



Short Communication

Complete genome sequence of canine astrovirus with molecular and epidemiological characterisation of UK strains

Sarah L. Caddy^{a,b,*}, Ian Goodfellow^{a,**}^a Division of Virology, Department of Pathology, University of Cambridge, Addenbrookes Hospital, Cambridge CB2 2QQ, UK^b Section of Virology, Faculty of Medicine, Imperial College London, St. Mary's Campus, London W2 1PG, UK

ARTICLE INFO

Article history:

Received 21 January 2015

Received in revised form 10 March 2015

Accepted 12 March 2015

Keywords:

Canine astrovirus

Gastroenteritis

Complete genome

ABSTRACT

Astroviruses are a common cause of gastroenteritis in children worldwide. These viruses can also cause infection in a range of domestic and wild animal species. Canine astrovirus (CaAstV) was first identified in the USA, and has since been reported in dogs from Europe, the Far East and South America. We sought to determine whether CaAstV is circulating in the UK dog population, and to characterise any identified strains. Stool samples were collected from pet dogs in the UK with and without gastroenteritis, and samples were screened for CaAstV by qPCR. Four CaAstV positive samples were identified from dogs with gastroenteritis (4/67, 6.0%), whereas no samples from healthy dogs were positive ($p < 0.001$). Sequencing of the capsid sequences from the four CaAstV strains found significant genetic heterogeneity, with only 80% amino acid identity between strains. The full genome sequence of two UK CaAstV strains was then determined, confirming that CaAstV conforms to the classic genome organisation of other astroviruses with ORF1a and ORF1b separated by a frameshift and ORF2 encoding the capsid protein. This is the first report describing the circulation of CaAstV in UK dogs with clinical signs of gastroenteritis, and the first description of the full-length genomes of two CaAstV strains.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Astroviruses are small non-enveloped, positive sense RNA viruses with a wide host range. Astroviruses are classified into two genera; the *Mamastrovirus* genera includes astroviruses isolated from humans, pigs, cattle, cats and dogs, whereas astrovirus isolates from birds are categorised into the *Avastrovirus* genera (King et al., 2011). Astroviruses were first identified in 1975 in the stools of children with diarrhoea (Appleton and Higgins, 1975;

Madeley and Cosgrove, 1975), and are now estimated to cause up to 10% of gastroenteritis cases in children worldwide (Moser and Schultz-Cherry, 2005).

Canine astrovirus (CaAstV) was first described in the USA, following the identification of star-shaped particles in diarrhoeic stools from a litter of beagles (Williams, 1980). Later studies have identified CaAstV in Italy, France, China, Korea and Brazil (Castro et al., 2013; Choi et al., 2014; Grellet et al., 2012; Martella et al., 2011; Zhu et al., 2011).

The genome of astroviruses is typically 6.4–7.3 kb and divided into three open reading frames (ORFs), ORF1a, ORF1b and ORF2 with a 5' untranslated region and a 3' poly-A tail (King et al., 2011). ORF1 codes for the non-structural proteins and ORF2 encodes the capsid precursor protein. The complete coding sequence of ORF2, and a partial sequence of ORF1b has previously been determined for a number of CaAstV isolates (Toffan et al., 2009; Zhu

* Corresponding author at: Division of Virology, Department of Pathology, University of Cambridge, Addenbrookes Hospital, Cambridge CB2 2QQ, UK.

** Corresponding author. Tel.: +44 1223 762653.

E-mail addresses: slc50@cam.ac.uk (S.L. Caddy), ig299@cam.ac.uk (I. Goodfellow).

et al., 2011), but to date, no full-length genome sequence of CaAstV has been reported.

The purpose of this study was to determine the prevalence of CaAstV in the UK dog population and to obtain where possible the complete genetic sequence of circulating strains. Four CaAstV strains were identified in dogs showing clinical signs of gastroenteritis and sequencing of ORF2 found significant genetic diversity between strains. The first full-length sequences of two CaAstV strains was subsequently determined.

2. Methods

2.1. Sample collection

Stool samples were collected from dogs admitted to five participating veterinary clinics and a single animal shelter in counties in the south and east of England; Cambridgeshire, Kent, Lincolnshire, Middlesex and Suffolk. Ethical approval was not required for this study as all samples collected were animal waste products. With owner consent, dogs were recruited to the study if they passed stools whilst hospitalised. Stool samples were collected from dogs from the animal shelter if they passed diarrhoea. All stool samples were stored at -20°C until and during transportation to the laboratory, where they were stored at -80°C prior to nucleic acid extraction. As controls, stool samples were also collected from healthy dogs owned by veterinary staff at each clinic, as well as from dogs at participating boarding kennels. Basic case data was recorded for each dog from which a stool sample was collected, including age, breed, sex, reason for admission, and any recent history of enteric disease.

