1 **Genome analysis of canine astroviruses reveals genetic heterogeneity and**

2 **suggests possible inter-species transmission**

3

4 *Eszter Mihalov-Kovács[a], Vito Martella[b], Gianvito Lanave[b], Livia Bodnar[b], Enikő Fehér[a], Szilvia*

5 *Marton[a], Gábor Kemenesi[c], Ferenc Jakab[c], Krisztián Bányai[a*]*

6

7 [a] Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian

8 Academy of Sciences, Hungária krt. 21,Budapest, H-1143, Hungary

9 [b] Dipartimento di Medicina Veterinaria, Universitá Aldo Moro di Bari, S.p. per Casamassima

10 km 3, 70010 Valenzano, Italy

11 [c] Virological Research Group, Faculty of Sciences, Institute of Biology, University of Pécs,

12 Ifjúság útja 20, H-7624 Pécs, Hungary

13

14 * Corresponding author: Krisztián Bányai, Institute for Veterinary Medical Research,

15 Hungária krt. 21, Budapest, H-1143, Hungary; Phone: +361 467 4060, Fax: +361 467 4067;

16 Email: bkrota@hotmail.com

17

18 **Abstract**

19 Canine astrovirus RNA was detected in the stools of 17/63 (26.9%) samples, using either a
20 broadly reactive consensus RT-PCR for astroviruses or random RT-PCR coupled with massive
21 deep sequencing. The complete or nearly complete genome sequence of five canine astroviruses
22 was reconstructed that allowed mapping the genome organization and to investigate the genetic
23 diversity of these viruses. The genome was about 6.6 kb in length and contained three open
24 reading frames (ORFs) flanked by a 5' UTR, and a 3' UTR plus a poly-A tail. ORF1a and
25 ORF1b overlapped by 43 nucleotides while the ORF2 overlapped by 8 nucleotides with the 3'
26 end of ORF1b. Upon genome comparison, four strains (HUN/2012/2, HUN/2012/6,
27 HUN/2012/115, and HUN/2012/135) were more related genetically to each other and to UK
28 canine astroviruses (88-96% nt identity), whilst strain HUN/2012/126 was more divergent (75-
29 76% nt identity). In the ORF1b and ORF2, strains HUN/2012/2, HUN/2012/6, and
30 HUN/2012/135 were related genetically to other canine astroviruses identified formerly in
31 Europe and China, whereas strain HUN/2012/126 was related genetically to a divergent canine
32 astrovirus strain, ITA/2010/Zoid. For one canine astrovirus, HUN/2012/8, only a 3.2 kb portion
33 of the genome, at the 3' end, could be determined. Interestingly, this strain possessed unique
34 genetic signatures (including a longer ORF1b/ORF2 overlap and a longer 3'UTR) and it was
35 divergent in both ORF1b and ORF2 from all other canine astroviruses, with the highest
36 nucleotide sequence identity (68% and 63%, respectively) to a mink astrovirus, thus suggesting a
37 possible event of interspecies transmission. The genetic heterogeneity of canine astroviruses may
38 pose a challenge for the diagnostics and for future prophylaxis strategies.

39 **Keywords** *Astroviridae*; dog; semiconductor sequencing; Hungary

40

1. Introduction

Astroviruses (AstV), family *Astroviridae,* are non-enveloped viruses with a diameter of 28-30 nm and with a typical star-like shape. The genome is a single strand of positive sense, RNA of 6.4-7.3 kb in size, containing three overlapping ORFs (ORF1a, ORF1b and ORF2) with a 3' poly(A) tail (Mendez and Arias, 2007). ORF1a encodes a serine 3C type of viral protease. ORF1b is separated from ORF1a by a heptameric frame shift signal (AAAAAAC) and encodes the viral RNA polymerase. ORF2 encodes an 87-kDa polypeptide, which functions as the capsid precursor. AstV infection is associated with gastroenteritis in many animal species and humans, and they are also associated with extra-intestinal diseases, such as nephritis in chickens, hepatitis in ducks and shaking syndrome in minks (Imada et al., 2000; Fu et al., 2009; Blomström et al., 2010). The evolution of AstVs is driven by mechanisms of genetic drift, recombination and, possibly, inter-species transmission (Finkbeiner et al., 2008; De Benedictis et al., 2011; Martella et al., 2014).

So far, AstVs are classified into two distinct genera, *Mamastrovirus* (MAstV) and *Avastrovirus* (AvAstV) with 19 MAstV (mammalian) species and 3 AvAstV (avian) species listed officially by the International Committee on Taxonomy of Viruses (ICTV) (Bosch et al., 2010). However, taking advantage of broad-range PCR primers for AstVs and metagenomic protocols, several novel AstVs have been identified in a number of mammalian and avian species (Chu et al., 2008; Finkbeiner et al., 2008; Bosch et al., 2010; De Benedictis et al., 2011). Using the classification criteria adopted in the 9th ICTV report that is based on genetic analysis of the full-length ORF2, an additional 14 MAstV and 4 AvAstV candidate species have been defined recently (Schultz-Cherry, 2013).

