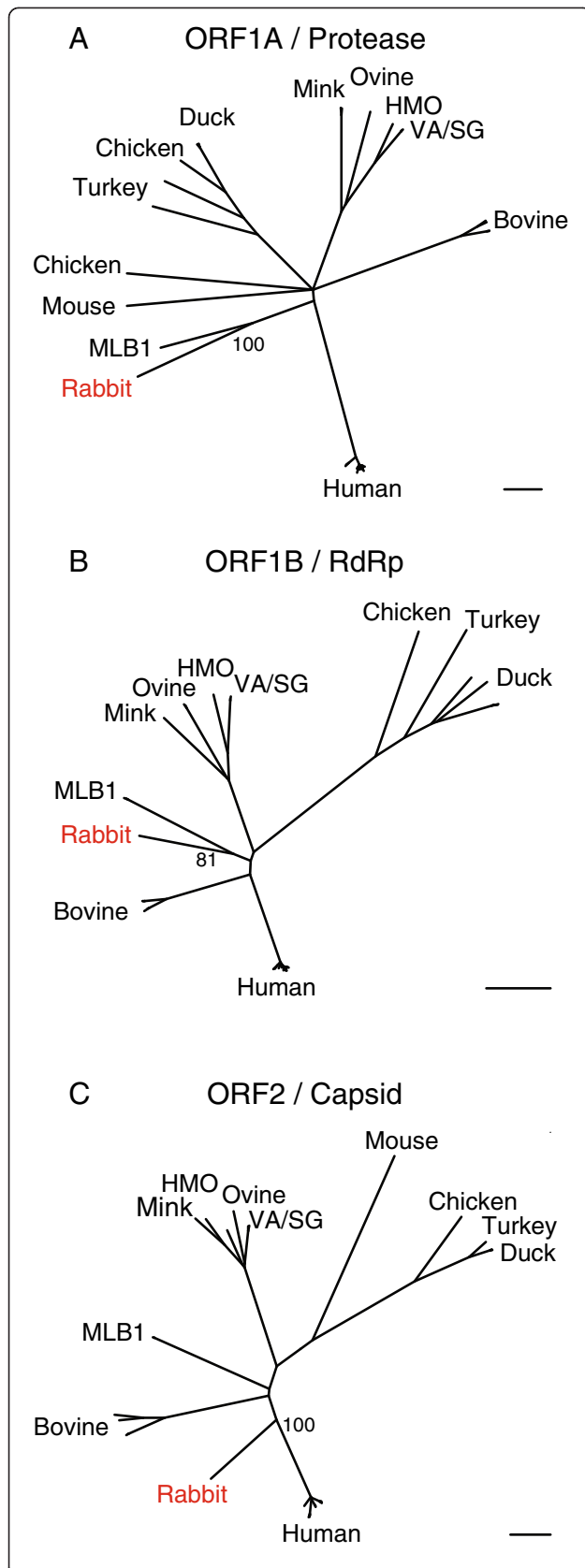
Figure 3 Maximum likelihood phylogenies created from multiple sequence alignments of the predicted rabbit astrovirus proteins and the related proteins of all astroviruses in Genbank with complete genome sequences. (**A**) Phylogeny based on ORF1A sequences. (**B**) Phylogeny based on ORF1B sequences. (**C**) Phylogeny based on the first ~420 amino acids of the indicated ORF2 (capsid) sequences. See Methods for accession numbers of sequences used. The bootstrap percentages of selected nodes are indicated. Scalebars indicate 0.2 amino acid substitutions per site.

sequences were compared to those of all the astroviruses in GenBank for which a complete genome sequence exists (see Methods). Analysis with the three major viral polyproteins, nsp1A, nsp1B (RdRp), and capsid (see Discussion) produced overall similar tree topologies (Figure 3). The rabbit astrovirus sequences branch basally on the phylogram, with the closest related sequences being those of Astrovirus MLB1, which was isolated from a human [46]. The rabbit virus nsp1A, nsp1B, and capsid proteins share 30%, 59%, and 25% pairwise amino acid identity with the corresponding Astrovirus MLB1 protein sequences. In the phylogenies based on nsp1A and nsp1B, the clade formed by the rabbit virus and Astrovirus MLB1 is supported by 100% and 81% of boostrap replicates, respectively (Figure 3A and B). In contrast, the rabbit virus capsid sequence does not form a well-supported clade with the astrovirus MLB1 capsid sequences, but instead branches along the lineage leading to the canonical human astroviruses.