2.2. Detection of CaAstV in clinical samples

Stools were diluted to a final concentration of 10% w/v in phosphate-buffered saline, pH 7.2, and solids were removed by centrifugation at $8000 \times g$ for 5 min. Viral nucleic acid was extracted from 140 μl of each clarified stool suspension using the GenEluteTM Mammalian Total RNA Miniprep Kit (Sigma-Aldrich) according to the manufacturers' instructions. An internal extraction control (equine arteritis virus RNA) was added during the nucleic acid extraction process as previously described (Caddy et al., 2013). cDNA was generated by reverse transcription using MMLV reverse transcriptase enzyme (Life Technologies) and random hexamers (Life Technologies) with the reaction performed at 42°C for 1 h, followed by an

inactivation step at 70°C for 10 min. qPCR was performed using primers targeting a conserved region of the viral RNA-dependent RNA polymerase as previously described (Martella et al., 2011). The sequences of primers used for PCR in this study are presented in Table 1. qPCR reactions were prepared using the MESA Blue qPCR MasterMix Plus for SYBR Assay (Eurogentech). Briefly, 2 μl cDNA was mixed with $2 \times$ MasterMix and 0.5 μM primers, then incubated at 95°C for 10 min. The thermal cycle protocol used with a ViiA7 qPCR machine (AB Applied Biosystems) was as follows: 40 cycles of 94°C , 15 s; 56°C , 30 s; 72°C , 30 s, followed by generation of a melt curve. Viral genome copy number was calculated by interpolation from a standard curve generated using serial dilutions of a standard DNA amplicon cloned from a positive control. The limit of detection was determined by the lowest dilution of control standard CNA reproducibly detected in the assay. All samples were additionally screened for the presence of canine parvovirus (CPV), canine enteric coronavirus (CECoV) and canine norovirus (CNV) using a 1-step qRT-PCR protocol (Caddy et al., 2013).

2.3. Capsid sequencing

All samples positive to CaAstV by qPCR were subjected to conventional PCR to confirm the presence of CaAstV and to enable genome sequencing. The capsid of all positive samples was amplified from cDNA synthesised using SuperScript II Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol with 0.5 μM AV12 primer (Table 1). The PCR reaction was performed using KOD hot start polymerase (EMD Millipore), with reverse primer s2m-rev, and the forward primer 625F-1 from the original qPCR assay. The amplification programme consisted of an initial 5 min step at 95°C , followed by 35 cycles with 95°C for 20 s, 58°C for 30 s and 72°C for 90 s. A final elongation step at 72°C for 5 min was performed, followed by chilling to 4°C .

PCR products were subsequently cloned into pCR-BluntTM using the Zero Blunt PCR Cloning Kit (Life Technologies) according to the manufacturers protocol. Sequencing of the 5' and 3' regions of the plasmid insert was performed using pCR-BluntTM specific primers by the University of Cambridge Biochemistry DNA Sequencing Facility. Sequencing primers for the central region of insert were then designed based on the primary sequence data to give a 200 nt overlap with each predecessor, and a second round of sequencing reactions was performed (sequencing primer details available on request). The complete capsid

Table 1
CaAstV specific primers used in this study.

Primer target		Sequence	Reference
RdRp	625F-1	GTACTATACCRCTGATTAAAT AGACCAARGTGTATAGTTTCTAG	Martella et al. (2011)
ORF1b-ORF2	AV12	TTT TTT TTT TTT TTT TTT GC	Toffan et al. (2009)
	s2m-rev	CCC TCG ATC CTA CTC GG	
ORF1a	683	TGAAGGACTGCTCAGAGTG	This paper
5' RACE	682	ATGCAGCGACAAACACAACA	This paper
3' RACE	894	CACAGCCCATTTGAAGATG	This paper

nucleotide sequence generated using this method was then confirmed by sequencing of PCR products generated from cDNA directly.

2.4. Full genome sequencing and phylogenetic analysis

PCR amplification of ORF1a and ORF1b was achieved using cDNA synthesised using SuperScript II Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol with 626R-1 reverse primer (Table 1). A forward primer was designed approximately 300 bp from the 5' end of ORF1a, based on a small genome fragment fortuitously amplified by mis-priming in a preliminary screen. The Zero Blunt PCR Cloning Kit was then used to clone the PCR product and sequencing performed using the protocol described above. Primers were designed sequentially to enable sequencing of the entire cloned product in both the 5'–3' and 3'–5' direction, with 200 nt overlaps between segments. Sequencing of PCR products directly was finally performed to ensure no mutations has arisen as a result of cloning. Each nucleotide was sequenced at least once in each direction, with an average coverage of 3.

Sequences at the extremity of the viral genome were determined using 5' and 3' RACE utilising a kit (Life Technologies) according to the manufacturers instructions, and gene specific primers as listed in Table 1. Direct sequencing was then performed as above.

Software used for sequence analysis included Vector NTI (Life Technologies), BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2>) and Protein Variability Server software (Garcia-Boronat et al., 2008). Evolutionary analysis was conducted using MEGA version 6 (Tamura et al., 2013).

3. Results

3.1. Sample collection

Stool samples and clinical data were collected from 67 dogs with severe gastroenteritis admitted to veterinary clinics or an animal shelter distributed across the UK between August 2012 and June 2014. Control samples were collected from 181 dogs without signs of gastroenteritis, from either veterinary inpatients with non-gastrointestinal illness, or dogs at boarding kennels or belonging to veterinary staff. In total 56 breeds of dog were represented. The mean age of dogs with gastroenteritis was 4.3 years (standard deviation 4.1 years) and the mean age of control animals was 6.1 years (standard deviation 3.9 years).

3.2. Identification of CaAstV isolates

Nucleic acid extraction and qPCR were performed on 248 stool samples. Samples were systematically tested for the presence of CaAstV using a SYBR-based qPCR screen. In addition, screening for CPV, CECov, CNV and the internal extraction control (EAV) was also performed using a 1-step qPCR protocol. CaAstV was detected in a total of four samples. All four positive dogs were showing signs of gastroenteritis, whereas CaAstV was not detected in any dogs without gastroenteritis. The difference between the prevalence of CaAstV in dogs with gastroenteritis (6.0%) and prevalence in dogs without gastroenteritis was statistically significant ($p = <0.001$). CPV was detected in 10/67 dogs with gastroenteritis, and CECov in 2/67 of the same group. Two of the four CaAstV positive dogs were also co-infected with CPV, but no co-infections with CECov and CaAstV were identified. The age range of CaAstV positive dogs was from 7 weeks to 7 years, with a mean age 2.1 years (SD 3.3 years). Table 2 summarises the clinical information and viral screening results of the four CaAstV positive dogs.