AstV-like particles have been detected only occasionally in dogs by EM, either alone or in co-infection with other enteric viruses (Williams, 1980; Marshall et al., 1984; Vieler and Herbst,

1995; Toffan et al., 2009). More recently, AstVs have been identified in dogs with enteric signs and characterized molecularly, suggesting that the detected viruses may represent a distinct AstV species (Toffan et al., 2009). Also, a canine AstV, strain ITA/08/Bari, was successfully adapted to replicate *in vitro* on MDCK cells, and AstV-specific antibodies were detected in convalescent canine sera (Martella et al., 2011, 2012). The prevalence of AstV infection seems higher in dogs with enteric disease than in asymptomatic animals (Martella et al., 2011; Zhu et al., 2011; Caddy and Goodfellow, 2015; Takano et al., 2015). Also, monitoring of natural infection by AstV in dogs has revealed that the acute phase of gastroenteritis overlaps with peaks of viral shedding (Martella et al., 2012). At present, limited and partial information on canine AstV genomes is available in GenBank. This limited amount of information seems to suggest that canine AstVs are genetically heterogeneous (Martella et al., 2011; Caddy and Goodfellow, 2015), thus posing a challenge for the diagnostic and for the understanding of the genetic and biological properties of these viruses in dogs.

In this study, the complete or nearly complete genome sequence of five canine AstVs and the partial genome sequence of one canine AstV were determined and analysed, providing information on the genome organization and genetic diversity of these viruses.

2. Materials and methods

2.1. Samples

During 2012 samples were collected from diarrheic and non-diarrheic dogs from a Hungarian shelter. There was no age restriction. A total of 63 samples obtained from 50 animals were tested for AstV by using a pan-astrovirus specific primer set (Chu et al, 2008) as described elsewhere (Mihalov-Kovács et al., 2014) and 37 (from 33 dogs) were randomly selected for viral metagenomics.

89 Fecal samples were diluted 1: 10 in PBS (phosphate buffered saline) and homogenized.
90 The homogenates were centrifuged at 10000×g for 5 min and the supernatant was collected for
91 extraction of viral RNA (Zymo DirectZol viral RNA extraction kit, Zymo Research).

92

93 *2.2. Semiconductor sequencing*

94 Templates for deep sequencing were prepared as described previously (Mihalov-Kovács
95 et al., 2015). In brief, viral RNA samples were denatured at 97°C for 5 min in the presence of 10
96 μM random hexamer tailed by a common PCR primer sequence (Djikeng et al., 2008). Reverse
97 transcription was performed with 1 U AMV reverse transcriptase (Promega, Madison, WI, USA),
98 400 μM dNTP mixture, and 1× AMV RT buffer at 42°C for 45 min following a 5 min incubation
99 at room temperature. Then, 5 μL cDNA was added to 45 μL PCR mixture to obtain a final
100 volume of 50 μL and a concentration of 500 μM for the PCR primer, 200 μM for dNTP mixture,
101 1.5 mM for MgCl2, 1× Taq DNA polymerase buffer, and 0.5 U for Taq DNA polymerase
102 (Thermo Scientific, Vilnius, Lithuania). The reaction conditions consisted of an initial
103 denaturation step at 95°C for 3 min, followed by 40 cycles of amplification (95°C for 30 sec,
104 48°C for 30 sec, 72°C for 2 min) and terminated at 72°C for 8 min.

105 We subjected 0.1 μg of random PCR product to enzymatic fragmentation and adaptor
106 ligation (NEBNext Fast DNA Fragmentation & Library Prep Set for Ion Torrent kit, New
107 England Biolabs, Ipswich, MA, USA). The barcoded adaptors were retrieved from the Ion Xpress
108 Barcode Adapters (Life Technologies, Carlsbad, CA, USA). The resulting cDNA libraries were
109 measured on an Qubit 2.0 device using the Qubit dsDNA BR Assay kit (Invitrogen, Eugene, OR,
110 USA). The emulsion PCR that produced clonally amplified libraries was carried out according to
111 the manufacturer's protocol using the Ion PGM Template kit on an OneTouch v2 instrument.
112 Enrichment of the templated beads (on an Ion One Touch ES machine) and further steps of

113 presequencing setup were performed according to the Page 3 of 11 200-bp protocol of the
114 manufacturer. The sequencing protocol recommended for Ion PGM Sequencing Kit on an 316
115 chip was strictly followed.