Diagnostic primers were designed to specifically detect rabbit astrovirus (MDS-119 and –120; Table 2). RT-PCR using these primers indicated that, of the five stool samples collected (from four sick and one control rabbit), only the EM-positive sick rabbit tested positive for the virus (Table 1).

Discussion

In this study, we describe the identification and genomic characterization of a member of the *Astroviridae* family of viruses. Traditionally, astroviruses have been named after the host species from which they were identified, so we propose naming this virus species *Astrovirus rabbit/TN/2009/USA*.

This virus was isolated from the feces of a rabbit suffering from enterocolitis, but it is not clear that the virus caused the disease in this instance. Indeed, samples from three of the other four sick rabbit from the same outbreak tested negative for the virus (Table 1). Nevertheless, astroviruses cause gastroenteritis in other mammals, and it is possible that this virus caused or exacerbated the illness in a subset of the animals [13,24,25,28,29,47]. The coccidian parasites and bacteria evident in the samples may have also contributed to the

enteritis and from this data alone, it is impossible to ascertain whether this astrovirus is capable of causing disease. Unfortunately, the sample material was limited, and attempts to propagate the virus in culture were unsuccessful, so experiments to directly demonstrate disease causality were not possible. It may be possible to generate an infectious clone of the virus, as has been accomplished for other astrovirus species [48]. Screening of additional rabbits with diarrhea will provide additional epidemiological information and another source of virus for follow-up experiments.

Indeed, as this manuscript was being prepared, the partial sequence of an astrovirus recovered from rabbits in Italy was reported [22]. In this report, viral RNA was detected by RT-PCR in 10 of 23 rabbits (43%) with enteritis and in 18% of 139 asymptomatic animals, and the mean copy number was ~100-fold higher in the feces of the sick rabbits compared to the apparently healthy animals. The 3395 nt sequence (accession JN052023) in the Martella *et al.* report includes a portion of ORF1B and ORF2, and is 92% identical at the nucleotide level to the sequence recovered from the Tennessee rabbit. Thus, similar astroviruses appear to be a common and geographically widespread infection of rabbits, and an increased viral load is associated with enteritis.

Astroviruses were so named because of their star-shaped appearance in electron micrographs [3]. Although non-enveloped virus-like particles of a size typical of astroviruses (~30 nm) were observed in the astrovirus RNA-positive sample, it is not clear that these particles correspond to the astrovirus. It is important to note that as few as 10% of particles in a preparation may exhibit the characteristic star-shaped morphology [3].

The rabbit astrovirus genome contains characteristics typical of astrovirus genomes, and it is therefore possible to make predictions about viral gene expression and function [1]. Like other astroviruses, the rabbit astrovirus genome is about 7 kb in length and contains three large open reading frames, named ORF1A, ORF1B, and ORF2. It is likely that ORF1A can be translated by itself or in conjunction with ORF1B via a −1 ribosomal framseshift at a conserved AAAAAAC motif (Figure 2) [49-51]. ORF1A and ORF1AB encode the viral non structural polyproteins nsp1A and nsp1ab. ORF2 encodes the viral structural capsid polyprotein. These polyproteins are predicted to be proteolytically processed into multiple domains with discrete functions [52-57].

Astrovirus nsp1A proteins contain protease domains, but apart from this their function is not well characterized. According to NCBI's Conserved Domain Database search tool (CDD) the predicted rabbit astrovirus nsp1a protein contains two conserved domains: a serine protease motif, and a chromosome segregation protein DNA-binding domain [58]. This DNA binding-related domain

may bind RNA and be involved in viral replication. As has been reported for other astroviruses, the nsp1A protein is predicted to contain several transmembrane and coiled coil domains [59]. It is likely that the nsp1A contains other functional domains of as yet unknown function. rabbit astrovirus nsp1A contains a 76 amino acid insertion between residues 674 and 749 not found in related protein sequences. These residues contain no obvious identifiable domain.

ORF1B is predicted to encode nsp1B, which results from proteolytic cleavage of nsp1ab. Nsp1b is the RNA-dependent RNA polymerase (RdRp), and this is the only domain identifiable by CDD or BLAST search. The nsp1B protein is the most conserved of the three polyproteins (Figure 2).