3.3. Comparison of CaAstV capsid sequences

The complete coding sequence of the four CaAstV strains identified was determined using conventional PCR and cloning techniques. The ORF2 nucleotide and amino acid sequences of these strains were aligned using ClustalW2. The overall nucleotide identity between strains was 77.1–81.1%, whereas the amino acid identity was 79.3–86.3%. The four sequences were deposited in the GenBank database and assigned accession numbers KP404149–KP404152 (cases 1–4, respectively).

A number of studies have reported the N-terminal and C-terminal regions of astrovirus capsids to be relatively conserved, whereas the central region is hypervariable. The human astrovirus capsid protein has previously been divided into three regions; the N terminus (amino acids 1–415), a variable central region (416–707), which includes a hypervariable section from 649 to 707, and a conserved C terminus (708–786) (Willcocks et al., 1995). An analogous approach for the CaAstV capsid was taken by Zhu et al. (2011), who divided the capsid into three regions for analysis: amino acids 1–446, 447–730, and 731–end. The four capsid sequences derived from this study have been analysed according to the latter scheme, and sequence identity compared (Fig. 1). Sequence analysis of the three regions clearly shows that the majority of sequence variation is concentrated in region II. A 24 nt

Table 2
Clinical and virological data for CaAstV positive cases.

Case	CaAstV strain	Age	Breed	Collection date	Additional viruses identified
1	Gillingham/2012/UK KP404149	7 weeks	Crossbreed	August 2012	CPV
2	Lincoln/2012/UK KP404150	7 years	Border Collie	August 2012	None
3	Huntingdon/2014/UK KP404151	1 year	Staffordshire Bull Terrier	February 2014	None
4	Braintree/2014/UK KP404152	10 weeks	Crossbreed	May 2014	CPV

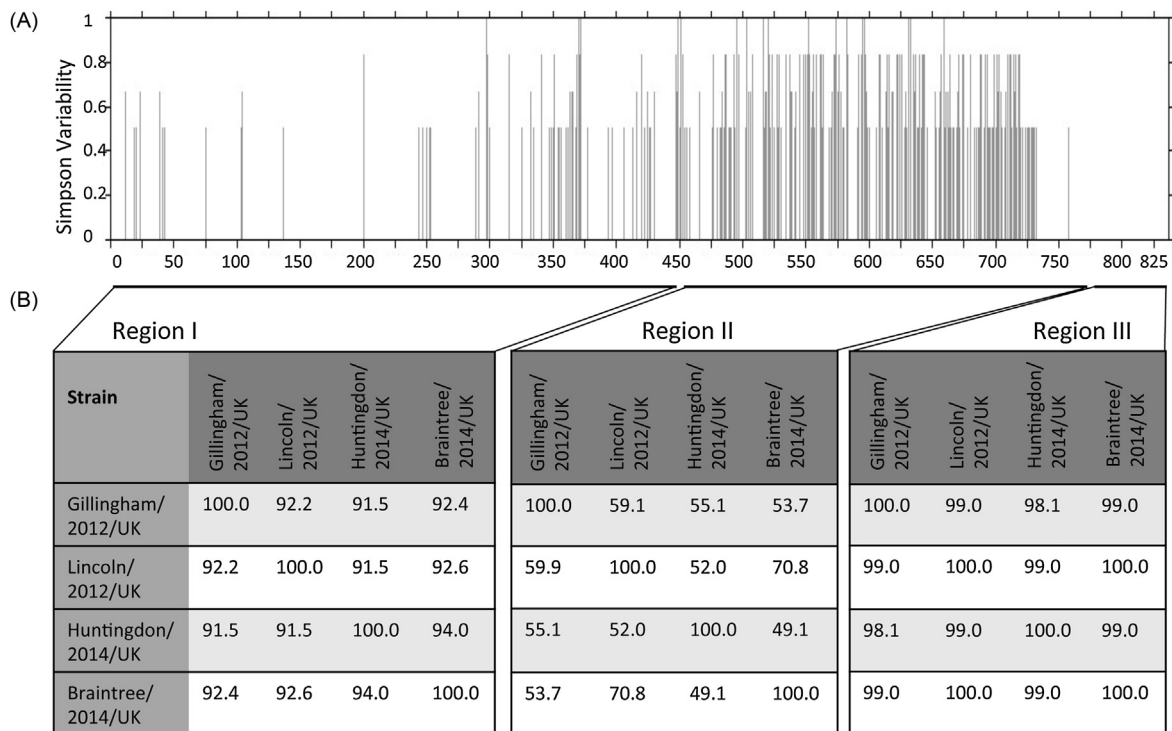


Fig. 1. Alignment and variability analysis of capsid amino acid sequences of the four CaAstV strains identified in the UK. (A) CaAstV capsid amino acid sequences were aligned using ClustalW2, then a variability scan was constructed using the Simpson diversity index and Protein Variability Server software (Garcia-Boronat et al., 2008). (B) Amino acid percentage identity between CaAstV strains in three different regions of the capsid protein.

deletion was identified in samples 2–4 in the 5' end region II, which has previously been reported in Chinese CaAstV strains (Zhu et al., 2011). It is not possible to predict the location of this deletion on the capsid structure as it is beyond the region of the astrovirus capsid spike for which the crystal structure has been solved (Dong et al., 2011). However, sequence alignment with the human astrovirus type 8 capsid (GenBank AAF85964.1) shows this deletion to be located downstream of the caspase cleavage site required for virion maturation, which truncates the full length capsid protein (VP90) into the mature VP70 form (Banos-Lara and Méndez, 2010). Therefore it is predicted that the 24 nt deletion will not alter the mature virion.