116

117 *2.3. 5' rapid amplification of cDNA ends (5' RACE)*

118 To determine the actual 5' end genomic sequence of two AstV strains, a 5' RACE
119 protocol was used, following the manufacturer's instructions (Invitrogen Ltd). The amplicons
120 were visualized on 1.5% agarose gel, and cleaned up with Qiagen Qiaquick Gel Extraction Kit,
121 according to the manufacturer's instructions. The amplicons were cloned using TOPO XL PCR
122 Cloning (Invitrogen Ltd) and the clones were subjected to sequencing in both directions using
123 Big Dye v3.1 chemistry on a 3730xl instrument from Applied Biosystems (Foster, CA). For
124 sequencing accuracy, a minimum of three independent clones for each fragment type were
125 selected for sequencing in both directions using the universal M13F/R primers.

126

127 *2.4. 3' rapid amplification of cDNA ends (3' RACE)*

128 The First Strand III kit (Invitrogen Ltd) was used to generate cDNA from the poly-A
129 tailed RNA target, as described previously (Martella et al., 2011). Forward primers were designed
130 about 1 kb upstream of the 3' terminus of the genome of the various AstV strains and used with a
131 reverse poly-T oligonucleotide with a 5'-anchored tail. The 3' RACE products were sequenced
132 on Ion Torrent PGM.

133

134 *2.5. Sequence and phylogenetic analysis*

135 Sequence data generated by the Ion Torrent PGM were trimmed and analyzed by the CLC
136 Bio software (www.clcbio.com). The same software package was utilized to map sequence reads

137 to reference eucaryotic viral sequences retrieved from GenBank. Sequence editing, annotation
138 and analysis were carried out using Geneious software v9.1.4 (Biomatters LTD, New Zealand).
139 Open reading frames (ORF) were predicted with the ORF finder. Multiple alignments were
140 prepared and manually adjusted, whereas phylogenetic analysis was performed by using the
141 MEGA6 software (Tamura et al., 2013) with a neighbor-joining method, Jukes Cantor genetic
142 distance model and bootstrapping over 1000 replicates. Mean amino acid genetic distances were
143 computed using the p-distance method of MEGA6. The protein sequence analysis and
144 classification were prepared with the aid of EMBL-EBI website (http://www.ebi.ac.uk/interpro/).
145 Transmembrane domains were determined using Phobius web server (Käll et al., 2007). Nuclear
146 localization signals (NLS) were predicted by
147 http://www.moseslab.csb.utoronto.ca/NLStradamus/.
148

149 *2.6. GenBank accession numbers*

150 AstV genomic sequences were deposited in the GenBank with the following accession
151 numbers: HUN/2012/2, KX599349; HUN/2012/6, KX599350; HUN/2012/115, KX599351;
152 HUN/2012/126, KX599352; HUN/2012/135, KX599353; HUN/2012/8, KX599354.
153

154 **3. Results**

155 *3.1. Detection of AstV sequences in the canine fecal virome*

156 During 2012, 63 stool specimens were screened for AstVs using broad-range AstV-
157 specific primer set. Of these, seven tested positive; 4 out of 25 diarrheic animals (16%) and 3 out
158 of 38 non-diarrheic animals (8%). The difference between groups was not statistically significant
159 (chi-square value, 1.003; p=0.316). A subset (n=37) of samples was selected from 33 animals for
160 viral metagenomics to detect virus diversity in canine fecal specimens. Among these 37 samples

161 26 specimens contained eukaryotic origin viral sequences (Figure 1). By viral metagenomics
162 (Table 1), AstV RNA was initially identified in 11 out of 37 fecal specimens (29.7%). However,
163 the proportion of AstV specific reads ranged from 1 out of 18711 (<0.1%) to 1627 out of 1635
164 (99.5%) eukaryotic viral sequence reads, respectively, and five samples that had very low AstV
165 sequence read numbers (range, 1 to 4) were not considered for further processing. Other enteric
166 viruses detected in the stools included parvovirus (in 20 samples), coronavirus (in 14 samples),
167 rotavirus (in 11 samples), picornavirus (in 7 samples), gyrovirus (in 1 sample), adenovirus (in 1
168 sample) and calicivirus (in 1 sample). Overall, by either the pan-astrovirus RT-PCR or by the
169 viral metagenomic screening, we identified AstV RNA in 17/63 (26.9%) samples. Interestingly,
170 some samples negative in RT-PCR generated AstV-specific reads by deep sequencing, whilst, in
171 turn, some samples containing AstV-specific reads tested negative by the pan-astrovirus RT-
172 PCR, indicating a weak correlation between the two methodologies (Table 1).

173

174 *3.2. General features of the canine AstV genomes*

175 For further analysis the complete, coding complete and partial genome sequence was
176 determined for four, one and one AstV strains, respectively. In general, the length of the coding
177 sequences ranged between 6441-6474 nucleotide (nt), with a GC content of 45.19-45.48%. The
178 total genome length, considering the 5' and 3' UTRs, ranged between 6535 and 6587 nt. The 5'
179 UTR was sequenced completely for strains HUN/2012/2, HUN/2012/115, HUN/2012/126 and
180 HUN/2012/135, and ranged between 42 and 60 nt in length. The 3' UTR was sequenced
181 completely for all the six strains and ranged between 49 and 100 nt in length.