ORF2 encodes the viral structural (capsid) precursor polyprotein. As is the case with other astroviruses, the structural polyprotein may be translated from a subgenomic RNA, although no evidence of such a species is evident in the deep sequencing coverage levels (Figure 2) [60,61]. It is possible that the subgenomic RNA is present is at higher levels in infected cells than in extracellular virus particles, which are likely the source of the filtered material sequenced. The region defined by the first 70 amino acids of the ORF2-encoded protein is basic, with a predicted isoelectric point of 13.3. This is consistent with related astrovirus proteins, in which this domain has been reported to be involved in nucleic acid binding [55]. The N-terminal ~450 amino acids of the capsid protein align by BLAST with other astrovirus capsid proteins, but no recognizable protein domains are contained in the C-terminal 400 amino acids, and the latter half of the protein exhibits a much lower degree of sequence similarity to other astrovirus proteins sequences (Figure 2). This is in agreement with the model that the N-terminal half of astrovirus capsid polyproteins form a conserved core domain and the C-terminal halves of the proteins are highly variable receptor interacting domains [62]. ORF1A is preceeded by a 12 nt untranslated region, which is short but within the size range (11–85 nt) previously reported for astroviruses [1]. The 87 nt predicted 3′ untranslated region is followed by a polyA tail.

Conclusions

Here, we identify and fully characterize the genome of an astrovirus present in the stool of a rabbit suffering from a fatal intestinal disease. This was enabled by the combination of two complementary genomics techniques: the Virochip microarray and high throughput sequencing on the Ion Torrent PGM platform. Although the precise contribution of this virus to the observed pathology remains to be determined, related viruses cause similar disease in other mammals, and this virus is

a plausible etiologic agent. The molecular information presented here sets the stage for confirmatory follow-up studies.

Methods

Clinical methods

Fresh fecal samples were submitted to the Clinical Virology Laboratory (University of Tennessee Veterinary Medical Center) for electronmicroscopic examination. Samples were processed as described in [63]. Briefly, samples were centrifuged in distilled water at approximately 15,000×g for one hour. The supernatant was discarded and pellet resuspended in 2–3 ml distilled water. 100ul of the resuspended material was added to 1 ml distilled water containing 3% phosphotungstic acid. The solution was aerosolized onto a carbonate-coated copper grid and examined by EM at 30,000 and 300,000X.

Nucleic acid extraction

Total nucleic acid was extracted from stool samples using the Qiagen QiaAmp DNA mini kit, which purifies RNA and DNA, according to the manufacturer's protocol. Additionally, the EM VLP-positive stool sample (sample ID 10–2208) was diluted in PBS, clarified by centrifugation at 10,000 g for 1 minute, and passed through a 0.22 um filter (Millipore). The Filtrate was treated with microccoal nuclease (MNase; NEB) in order to degrade non-encapsidated (or otherwise not protected) nucleic acid. MNase reactions containing 1x reaction buffer, 100 µg/ml BSA, and 4, 1, 0.2, or 0 units MNase were incubated at room temperature for 15 minutes. MNase was inactivated by addition of EGTA to 20 mM.

Library preparation

For virochip analysis, randomly-primed libraries were prepared from total nucleic acid. RNA was reverse transcribed in reactions containing 1x reaction buffer, 5 mM dithiothreitol, 1.25 mM dNTPs, 20 pmoles primer MDS-4 (Table 2), 100 U Superscript III (Invitrogen), and 100 ng template. Following reverse transcription, Sequenase reaction buffer and 2 U of Sequenase DNA polymerase (Affymetrix) were added to samples for second strand synthesis. The Sequenase reactions were performed twice so that starting DNA templates would be converted into end-tagged library molecules. The resulting libraries were amplified by PCR using primer MDS-189 (Table 2). PCRs contained 1x reaction buffer, 2 µM primer, 0.25 mM dNTPs, 2 U taq DNA polymerase, and 2 µl library template. Thermocycling was: 95°C for 2 min; 2 cycles of 95°C for 30 sec, 40°C for 30 sec, and 72°C for 1 min, then 25 cycles of 95°C for 30 sec, 58°C for 30 sec, and 72°C for 1 min, with a final extension of 72°C for 5 min.