Evolutionary analysis of the four CaAstV sequences from the study, alongside the seven previously reported full-length CaAstV capsid sequences are presented in Fig. 2. This analysis indicated that the UK strains do not cluster, contrary to a previous study, which analysed CaAstV strains from a single city (Zhu et al., 2011). Each UK strain is distinct from each other, and whereas one strain clusters most closely with the Chinese strains, the remainder group with strains identified in Italy over a number of years.

3.4. Full-length CaAstV genome sequenced

The complete CaAstV genome was determined from two samples, isolated from a 7-week-old crossbreed dog (strain Gillingham/2012/UK, GenBank accession number

KP404149) and a 7-year-old Border Collie (strain Lincoln/2012/UK, KP404150). The total length of CaAstV Gillingham/2012/UK is 6600 nt and CaAstV/Lincoln/2012/UK is 6572 nt. Each genome encodes three ORFs; ORF1a, ORF1b, and ORF2 flanked by a 5' UTR, and a 3' UTR plus a poly-A tail (Fig. 3). In human astroviruses, the 5' UTR is 85 nt in length, whereas data from both CaAstV strains would indicate that the CaAstV 5' UTR is 45 nt. The 83 nt 3' UTR of human astrovirus is identical to 3' UTR in CaAstV Lincoln/2012/UK, whereas the 3' UTR of CaAstV Gillingham/2012/UK is 2 nt shorter. The nucleotide composition of both CaAstV strains is 29% A, 22% G, 26% T and 23% C. The G/C composition is 45%.

The ORF1a of non-canine astroviruses encodes a serine protease (Willcocks et al., 1994). Sequence alignment of CaAstV with astroviruses of other species shows a high degree of conservation in the predicted serine protease region (Fig. 3B). This is especially pronounced in the regions around the proposed catalytic triad of the serine protease.

ORF1a also encodes the viral genome-linked protein, VPg (Fuentes et al., 2012) (Fig. 4C). CaAstV VPg is predicted to start at aa 656, at a conserved QK cleavage site and is 90 aa in length. In other astroviruses, it has been proposed that the C-terminal VPg cleavage site is coded by Q(P/A/S/L) (Al-Mutairy et al., 2005), and the presence of a QS dipeptide at the same site in CaAstV is consistent with this prediction. The amino acid motif KGK(N/T)K is conserved at the N-terminal end of VPg sequences from both astroviruses and caliciviruses, and this is also

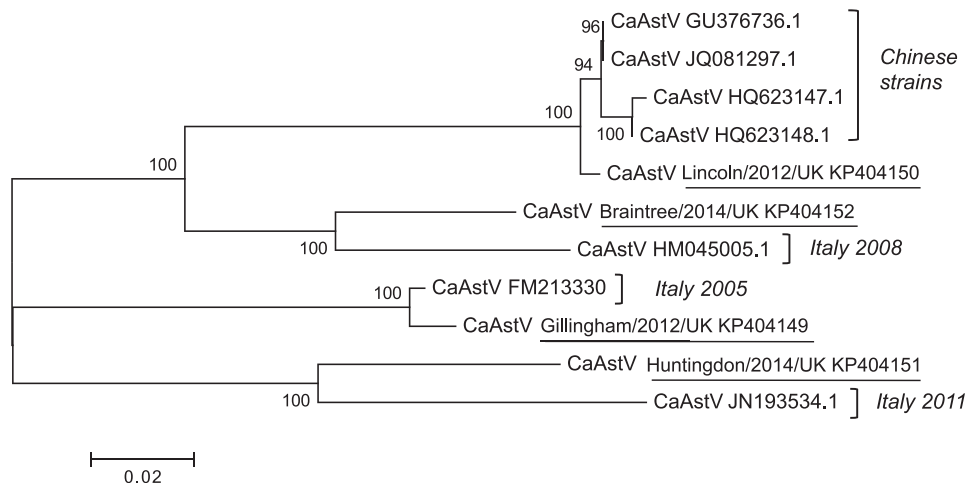


Fig. 2. Phylogenetic tree based on the full-length amino acid sequence of the capsid protein of CaAstVs. This includes the four UK strains identified in this study (underlined) and the seven strains previously sequenced and listed in GenBank. The tree was determined using the neighbour-joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (100 replicates) is shown next to the branches. Evolutionary analyses were conducted in MEGA6.

identifiable in the CaAstV genome. Another conserved VPg motif is TEXEY, with mutagenesis studies indicating that the Y (tyrosine) residue covalently links VPg to viral RNA (Fuentes et al., 2012). Analysis of the CaAstV ORF1a

sequence also identifies the conserved TEXEY motif at 684–688, thus this tyrosine is predicted to covalently link to the RNA genome. However, the CaAstV sequence diverges slightly from the other mamastroviruses studied,