182 The genome of strain HUN/2012/115 was 6569 nt long, excluding the polyA-tail. The 5'
183 UTR and 3' UTR were, respectively, 45 nt and 83 nt long. The genome of strain HUN/2012/135
184 was 6587 nt long and it was organized similarly to strain HUN/2012/115, even if the 3' UTR was

185 2 nt longer. The genome of strain HUN/2012/2 was 6576 nt long and it differed for the longer 5' UTR region (60 nt) from the other canine AstVs. Despite several attempts, for strain HUN/2012/6, it was not possible to generate the actual 5' UTR sequence. The genome of this strain was organized in an identical manner to the canine AstV strain GBR/2014/Gillingham. The genome length of the AstV strain HUN/2012/126 was 6535 nt long, excluding the polyA-tail. The 5' UTR was 42 nt in length, with an additional in-frame ATG codon located 18 nt upstream (bases 25-27) of the ATG codon of ORF1. The 3' UTR was 49 nt long. Only the partial genome (3187 nt in length) of strain HUN/2012/8 was determined. The virus possessed an early stop codon in the ORF2, resulting in a longer (100 nt) 3' UTR (Table 2).

194 The coding sequences showed the typical organization of AstV genome. There were three overlapping ORFs, coding the non-structural and structural proteins. The predicted genome organization of the viruses is shown in Figure 2 and summarized in Table 2.

197

198 *3.3. Analysis of ORF1a*

199 The complete ORF1a was 2670 nt long for all canine AstVs, except strain HUN/2012/126, and codes for a 889 aa long polyprotein. For strain HUN/2012/126, an in-frame start codon is located 18 nt upstream of the ORF1 start codon (the length of the ORF is 2667/2685 depending on the start). A ribosomal frameshift signal, the heptameric (slippery) AAAAAAC sequence, was recognized at nt 2622-2628 (2619/2637 in HUN/2012/126 depending on the start used), with the ORF1b initiating after the slippery sequence. The predicted 95 kDa nsp1a protein, encoded by ORF1a, contained conserved domains typical for the AstV serine protease between aa 439 and 583 aa. Five transmembrane domains were identified on the predicted ORF1a protein at residues aa151-168, aa240-266, aa287-313, aa319-339 and aa351-

380. A possible coiled coil structure was also found at residues aa595-616 and aa635-656 (Figure 2).

Upon sequence comparison, strains HUN/2012/2, HUN/2012/6, HUN/2012/115 and HUN/2012/135 displayed 96-99% nt identity to each other and 93-98% nt to the canine strains UK/2014/Lincoln and UK/2014/Gillingham, whilst strain HUN/2012/126 displayed only 74% nt identity to the other canine AstVs. Identity to non-canine AstVs was markedly lower (33-51% nt) (Table 3).

3.4. Analysis of ORF1b

The complete ORF1b was 1530 nt long in all canine AstVs, except the partially sequenced strain, HUN/2012/8, where a 770 nt long portion was determined. The ORF1b begins with the last nucleotide of the slippery heptamer and with a -1 frameshift with respect to ORF1a. Also, the downstream stem loop structure was highly conserved among canine AstVs (Caddy and Goodfellow, 2015). ORF1b codes for a 509 aa long non-structural protein, the RNA-dependent RNA polymerase (RdRp). The RdRp domain was recognized between aa168-373, with conserved motifs at positions aa321 (GNPSG), aa364 (YGDD), and aa391 (FGMWVK).

Sequence comparison in the ORF1b (RdRp) was calculated using a 415 nt (138 aa) long fragment located at the very 3' end of the gene. Upon sequence comparison, strains HUN/2012/2, HUN/2012/6, HUN/2012/115 and HUN/2012/135 displayed 98-99% nt (99% aa) identity to each other and 96-99% nt (99% aa) to most canine strains. Strain HUN/2012/126 displayed the highest identity to strain ITA/2010/Zoid (99% nt and 99% aa) and 85-86% nt (87-88% aa) to the other canine AstVs. Strain HUN/2012/8 displayed the highest identity to a mink AstV (68% nt and 70% aa), whilst identity to other canine AstVs was lower (56-58% nt and 50-52% aa) (Table 3**).**

232 ### 3.5. *Analysis of ORF2*

233 The length of ORF2 is variable among the canine AstV strains. In all but one strains
234 analysed in this study, the ORF1b and ORF2 overlapped by 8 nt. In strain HUN/2012/8 the
235 overlapping was 11 nt in length. In addition, there was a second in-frame ATG start codon,
236 located 180 nt upstream of the start codon homologous to other *Mamastrovirus* genomes. Only in
237 strain HUN/2012/8, this additional in-frame start codon was located 141 nt upstream of the first
238 start codon. During the RNA replication, ORF2 is expressed from a sgRNA (Walter and Mitchell,
239 2003; Mendez et al., 2014). The sgRNA promoter sequence has a highly conserved nucleotide
240 sequence motif and is present upstream of the ORF2 start codon (Mendez et al., 2014). The
241 highly conserved nucleotide stretch upstream of ORF2,
242 ATATGGAGGGGAGGACCAAAAATTGCG**ATG**GC, believed to be part of a promoter
243 region for synthesis of sgRNA, was completely conserved in the sequence of canine AstV strains
244 except for the sequence of strain HUN/2012/8, that was
245 CTTTGGAGGGGAGGACCAAAGCAGCGCT**ATG**GC.