For microarray hybridzation, a fraction of each library was amplified by PCR as above but with a modified dNTP mixture including 5-(3-aminoallyl)-dUTP (Ambion) in lieu of 75% of the dTTP normally in the mixture. The resulting amino-allyl-containing DNA was purified using a DNA Clean and Concentrator-5 column (DNA-CC-5; Zymo Research). The eluate was heat denatured at 95°C for 2 min, cooled briefly on ice, then fluorescently labeled in reactions containing 100 mM sodium bicarbonate pH 9, 10% DMSO, and 667 µM Cy3 mono NHS ester (GE Healthcare) for 1 hour at 25 °C. Labeled DNA was purified using the DNA-CC-5 columns and added to hybridization reactions containing 3xSSC, 25 mM HEPES pH 7.4, and 0.25% SDS. Hybridization mixtures were heated at 95°C for 2 minutes then applied to microarrays and hybridized overnight at 65 °C. Following hybridization, arrays were washed twice in 0.57xSSC and 0.028% SDS and twice in 0.057x SSC, then scanned on an Axon GenePix 4000B microarray scanner. Three tools were used to analyze Virochip data: E-predict [32], Z-score analysis [33], and cluster analysis [34].

PCR/Sanger sequencing

Oligonucleotide sequences used to amplify the rabbit astrovirus genome are listed in Table 2. PCRs contained 1x reaction buffer, 2 µM primer, 0.25 mM dNTPs, 2 U Taq DNA polymerase, and 2 µl library template. Thermocycling was: 95°C for 2 min, then 30 cycles of 95°C for 30 sec, 58°C for 30 sec, and 72°C for 2 min, with a final extension of 72°C for 5 min. Amplicons were purified from agarose gels using the PureLink gel extraction kit (Invitrogen) and cloned into the pCR2.1 TOPO vector (Invitrogen) according to the manufacturer's protocols. Cloned amplicons were sequenced on an ABI 3700 instrument. For diagnostic testing, primers MDS-119 and –120 (Table 1) were used with reaction and thermocycling as described above.

5′ and 3′ RACE

5′ and 3′ RACE were performed essentially as described [64,65], with primers listed in Table 2. RACE amplicons were cloned and sequenced as described above.

Ion Torrent metagenomic sequencing

Randomly primed cDNA was prepared in reverse transcription reactions containing 1x reaction buffer, 5 mM dithiothreitol, 1.25 mM dNTPs, 20 pmoles random hexamer primer, 100 U Superscript III (Invitrogen), and 100 ng RNA template. Second strand DNA was synthesized as above using Sequenase DNA polymerase. cDNA was adapted for Ion Torrent sequencing using the Ion Xpress Library Kit (catalog # 4468987) according to the manufacturer's protocols. Although this protocol includes PCR amplification, identical reads were removed during

analysis to eliminate possible PCR duplicates. Sequencing was performed on an Ion Personal Genome Machine on one model 314 chip and one model 316 chip according to the manufacturer's protocols.

Tissue culture

100 μl of stool positive for rabbit astrovirus RNA by RT-PCR was diluted 1:10 in PBS and filtered through a 0.22 μm filter (Millipore). Filtrate was used to inoculate cultures of RK, Vero, HT-29, or Caco-2 cells, which were grown in DMEM supplemented with 10% fetal bovine serum and 50 units/ml penicillin and 50 μg/ml streptomycin at 37°C and 5% CO_2. Cultures were supplemented with 0, 1, 10, or 100 μg/ml trypsin, which has been shown to be necessary for astrovirus infectivity in tissue culture. Culture supernatant was harvested and replaced at the indicated time points and stored at −80°C until processing. RNA was extracted from supernatant and reverse transcribed as described above. PCR using primers MDS-119 and −120 for rabbit astrovirus and MDS-156 and −157 for rabbit rRNA (as a positive RT-PCR control) was used to detect viral RNA, with reaction conditions as described above (See Table 2 and [66]).