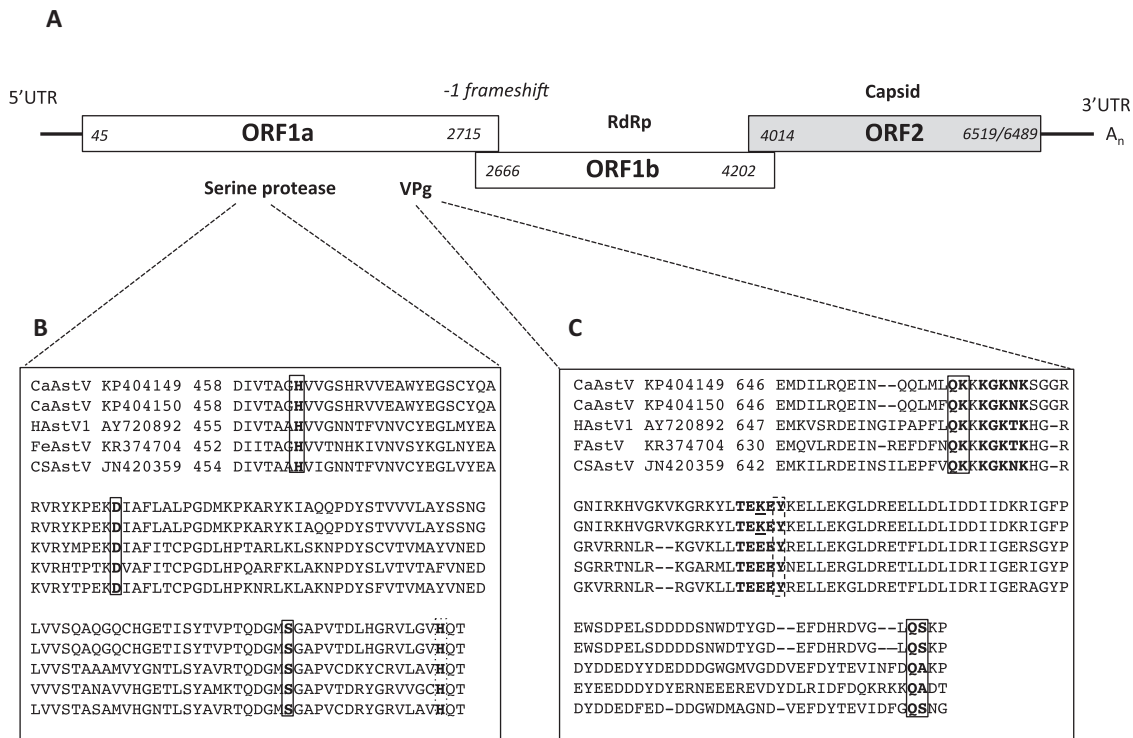


Fig. 3. Genome organisation of CaAstV. The genome is divided into three ORFs, with ORF1 divided by a frameshift site (A). The CaAstV of ORF1 is predicted to encode a serine protease (B) and a VPg protein (C). This is supported by alignment of the two CaAstV sequences with HAstV (human), FeAstV (feline) and CSLAstV (Californian sea lion) sequences. The suspected catalytic triad of the serine protease (B) is highlighted by the boxed residues, with the dotted box H residue (histidine) representing the substrate binding region. The boxed residues in (C) represent the C and N termini cleavage sites of VPg. The conserved KGK(N/T)K and TEXEY motifs of VPg proteins are highlighted in bold, with the TEXEY variation identified in CaAstV underlined. The tyrosine residues identified by a dashed box represent the site of covalent linkage of VPg to the RNA genome.

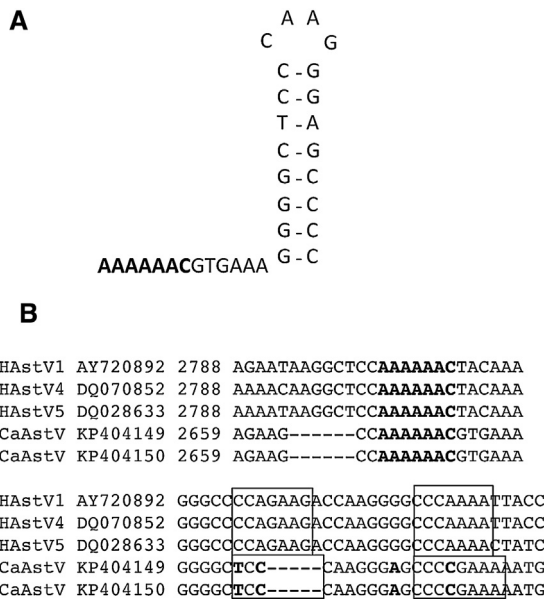


Fig. 4. Analysis of ORF1a/1b –1 frameshift site. (A) Predicted structure of the –1 frameshift site, directed by a slippery heptamer sequence and a downstream stem loop structure. (B) Sequence alignment of CaAstV and a range of human astrovirus (HuAstV) isolates shows this region to be highly conserved. The slippery heptamer sequence is highlighted in bold, and the stem loop structure is boxed, with nucleotides in bold representing differences.

in that X of the motif corresponds to K, whereas this is E/Q in all other mamastroviruses.

A –1 ribosomal frameshift site between ORF1a and ORF1b, present in human astroviruses (Marczinke et al., 1994), is also conserved in CaAstV. This translational frameshift is directed by the slippery heptamer sequence AAAAAAC at position 2666 in the CaAstV genome. A stem loop structure is predicted downstream of the slippery sequence, as shown in Fig. 4. The slippery sequence and downstream stem loop are highly conserved amongst mamastroviruses. The 3' end of ORF1a overlaps with ORF1b by 49 nucleotides. This is shorter than the 71 nt overlap reported for human astroviruses.