246 The ORF2 was 2292/2472 nt long for strain HUN/2012/115 and HUN/2012/135,
247 2295/2475 nt long for strain HUN/2012/2, 2325/2505 nt long for strain HUN/2012/6, 2298/2478
248 nt for strain HUN/2012/126 and 2328/2469 nt for HUN/2012/8 (Table 2). The terminal nt stretch
249 of ORF2 and the adjacent 3'UTR (s2m) is highly conserved among most AstVs and similarities
250 have been observed in the sequence and folding of the 3'UTR of viruses from other viral
251 families, suggesting that it is relevant for AstV genome replication (Monceyron et al., 1997). The
252 conserved s2m sequence CCGCGGCCACGCCGAG**TAG**GAACGAGGGTACAG overlaps the 6
253 aa C terminus of the capsid (SRGHAE), that is highly conserved in several mammalian AstVs,
254 including the canine AstVs (Figure 2). However, in strain HUN/2012/8 the highly conserved s2m

255 stretch was found downstream of the ORF2 termination codon, in the 3'UTR. As a consequence,
256 the 6 aa C terminus of strain HUN/2012/8 was TLSTKH.

257 Upon sequence comparison, the Hungarian canine AstVs were found to differ markedly.
258 Strains HUN/2012/2 and HUN/2012/135 were highly related to each other (92% nt and 96% aa
259 identity) and to Chinese CAstVs (91-96% nt and 95-99 % aa identity) as well as to the UK strain
260 GBR/2014/Lincoln (91-95% nt and 95-99% aa). Strain HUN/2012/6 displayed the highest
261 identity (91-95% nt and 96-98% aa) to three Italian strains (ITA/2005-3, ITA/2005-6, ITA/2005-
262 8) and to the UK strain GBR/2014/Gillingham. Strain HUN/2012/126 displayed the highest
263 identity (97% nt and 98% aa) to the Italian strain ITA/2010/Zoid. Strain HUN/2012/115
264 displayed the highest identity (79% nt and 79% aa) to the Italian strain ITA/2010/Zoid, while
265 sequence identity to other canine AstVs ranged from 72 to 75% at the nt level and 75 to 80% at
266 the aa level. Interestingly, strain HUN/2012/8 displayed the highest identity (63% nt and 64% aa)
267 to a mink AstV (Table 3).

268

269 *3.6. Phylogenetic analysis of canine AstVs*

270 Phylogenetic trees were constructed using the nucleotide alignments of ORF1a, ORF1b
271 and ORF2 (Figure 3). The ORF1a and ORF2 alignments were based on the full-length ORF
272 sequences and also included a selection of sequences of mammalian AstV strains. The ORF1b
273 alignment included 415 nt at the 3' end (138 aa at the C terminus of the protein).

274 In all phylogenetic trees, the majority of canine AstVs formed a distinct monophyletic
275 group. In the ORF1a-based tree (Figure 3A), only five canine AstV strains generated in this study
276 and two AstV strains from UK could be included, as information on the ORF1a is not available
277 for other canine AstVs. Although monophyletic, the canine AstVs segregated in two distinct
278 branches, with strain HUN/2012/126 in the basal position. When enlarging the number of