Bioinformatics

Astrovirus protein sequences for all astroviruses with complete genome sequences were downloaded from Gen-Bank. The accessions of these sequences are: AB308374, AF260508, AY179509, AY720891, AY720892, DQ028633, DQ070852, DQ344027, EU111742, FJ222451, FJ402983, FJ434664, FJ755402, FJ755403, FJ755404, FJ755405, FJ919225, FJ919226, FJ919227, FJ919228, FJ973620, GQ415660, GQ495608, GQ891990, GU985458, HM237363, HM450380, HQ398856, HQ916313, HQ916314, HQ916316, HQ916317, JF414802, JF755422, L23513, NC_001943, NC_002469, NC_002470, NC_003790, NC_004579, NC_005790, NC_010646, NC_011400, NC_012437, NC_013060, NC_013443, NC_014320, NC_015935, and Y15937. Sequences were aligned using Clustal (version 2.0.12) using default parameters [67]. Alignments were manually inspected and trimmed to the point of clear homology. The ORF2 (capsid) alignment is based on the first ~420 amino acids of the protein sequences, which correspond to the relatively conserved core region of the capsid protein. Neighbor joining trees were constructed using PhyML software (PhyML plugin for Geneious version 2.0.12) using 100 bootstrap replicates and default parameters [68]. The sequence of the rabbit astrovirus has been deposited in Gen-Bank, with accession JF729316.

The BLAST alignment software (version 2.2.25+) was used to taxonomically categorize the sequences in the Ion Torrent dataset [69]. The NCBI non-redundant nucleotide database was searched using the blastn algorithm with an expect value cutoff of 1e-6. For each query producing an alignment, the taxonomic ID of the best alignment was determined and tallied. The Paired-Read Iterative Contig Extenion (PRICE, version 0.13) *de novo* assembler (freely available at: http://derisilab.ucsf.edu/software/price/index.html) was used to assemble the rabbit astrovirus genome from the Ion Torrent dataset.

Several tools were combined to determine coverage, quality, and error metrics. First, CD-HIT was used to collapse identical sequences (cd-hit-est version 4.5.4 run with parameter −c 1), which may be PCR duplicates [70]. Then, the bowtie2 aligner was used to map unique reads to the Sanger-verified viral genome assembly (version 2.0.0 run in −local mode with otherwise default parameters; [71]). The number of gap openings (XO field) and mismatches (XM field) and the total number of aligned bases in the bowtie2 SAM output were tallied to determine the frequency of indels and mismatches. Average per base quality scores were determined directly from the Ion Torrent FASTQ output.

Author details

[1]Departments of Medicine, Biochemistry and Biophysics, and Microbiology, University of California San Francisco, San Francisco, CA, USA. [2]Department of Small Animal Clinical Sciences, University of Tennessee College of Veterinary Medicine, Knoxville, TN, USA. [3]Department of Biomedical and Diagnostic Sciences, University of Tennessee College of Veterinary Medicine, Knoxville, TN, USA. [4]Howard Hughes Medical Institute, Chevy Chase, MD, USA.

References

1. Méndez E, Arias CF: **Astroviruses**. In *Fields Virology*. 5th edition. Edited by Knipe DM, Howley PM. Philadelphia, PA, USA: Lippincott Williams & Wilkins; 2007:981–1000.
2. Madeley CR, Cosgrove BP: **Letter: Viruses in infantile gastroenteritis**. *Lancet* 1975, **2**:124.