ORF1b is predicted to code for an RNA dependent RNA polymerase (RdRp). The CaAstV sequence contains a YGDD motif at aa 1252, common to RdRps of a variety of RNA viruses, supporting this conclusion. ORF1b of CaAstV aligns with the RdRp of human astroviruses with 58–60% aa identity (HAV-1, 4, 5, and 8). There is a similar identity to feline astrovirus (56%) and porcine astrovirus (55%). CaAstV aligns most closely with the Californian sealion astroviruses (67–71%), though incomplete Californian sea lion astrovirus sequences were available.

As for all other astroviruses studied to date, an overlapping reading frame exists at the ORF1b–ORF2 junction of CaAstV. The CaAstV ORF1b–ORF2 overlap sequence is 8 nt, as reported for other mamastroviruses, hence ORF2 is in the same frame as ORF1a. As previously reported, the capsid sequence was found to have an in-frame start codon 180 nt upstream of the start codon homologous to other mamastrovirus genomes (Toffan et al., 2009).

Table 3

Identity of the two full-length CaAstV sequences.

	Nucleotide (%)	Amino acid (%)
ORF1a	93.8	98.4
ORF1b	97.6	99.8
ORF2	77.5	80.8

The 6 aa C terminus of the VP1 (SRGHAE) is highly conserved with several mammalian AstVs. This motif is within a highly conserved nucleotide stretch, s2m, overlapping the termination codon of ORF2, and has been identified in both CaAstV strains sequenced.

The overall nucleotide identity between the two CaAstV strains sequenced was 88.5%. Sequence comparison of the individual ORFs is presented in Table 3. This clearly shows that ORF1b (RdRp) is most closely conserved, making this an ideal target for qPCR screens. Conversely, the capsid sequence is most diverse.

A phylogenetic tree was constructed by multiple alignment of the full-length genome of the two CaAstV strains isolated in this study, and a number of astrovirus reference strains isolated from different mammalian species. This was achieved using MEGA6 software (Fig. 5).

4. Discussion

CaAstV has previously been detected sporadically in dogs across the world, but the association with disease, prevalence levels and genetic diversity is largely unknown. This study presents the first identification and molecular characterisation of CaAstV cases in dogs in the UK. Sequencing of the viral capsid for all four strains revealed extensive genetic diversity and the first full-length sequences for CaAstV were determined for two strains.

The prevalence of CaAstV in gastroenteritis cases in this study was shown to be 6.0%. This prevalence was unexpectedly high given that CaAstV has not previously been reported in the UK. It may be predicted that this is an underestimation of prevalence based on the population of dogs surveyed. The majority of previous CaAstV epidemiological studies have focused on dogs less than 6 months old, whereas this study included dogs of any age. Serological studies have shown that exposure to CaAstV typically occurs in young animals, with dogs older than 3 months significantly more likely to be seropositive than younger dogs (Martella et al., 2011). This suggests that studies focusing only on young dogs will identify more positive CaAstV cases. However, the decision to survey dogs of any age in this study enabled detection of a CaAstV case in a 7-year-old dog. This is oldest case of CaAstV reported to date and highlights the need to have an index of suspicion for infectious causes of gastroenteritis in dogs of any age. Indeed although human astrovirus is more common in paediatric populations, infections in the elderly are reported (Fernández et al., 2011).

The pathology caused by CaAstV in dogs is uncertain as CaAstV has previously been detected in the stools of both healthy and diseased dogs (Grellet et al., 2012; Martella et al., 2011). However, our study shows a clear relationship between the presence of CaAstV RNA in stool samples, and

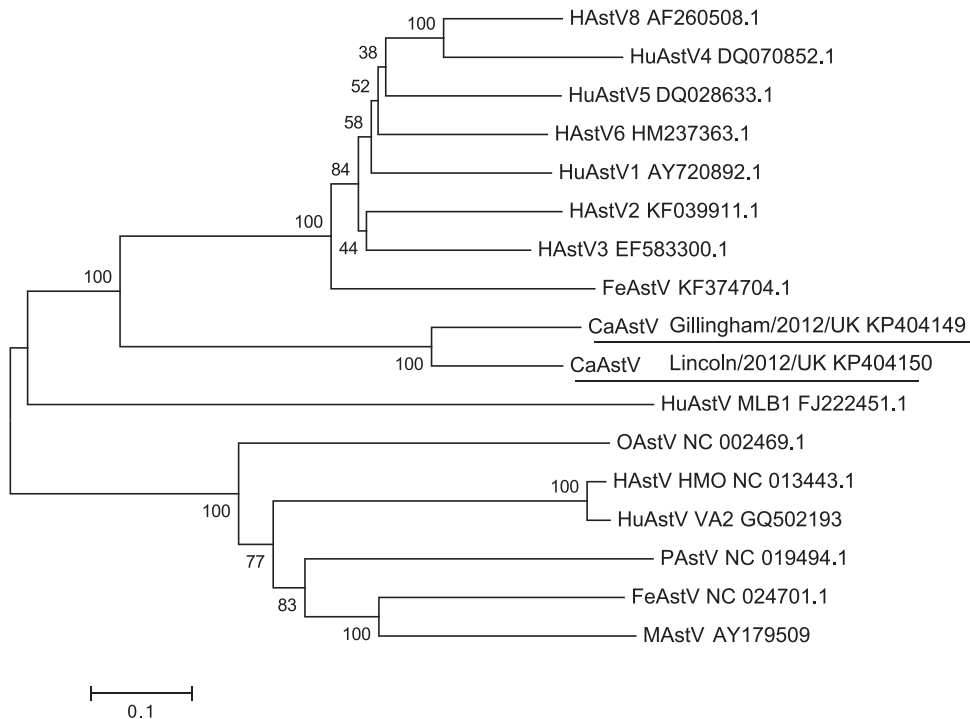


Fig. 5. Phylogenetic tree based complete genome nucleotide sequences of astroviruses from a range of mammalian species. GenBank accession numbers are listed for all sequences analysed, with the CaAstV sequences underlined. The tree was determined using the neighbour-joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (100 replicates) is shown next to the branches. Evolutionary analyses were conducted in MEGA6.