279 analysed sequences, in the ORF1b (Figure 3B), this pattern of segregation was confirmed. Four
280 strains (HUN/2012/2, HUN/2012/6, HUN/2012/115, and HUN/2012/135) were grouped with the
281 majority of canine AstVs identified in European and extra-European countries, whilst strain
282 HUN/2012/126 clustered with the highly virulent strain ITA/2010/Zoid. Interestingly, strain
283 HUN/2012/8 was distantly related to all other canine AstVs and it was grouped with AstVs
284 identified in minks and in Californian sea lions. The unique evolutionary relationship of this
285 canine strain was confirmed in the ORF2 (capsid)-based tree (Figure 3C). This gene is considered
286 as the fundamental target for classification and characterization of AstVs. When observing in
287 detail the large monophyletic group of canine AstVs which classified into species 5 (Schultz-
288 Cherry, 2013), it was possible to identify clearly sub-clustering patterns, with at least 6 sub-
289 clusters. All Chinese strains, the UK strain GBR/2014/Lincoln and the Hungarian strains
290 HUN/2012/2 and HUN/2012/135 formed a unique group, with more than 91% nt identity in the
291 full-length ORF2. A second group included three Italian strains (ITA/2005-3, ITA/2005-6, and
292 ITA/2005-8), the UK strain GBR/2014/Gillingham and strain HUN/2012/6, with more than 91%
293 nt identity to each other. A third group encompassed the virulent strain ITA/2010/Zoid and strain
294 HUN/2012/126 (97% nt identity to each other). Other canine strains (GBR/2014/Braintree,
295 GBR/2014/Huntingdon, HUN/2012/115) were intermingled between these three main sequence
296 groups and, as supported by sequence comparison (Table 3), they clearly appeared as distinct
297 genetic sub-lineages.
298
299 **4. Discussion**
300 AstVs have been reported repeatedly in dogs across the world, although the association
301 with the disease, prevalence rates and information on the genetic diversity are still largely
302 unknown. In this study, we could neither confirm nor refute an association between AstV

303 infection and diarrhea in dogs. Furthermore, we observed discrepancies between results obtained
304 by the pan-astrovirus PCR assay and those derived from viral metagenomics. However, none of
305 these methods were developed for routine diagnostic procedures and we do not have information
306 about the actual sensitivity of these assays. At present it seems more appropriate to consider these
307 methods, originally developed to detect viral diversity, as independent approaches that help
308 understand the viral community and diversity in clinical/pathological specimens. Thus, our
309 prevalence data for AstV must be interpreted with caution, even if the detection rates of AstVs
310 fell in the ranges reported in the literature. For example, canine AstV RNA has been detected in 6
311 to 27% of dogs with enteritis and 0 to 19% of aysmptomatic dogs in studies in Italy (Martella et
312 al., 2011), China (Zhu et al., 2011), Japan (Takano et al., 2015), France (Grellet et al., 2012) and
313 UK (Caddy and Goodfellow, 2015). Even if those studies differed in the number, provenience
314 and age of the animals enrolled in the investigations, in most studies the detection rates were
315 found to differ significantly between symptomatic and asymptomatic dogs, suggesting a possible
316 role of AstVs as enteric pathogens of dogs. Also, the onset and evolution of the clinical signs
317 have been found to correlate with the patterns of virus shedding and with seroconversion in two
318 dogs naturally infected by canine AstVs (Martella et al., 2012). Yet, experimental studies will be
319 needed to demonstrate firmly the association of canine AstVs with gastroenteritis.

320 Sequencing of the ORF1a, ORF1b and ORF2 regions revealed significant sequence
321 variation in the CAstVs detected in this study. Five strains were grouped with other canine AstVs
322 identified worldwide, but further separated into multiple capsid (ORF2) sub-clusters. By
323 parallelism with what is observed in human AstVs (MAstV species 1), that are classified
324 antigenically in 8 serotypes, it is possible to speculate that the genetic heterogeneity observed in
325 canine AstVs might account for a significant antigenic diversity (Caddy and Goodfellow, 2015).
326 Weak antigenic cross-reactivity has been observed in immune electron microscopy between

327 strain ITA/2008/Bari and ITA/2010/Zoid virus particles and the sera of convalescent dogs
328 (Martella et al., 2012). However, isolation in vitro of these viruses is fastidious and only one
329 strain, ITA/208/Bari, has been adapted to grow in vitro thus far (Martella et al., 2011), thus
330 hampering a precise evaluation of the antigenic relationships among different canine AstVs.

331 Interestingly, one Hungarian strain, HUN/2012/126, resembled a highly virulent canine
332 AstV, ITA/2010/Zoid, associated with enteric disease in an adult and young dog in Italy
333 (Martella et al., 2012). The question whether some canine AstV strains may differ in their
334 biological properties in terms of fitness or virulence remains open.

335 Even more interestingly, one of the canine AstVs identified in Hungary, HUN/2012/8,
336 was found to be genetically distantly related from all other canine AstVs identified worldwide.
337 Although we were successful to determine only the partial (3.2 kb) 3'end of the genome, the
338 virus displayed unique genetic signatures, such as a longer ORF1b/ORF2 overlap (11 nt), and a
339 longer 3'UTR. Also, this strain markedly varied from all the other canine AstVs in the highly
340 conserved promoter region for synthesis of subgenomic RNA, located upstream of the ORF2
341 (Walter and Mitchell, 2003; Mendez et al., 2014). Conversely, all other canine AstVs analysed in
342 this study displayed a conserved genomic architecture, which was shared with all currently
343 canine AstVs for whom the complete or partial (ORF1b and ORF2) genome sequence is available
344 (Martella et al., 2011, 2012; Takano et al., 2015). In addition, upon sequence comparison and
345 phylogenetic analysis, strain HUN/2012/8 differed markedly from other canine AstVs concerning
346 the partial ORF1b and the full-length ORF2 (Table 3**).** According to the 9th ICTV report
347 (http://talk.ictvonline.org/files/ictv_official_taxonomy_updates_since_the_8th_report/m/vertebrat
348 e-official/4178), for MAstVs the mean aa genetic distances (p-dist) range from 0.378 to 0.750%,
349 and from 0.006 to 0.312%, respectively, between and within groups (species). Using our set of
350 ORF2 sequences, the p-dist of strain HUN/2012/8 was calculated to range between 0.350 and

0.772. Accordingly, this strain appears distantly related from any other known mammalian AstV and may represent a novel AstV species.