3. Madeley CR, Cosgrove BP: **Letter: 28 nm particles in faeces in infantile gastroenteritis.** *Lancet* 1975, **2**:451–452.
4. Appleton H, Higgins PG: **Letter: Viruses and gastroenteritis in infants.** *Lancet* 1975, **1**:1297.
5. De Benedictis P, Schultz-Cherry S, Burnham A, Cattoli G: **Astrovirus infections in humans and animals - Molecular biology, genetic diversity, and interspecies transmissions.** *Infect Genet Evol* 2011, **11**:1529–1544.
6. Snodgrass DR, Angus KW, Gray EW, Menzies JD, Paul G: **Pathogenesis of diarrhoea caused by astrovirus infections in lambs.** *Arch Virol* 1979, **60**:217–226.
7. Williams FP Jr: **Astrovirus-like, coronavirus-like, and parvovirus-like particles detected in the diarrheal stools of beagle pups.** *Arch Virol* 1980, **66**:215–226.
8. Tzipori S, Menzies JD, Gray EW: **Detection of astrovirus in the faeces of red deer.** *Vet Rec* 1981, **108**:286.
9. Bridger JC, Hall GA, Brown JF: **Characterization of a calici-like virus (Newbury agent) found in association with astrovirus in bovine diarrhea.** *Infect Immun* 1984, **43**:133–138.
10. Gough RE, Collins MS, Borland E, Keymer LF: **Astrovirus-like particles associated with hepatitis in ducklings.** *Vet Rec* 1984, **114**:279.
11. Shirai J, Shimizu M, Fukusho A: **Coronavirus-, calicivirus-, and astrovirus-like particles associated with acute porcine gastroenteritis.** *Nippon Juigaku Zasshi* 1985, **47**:1023–1026.
12. Reynolds DL, Saif YM: **Astrovirus: a cause of an enteric disease in turkey poults.** *Avian Dis.* 1986, **30**:728–735.
13. Harbour DA, Ashley CR, Williams PD, Gruffydd-Jones TJ: **Natural and experimental astrovirus infection of cats.** *Vet Rec* 1987, **120**:555–557.
14. Atkins A, Wellehan JFX, Childress AL, Archer LL, Fraser WA, Citino SB: **Characterization of an outbreak of astroviral diarrhea in a group of cheetahs (Acinonyx jubatus).** *Vet Microbiol* 2009, **136**:160–165.
15. Mittelholzer C, Hedlund K-O, Englund L, Dietz H-H, Svensson L: **Molecular characterization of a novel astrovirus associated with disease in mink.** *J Gen Virol* 2003, **84**:3087–3094.
16. Baxendale W, Mebatsion T: **The isolation and characterisation of astroviruses from chickens.** *Avian Pathol.* 2004, **33**:364–370.
17. Zhu HC, Chu DKW, Liu W, Dong BQ, Zhang SY, Zhang JX, Li LF, Vijaykrishna D, Smith GJD, Chen HL, Poon LLM, Peiris JSM, Guan Y: **Detection of diverse astroviruses from bats in China.** *J Gen Virol* 2009, **90**:883–887.
18. Rivera R, Nollens HH, Venn-Watson S, Gulland FMD, Wellehan JFX: **Characterization of phylogenetically diverse astroviruses of marine mammals.** *J Gen Virol* 2010, **91**:166–173.
19. Chu DKW, Chin AWH, Smith GJ, Chan K-H, Guan Y, Peiris JSM, Poon LLM: **Detection of novel astroviruses in urban brown rats and previously known astroviruses in humans.** *J Gen Virol* 2010, **91**:2457–2462.
20. Phan TG, Kapusinszky B, Wang C, Rose RK, Lipton HL, Delwart EL: **The fecal viral flora of wild rodents.** *PLoS Pathog* 2011, **7**:e1002218.
21. Zhao W, Zhu AL, Yuan CL, Yu Y, Zhu CX, Lan DL, Yang ZB, Cui L, Hua XG: **Detection of astrovirus infection in pigeons (Columbia livia) during an outbreak of diarrhoea.** *Avian Pathol* 2011, **40**:361–365.
22. Martella V, Moschidou P, Pinto P, Catella C, Desario C, Larocca V, Circella E, Bànyai K, Lavazza A, Magistrali C, Decaro N, Buonavoglia C: *Astroviruses in Rabbits.* Emerging Infectious Diseases 2011, **17**:2287–2293.
23. Moser LA, Schultz-Cherry S: **Pathogenesis of astrovirus infection.** *Viral Immunol.* 2005, **18**:4–10.
24. Kurtz JB, Lee TW, Craig JW, Reed SE: **Astrovirus infection in volunteers.** *J Med Virol* 1979, **3**:221–230.
25. Midthun K, Greenberg HB, Kurtz JB, Gary GW, Lin FY, Kapikian AZ: **Characterization and seroepidemiology of a type 5 astrovirus associated with an outbreak of gastroenteritis in Marin County.** *California. J. Clin. Microbiol.* 1993, **31**:955–962.
26. Herrmann JE, Taylor DN, Echeverria P, Blacklow NR: **Astroviruses as a cause of gastroenteritis in children.** *N Engl J Med* 1991, **324**:1757–1760.
27. Herrmann JE, Hudson RW, Perron-Henry DM, Kurtz JB, Blacklow NR: **Antigenic characterization of cell-cultivated astrovirus serotypes and development of astrovirus-specific monoclonal antibodies.** *J Infect Dis* 1988, **158**:182–185.
28. Oishi I, Yamazaki K, Kimoto T, Minekawa Y, Utagawa E, Yamazaki S, Inouye S, Grohmann GS, Monroe SS, Stine SE: **A large outbreak of acute gastroenteritis associated with astrovirus among students and teachers in Osaka.** *Japan. J. Infect. Dis.* 1994, **170**:439–443.
29. Shimizu M, Shirai J, Narita M, Yamane T: **Cytopathic astrovirus isolated from porcine acute gastroenteritis in an established cell line derived from porcine embryonic kidney.** *J Clin Microbiol* 1990, **28**:201–206.