the presence of clinical signs of gastroenteritis ($p < 0.001$). This finding is in agreement with two previous studies from Italy and China (Martella et al., 2011; Zhu et al., 2011), but is at odds with a French study which found no significant difference in CaAstV identification between diarrhoeic (27%) or healthy puppies (19%) (Grellet et al., 2012). Determination of the specific pathology induced by CaAstV infection in dogs is often confounded in a clinical setting by co-infections with other gastroenteric pathogens (e.g. CPV in dogs 1 and 4 in this study). Experimental studies will be required to confirm or refute the association of CaAstV with gastroenteritis, but this study does suggest CaAstV can cause disease.

Sequencing of the capsid region of each CaAstV strain identified in this study revealed significant sequence variation. This mirrors the variation previously identified within human astrovirus isolates (De Benedictis et al., 2011). At present, astroviruses are named according to the species in which they are isolated, and subsequent classification is based upon serotypes; these are defined if there is a 20-fold or greater two-way cross neutralisation titre (King et al., 2011). Sequence analysis has verified this classification, with human astroviruses 1–8 having 86–100% nucleotide identity within a serotype, based on capsid sequences (Noel et al., 1995). The nucleotide variation within the capsid region of the four CaAstV strains was shown to be 77.1–81.1%, which strongly suggests these strains are also different serotypes. Confirmation of this requires serological analysis, but unfortunately repeated attempts to grow the CaAstV

isolates identified in this study in cell culture failed (data not shown).

Identification of four possible CaAstV serotypes circulating in the UK alone raises questions regarding the possible origins of these strains. Phylogenetic analysis of the UK capsid strains alongside the limited number of CaAstV sequences previously listed in GenBank was unexpected. There was no clustering of the UK strains, unlike the grouping of all Chinese strains. Instead UK strains each grouped with a different CaAstV isolate from either China or Italy. With such limited sequences available it is not possible to determine whether CaAstV strains have spread globally, or independent evolution has occurred. Clearly a high rate of evolution does occur within all astroviruses however, as their RNA genome facilitates introduction of point mutations and recombination events.

Given the strain diversity identified, it has been suggested that some CaAstV strains may be more pathogenic than others. This has previously been reported for astroviruses of mink, which show variation in their ability to invade the central nervous system in a strain related manner (Martella et al., 2012). Assessment of this risk will require wider epidemiological and clinical studies. Another concern raised by the existence of multiple circulating CaAstV strains is regarding future disease control. Management of viral causes of gastroenteritis in dogs is best achieved by widespread vaccination, exemplified by the widely used canine parvovirus vaccine. However, if CaAstV vaccine development is considered, the presence of multiple strains will make vaccine design challenging.

Full genome sequencing of two CaAstV isolates revealed them to be closely related and possess a typical astrovirus organisation. The first full length sequence of an astrovirus was for human astrovirus in 1994 (Willcocks et al., 1994), and relatively few full length sequences have since been determined. Sequence analysis of the CaAstV strains identifies the presence of a serine protease and VPg within ORF1a as for other astroviruses, which is separated from and conserved RdRp of ORF1b by a –1 frameshift.

In summary, this study has not only identified CaAstV circulating in the UK dog population, but also found significant genetic diversity within the CaAstV strains. Furthermore, full genome sequencing of two CaAstV isolates has enabled detailed molecular characterisation of this astrovirus species, and provides the astrovirus field with further examples of genome variation.

Acknowledgements

The authors would like to thank the veterinary teams at VetsNow Hospitals in Gillingham, Lincoln and Northolt, and Swayne and Partners, Bury St. Edmunds for assisting with stool sample collection. We would also like to thank Wood Green Animal Shelter, Abington Boarding Kennels and Greenlow Kennels in Cambridgeshire for samples. Further gratitude is expressed to James Warland at Cambridge Veterinary School for his work collecting samples and case management. This work was supported by a PhD studentship from the Medical Research Council to S.C. and a Wellcome Trust Senior Fellowship to I.G. (Ref: WT097997MA). I.G. is a Wellcome Senior Fellow.