It remains to be established whether strain HUN/2012/8 is actually a canine virus, or it is the result of inter-species transmission from an unidentified animal source. Although AstVs are regarded as viruses with a robust host-species restriction, several pieces of evidences in mammals (including humans) indicate that the host species barrier may be permeable in some occasions, and that heterologous viruses may stably adapt to the new host (Finkbeiner et al., 2008, 2009; Nagai et al., 2015). Moreover, occasional transmission of AstVs is thought to have occurred between avian and mammalian species in both directions (Sun et al., 2014; Pankovics et al., 2015).

In summary, data about canine AstVs is currently scarce. Nonetheless, the genetic heterogeneity among circulating AstV strains is evident representing some challenge for laboratory diagnosis. This diversity may also be challenging for future prophylaxis strategies, complicating the design of future vaccines, as it is currently unknown whether antigenic cross-reaction exists among and cross-protection is elicited by different canine AstV strains. Further efforts are needed to better understand the biology of canine AstVs and large-scale structured surveillance programs should be initiated to elucidate the relative veterinary importance of various canine AstV types.

Conflict of interest statement

The authors declare no conflicts of interest.

375

376 **Funding**

377

378 The study was supported by the Momentum program awarded by the Hungarian Academy of
379 Sciences. S.M. was a recipient of the János Bolyai fellowship (awarded by the Hungarian
380 Academy of Sciences). F.J. was supported by TÁMOP (4.2.4.A/2-11-1-2012-0001) and ÚNKP-
381 16-4-III, New Excellence Program of the Ministry of Human Capacities. G.K. received funding
382 from ÚNKP-16-3-III, New Excellence Program of the Ministry of Human Capacities The funders
383 had no role in study design, data collection, analysis and interpretation, or the decision to submit
384 the work for publication.

385

386 **References**

387

388 Blomström, A.L., Widén, F., Hammer, A.S., Belák, S., Berg, M., 2010. Detection of a novel
389 astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral
390 metagenomics. J. Clin. Microbiol. 48, 4392–4396. doi: 10.1128/JCM.01040-10.
391 Bosch, A., Guix, S., Krishna, N., Mendez, E., Monroe, S.S., Pantin-Jackwood, M., Schultz-
392 Cherry, S., 2010. Nineteen new species in the genus Mamastrovirus in the Astroviridae
393 family. ICTV 2010018aV2010b.
394 Caddy, S.L., Goodfellow, I., 2015. Complete genome sequence of canine astrovirus with
395 molecular and epidemiological characterisation of UK strains. Vet. Microbiol. 177,
396 206–213. doi: 10.1016/j.vetmic.2015.03.011.
397 Chu, D.K.W., Poon, L.L.M., Guan, Y., Peiris, J.S.M., 2008. Novel astroviruses in insectivorous
398 bats. J. Virol. 82, 9107–9114. doi: 10.1128/JVI.00857-08.

399 De Benedictis, P., Schultz-Cherry, S., Burnham, A., Cattoli, G., 2011. Astrovirus infections in
400 humans and animals - molecular biology, genetic diversity, and interspecies
401 transmissions. Infect. Genet. Evol. 11, 1529–1544. doi: 10.1016/j.meegid.2011.07.024.
402 Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., Feldblyum, J., Sengamalay, N., Afonso, C.,
403 Zhang, X., Anderson, N.G., Ghedin, E., Spiro, D.J., 2008. Viral genome sequencing by
404 random priming methods. BMC Genomics. 9, 5. doi: 10.1186/1471-2164-9-5.
405 Finkbeiner, S.R., Kirkwood, C.D., Wang, D., 2008. Complete genome sequence of a highly
406 divergent astrovirus isolated from a child with acute diarrhea. Virol. J. 5, 117. doi:
407 10.1186/1743-422X-5-117.
408 Finkbeiner, S.R., Holtz, L.R., Jiang, Y., Rajendran, P., Franz, C.J., Zhao, G., Kang, G., Wang, D.,
409 2009. Human stool contains a previously unrecognized diversity of novel astroviruses.
410 Virol. J. 6, 161. doi: 10.1186/1743-422X-6-161.
411 Fu, Y., Pan, M., Wang, X., Xu, Y., Xie, X., Knowles, N., Yang, H., Zhang, D., 2009. Complete
412 sequence of a duck astrovirus associated with fatal hepatitis in ducklings. J. Gen. Virol.
413 90, 1104–1108. doi: 10.1099/vir.0.008599-0.
414 Grellet, A., De Battisti, C., Feugier, A., Pantile, M., Marciano, S., Grandjean, D., Cattoli, G.,
415 2012. Prevalence and risk factors of astrovirus infection in puppies from French
416 breeding kennels. Vet. Microbiol. 157, 214–219. doi: 10.1016/j.vetmic.2011.11.012.
417 Imada, T., Yamaguchi, S., Mase, M., Kenji, T., Masanori, K., Akira, M., 2000. Avian Nephritis
418 Virus (ANV) as a new member of the family Astroviridae and construction of infectious
419 ANV cDNA. J, Virol. 74, 8487–8493.
420 Käll, L., Krogh, A., Sonnhammer, E.L.L., 2007. Advantages of combined transmembrane
421 topology and signal peptide prediction--the Phobius web server. Nucleic Acids Res. 35,
422 W429-432. doi: 10.1093/nar/gkm256.