30. Sebire NJ, Malone M, Shah N, Anderson G, Gaspar HB, Cubitt WD: **Pathology of astrovirus associated diarrhoea in a paediatric bone marrow transplant recipient.** *J Clin Pathol* 2004, **57**:1001–1003.
31. Woode GN, Pohlenz JF, Gourley NE, Fagerland JA: **Astrovirus and Breda virus infections of dome cell epithelium of bovine ileum.** *J Clin Microbiol* 1984, **19**:623–630.
32. Koci MD, Schultz-Cherry S: **Avian astroviruses.** *Avian Pathol.* 2002, **31**:213–227.
33. Imada T, Yamaguchi S, Mase M, Tsukamoto K, Kubo M, Morooka A: **Avian nephritis virus (ANV) as a new member of the family Astroviridae and construction of infectious ANV cDNA.** *J Virol* 2000, **74**:8487–8493.
34. Thouvenelle ML, Haynes JS, Reynolds DL: **Astrovirus infection in hatchling turkeys: histologic, morphometric, and ultrastructural findings.** *Avian Dis.* 1995, **39**:328–336.
35. Behling-Kelly E, Schultz-Cherry S, Koci M, Kelley L, Larsen D, Brown C: **Localization of astrovirus in experimentally infected turkeys as determined by in situ hybridization.** *Vet Pathol* 2002, **39**:595–598.
36. Relman DA: **Microbial genomics and infectious diseases.** *N Engl J Med* 2011, **365**:347–357.
37. Tang P, Chiu C: **Metagenomics for the discovery of novel human viruses.** *Future Microbiol* 2010, **5**:177–189.
38. Lipkin WI: **Microbe hunting.** *Microbiol Mol Biol Rev* 2010, **74**:363–377.
39. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, Derisi JL: **Microarray-based detection and genotyping of viral pathogens.** *Proc Natl Acad Sci USA* 2002, **99**:15687–92.
40. Kistler AL, Gancz A, Clubb S, Skewes-Cox P, Fischer K, Sorber K, Chiu CY, Lublin A, Mechani S, Farnoushi Y, Greninger A, Wen CC, Karlene SB, Ganem D, Derisi JL: **Recovery of divergent avian bornaviruses from cases of proventricular dilatation disease: identification of a candidate etiologic agent.** *Virol J* 2008, **5**:88.
41. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Leamon JH, Johnson K, Milgrew MJ, Edwards M, Hoon J, Simons JF, Marran D, Myers JW, Davidson JF, Branting A, Nobile JR, Puc BP, Light D, Clark TA, Huber M, Branciforte JT, Stoner IB, Cawley SE, Lyons M, Fu Y, Homer N, Sedova M, Miao X, Reed B, Sabina J, Feierstein E, Schorn M, Alanjary M, Dimalanta E, Dressman D, Kasinskas R, Sokolsky T, Fidanza JA, Namsaraev E, McKernan KJ, Williams A, Roth GT, Bustillo J: **An integrated semiconductor device enabling non-optical genome sequencing.** *Nature* 2011, **475**:348–352.
42. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ: **Performance comparison of benchtop high-throughput sequencing platforms.** *Nat Biotechnol* 2012, **30**:434–439.
43. Brinker JP, Blacklow NR, Herrmann JE: **Human astrovirus isolation and propagation in multiple cell lines.** *Arch Virol* 2000, **145**:1847–1856.
44. Lee TW, Kurtz JB: **Serial propagation of astrovirus in tissue culture with the aid of trypsin.** *J Gen Virol* 1981, **57**:421–424.
45. Willcocks MM, Carter MJ, Laidler FR, Madeley CR: **Growth and characterisation of human faecal astrovirus in a continuous cell line.** *Arch Virol* 1990, **113**:73–81.
46. Finkbeiner SR, Kirkwood CD, Wang D: **Complete genome sequence of a highly divergent astrovirus isolated from a child with acute diarrhea.** *Virol J* 2008, **5**:117.
47. Herrmann JE, Cubitt WD, Hudson RW, Perron-Henry DM, Oshiro LS, Blacklow NR: **Immunological characterization of the Marin County strain of astrovirus.** *Arch Virol* 1990, **110**:213–220.
48. Geigenmüller U, Ginzton NH, Matsui SM: **Construction of a genome-length cDNA clone for human astrovirus serotype 1 and synthesis of infectious RNA transcripts.** *J Virol* 1997, **71**:1713–1717.
49. Jiang B, Monroe SS, Koonin EV, Stine SE, Glass RI: **RNA sequence of astrovirus: distinctive genomic organization and a putative retrovirus-like ribosomal frameshifting signal that directs the viral replicase synthesis.** *Proc Natl Acad Sci U S A* 1993, **90**:10539–10543.
50. Marczinke B, Bloys AJ, Brown TD, Willcocks MM, Carter MJ, Brierley I: **The human astrovirus RNA-dependent RNA polymerase coding region is expressed by ribosomal frameshifting.** *J Virol* 1994, **68**:5588–5595.
51. Lewis TL, Matsui SM: **An astrovirus frameshift signal induces ribosomal frameshifting in vitro.** *Arch Virol* 1995, **140**:1127–1135.
52. Bass DM, Qiu S: **Proteolytic processing of the astrovirus capsid.** *J Virol* 2000, **74**:1810–1814.
53. Kiang D, Matsui SM: **Proteolytic processing of a human astrovirus nonstructural protein.** *J Gen Virol* 2002, **83**:25–34.
54. Geigenmüller U, Chew T, Ginzton N, Matsui SM: **Processing of nonstructural protein 1a of human astrovirus.** *J Virol* 2002, **76**:2003–2008.