References

- Al-Mutairy, B., Walter, J.E., Pothan, A., Mitchell, D.K., 2005. Genome prediction of putative genome-linked viral protein (VPg) of astroviruses. *Virus Genes* 31, 21–30, <http://dx.doi.org/10.1007/s11262-004-2196-1>.
- Appleton, H., Higgins, P.G., 1975. Letter: viruses and gastroenteritis in infants. *Lancet* 1, 1297.
- Banos-Lara, M.D.R., Méndez, E., 2010. Role of individual caspases induced by astrovirus on the processing of its structural protein and its release from the cell through a non-lytic mechanism. *Virology* 401, 322–332, <http://dx.doi.org/10.1016/j.virol.2010.02.028>.
- Caddy, S., Emmott, E., El-Attar, L., Mitchell, J., de Rougemont, A., Brownlie, J., Goodfellow, I., 2013. Serological evidence for multiple strains of canine norovirus in the UK dog population. *PLOS ONE* 8, e81596, <http://dx.doi.org/10.1371/journal.pone.0081596>.
- Castro, T.X., Cubel García, R.C.N., Costa, E.M., Leal, R.M., Xavier, M.D.P.T., Leite, J.P.G., 2013. Molecular characterisation of calicivirus and astrovirus in puppies with enteritis. *Vet. Rec.* 172, 557, <http://dx.doi.org/10.1136/vr.101566>.
- Choi, S., Lim, S., Kim, Y., Cho, Y., Song, J., An, D., 2014. Phylogenetic analysis of astrovirus and kobuvirus in Korean dogs. *J. Vet. Med. Sci.* 78, 1141–1145.
- De Benedictis, P., Schultz-Cherry, S., Burnham, A., Cattoli, G., 2011. Astrovirus infections in humans and animals – molecular biology, genetic diversity, and interspecies transmissions. *Infect. Genet. Evol.* 11, 1529–1544, <http://dx.doi.org/10.1016/j.meegid.2011.07.024>.
- Dong, J., Dong, L., Méndez, E., Tao, Y., 2011. Crystal structure of the human astrovirus capsid spike. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12681–12686, <http://dx.doi.org/10.1073/pnas.1104834108>.
- Fernández, J., de Oña, M., Melón, S., Alvarez-Argüelles, M.E., Boga, J.A., 2011. Noroviruses as cause of gastroenteritis in elderly patients. *Aging Clin. Exp. Res.* 23, 145–147, <http://dx.doi.org/10.3275/7040>.
- Fuentes, C., Bosch, A., Pintó, R.M., Guix, S., 2012. Identification of human astrovirus genome-linked protein (VPg) essential for virus infectivity. *J. Virol.* 86, 10070–10078, <http://dx.doi.org/10.1128/JVI.00797-12>.
- García-Boronat, M., Díez-Rivero, C.M., Reinherz, E.L., Reche, P.A., 2008. PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Res.* 36, W35–W41, <http://dx.doi.org/10.1093/nar/gkn211>.
- Grellet, A., De Battisti, C., Feugier, A., Pantile, M., Marciano, S., Grandjean, D., Cattoli, G., 2012. Prevalence and risk factors of astrovirus infection in puppies from French breeding kennels. *Vet. Microbiol.* 157, 214–219, <http://dx.doi.org/10.1016/j.vetmic.2011.11.012>.
- King, A.M., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., 2011. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. Elsevier, London, UK.
- Madeley, C.R., Cosgrove, B.P., 1975. Letter: 28 nm particles in faeces in infantile gastroenteritis. *Lancet* 2, 451–452.
- Marczinke, B., Bloys, A., Brown, T., Willcocks, M., Carter, M.J., Brierley, I., 1994. The human astrovirus RNA-dependent RNA polymerase coding region is expressed by ribosomal frameshifting. *J. Virol.* 68, 5588–5595.
- Martella, V., Moschidou, P., Lorusso, E., Mari, V., Camero, M., Bellacicco, A., Losurdo, M., Pinto, P., Desario, C., Bányai, K., Elia, G., Decaro, N., Buonavoglia, C., 2011. Detection and characterization of canine astroviruses. *J. Gen. Virol.* 92, 1880–1887, <http://dx.doi.org/10.1099/vir.0.029025-0>.
- Martella, V., Moschidou, P., Catella, C., Larocca, V., Pinto, P., Losurdo, M., Corrente, M., Lorusso, E., Bányai, K., Decaro, N., Lavazza, A., Buonavoglia, C., 2012. Enteric disease in dogs naturally infected by a novel canine astrovirus. *J. Clin. Microbiol.* 50, 1066–1069, <http://dx.doi.org/10.1128/JCM.05018-11>.
- Moser, L.A., Schultz-Cherry, S., 2005. Review pathogenesis of astrovirus infection. *Viral Immunol.* 18, 4–10.
- Noel, J.S., Lee, T.W., Kurtz, J.B., Glass, R.I., Monroe, S.S., Noel, J.S., Lee, T.W., Kurtz, J.B., Glass, R.I., Monroe, S.S., 1995. Typing of human astroviruses from clinical isolates by enzyme immunoassay and nucleotide sequencing. *J. Clin. Microbiol.* 33, 797–801.
- Tamura, K., Stecher, G., Peterson, D., Filipinski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729, <http://dx.doi.org/10.1093/molbev/mst197>.
- Toffan, A., Jonassen, C.M., De Battisti, C., Schiavon, E., Kofstad, T., Capua, I., Cattoli, G., 2009. Genetic characterization of a new astrovirus detected in dogs suffering from diarrhoea. *Vet. Microbiol.* 139, 147–152, <http://dx.doi.org/10.1016/j.vetmic.2009.04.031>.
- Willcocks, M.M., Brown, T.D., Madeley, C.R., Carter, M.J., 1994. The complete sequence of a human astrovirus. *J. Gen. Virol.* 75, 1785–1788.
- Willcocks, M.M., Kurtz, J.B., Lee, T.W., Carter, M.J., 1995. Prevalence of human astrovirus serotype 4: capsid protein sequence and comparison with other strains. *Epidemiol. Infect.* 114, 385–391.
- Williams, F., 1980. Astrovirus-like, coronavirus-like, and parvovirus-like particles detected in the diarrheal stools of beagle pup. *Arch. Virol.* 66, 215–226.
- Zhu, L.A., Zhao, W., Yin, H., Shan, T.L., Zhu, C.X., Yang, X., Hua, X.G., Cui, L., 2011. Isolation and characterization of canine astrovirus in China. *Arch. Virol.* 156, 1671–1675, <http://dx.doi.org/10.1007/s00705-011-1022-z>.