55. Geigenmüller U, Ginzton NH, Matsui SM: **Studies on intracellular processing of the capsid protein of human astrovirus serotype 1 in infected cells.** *J Gen Virol* 2002, **83**:1691–1695.

56. Méndez E, Fernández-Luna T, López S, Méndez-Toss M, Arias CF: **Proteolytic processing of a serotype 8 human astrovirus ORF2 polyprotein.** *J Virol* 2002, **76**:7996–8002.

57. Méndez E, Salas-Ocampo E, Arias CF: **Caspases mediate processing of the capsid precursor and cell release of human astroviruses.** *J Virol* 2004, **78**:8601–8608.

58. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH: **CDD: a Conserved Domain Database for the functional annotation of proteins.** *Nucleic Acids Res* 2011, **39**:D225–229.

59. Jonassen CM, Jonassen TTØ, Sveen TM, Grinde B: **Complete genomic sequences of astroviruses from sheep and turkey: comparison with related viruses.** *Virus Res.* 2003, **91**:195–201.

60. Monroe SS, Stine SE, Gorelkin L, Herrmann JE, Blacklow NR, Glass RI: **Temporal synthesis of proteins and RNAs during human astrovirus infection of cultured cells.** *J Virol* 1991, **65**:641–648.

61. Koci MD, Seal BS, Schultz-Cherry S: **Molecular characterization of an avian astrovirus.** *J Virol* 2000, **74**:6173–6177.

62. Krishna NK: **Identification of structural domains involved in astrovirus capsid biology.** *Viral Immunol.* 2005, **18**:17–26.

63. Doane F: *Electron microscopy in diagnostic virology.* Cambridge [u.a.]: Cambridge University Press; 1987.

64. Scotto-Lavino E, Du G, Frohman MA: **5′ end cDNA amplification using classic RACE.** *Nat Protoc* 2006, **1**:2555–2562.

65. Scotto-Lavino E, Du G, Frohman MA: **3′ end cDNA amplification using classic RACE.** *Nat Protoc* 2006, **1**:2742–2745.

66. Kitano T, Umetsu K, Tian W, Osawa M: **Two universal primer sets for species identification among vertebrates.** *Int. J. Legal Med.* 2007, **121**:423–427.

67. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: **Clustal W and Clustal X version 2.0.** *Bioinformatics* 2007, **23**:2947–2948.

68. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol* 2003, **52**:696–704.

69. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389–3402.

70. Li W, Godzik A: **Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences.** *Bioinformatics* 2006, **22**:1658–1659.

71. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nature Methods* 2012, **9**:357–359.