423 Marshall, J.A., Healey, D.S., Studdert, M.J., Scott, P.C., Kennett, M.L., Ward, B.K., Gust, I.D.,
424 1984. Viruses and virus-like particles in the faeces of dogs with and without diarrhoea.
425 Aust. Vet. J. 61, 33–38.
426 Martella, V., Moschidou, P., Buonavoglia, C., 2011. Astroviruses in dogs. Vet. Clin. North. Am.
427 Small. Anim. Pract. 41, 1087–1095. doi: 10.1016/j.cvsm.2011.09.001.
428 Martella, V., Moschidou, P., Lorusso, E., Mari, V., Camero, M., Bellacicco, A., Losurdo, M.,
429 Pinto, P., Desario, C., Bányai, K., Elia, G., Decaro, N., Buonavoglia, C., 2011.
430 Detection and characterization of canine astroviruses. J. Gen. Virol. 92, 1880–1887. doi:
431 10.1099/vir.0.029025-0.
432 Martella, V., Moschidou, P., Catella, C., Larocca, V., Pinto, P., Losurdo, M., Corrente, M.,
433 Lorusso, E., Bànyai, K., Decaro, N., Lavazza, A., Buonavoglia, C., 2012. Enteric
434 disease in dogs naturally infected by a novel canine astrovirus. J. Clin. Microbiol. 50,
435 1066–1069. doi: 10.1128/JCM.05018-11.
436 Martella, V., Pinto, P., Tummolo, F., De Grazia, S., Giammanco, G., Medici, M.C., Ganesh, B.,
437 L'Homme, Y., Farkas, T., Jakab, F., Bányai, K., 2014. Analysis of the ORF2 of human
438 astroviruses reveals lineage diversification, recombination and rearrangement and
439 provides the basis for a novel sub-classification system. Arch. Virol. 159, 3185–3196.
440 doi: 10.1007/s00705-014-2153-9.
441 Mendez, E., Arias, C.F., 2007. Astroviruses. In: Knipe, D.M., Howley, P.M. (Eds.), Fields
442 Virology, 5th edn. Lippincott Williams and Wilkins, Philadelphia pp. 981-1000.
443 Mendez, E., Muñoz-Yañez, C., Sánchez-San Martín, C., Aguirre-Crespo, G., Baños-Lara Mdel,
444 R., Gutierrez, M., Espinosa, R., Acevedo, Y., Arias, C.F., López, S., 2014.
445 Characterization of human astrovirus cell entry. J. Virol. 88, 2452–2460. doi:
446 10.1128/JVI.02908-13.

447 Mihalov-Kovács, E., Marton, S., Fehér, E., Lengyel, G., Jakab, F., Tuboly, T., Bányai, K., 2014.
448 Enteric viral infections of sheltered dogs in Hungary [in Hungarian]. Magy. Allatorv.
449 Lapja. 136, 661–670.
450 Mihalov-Kovács, E., Gellért, Á., Marton, S., Farkas, S.L., Fehér, E., Oldal, M., Jakab, F.,
451 Martella, V., Bányai, K., 2015. Candidate new rotavirus species in sheltered dogs,
452 Hungary. Emerg. Infect. Dis. 21, 660–663. doi: 10.3201/eid2104.141370.
453 Monceyron, C., Grinde, B., Jonassen, T.O., 1997. Molecular characterisation of the 3'-end of the
454 astrovirus genome. Arch. Virol. 142, 699–706.
455 Nagai, M., Omatsu, T., Aoki, H., Otomaru, K., Uto, T., Koizumi, M., Minami-Fukuda, F., Takai,
456 H., Murakami, T., Masuda, T., Yamasato, H., Shiokawa, M., Tsuchiaka, S., Naoi, Y.,
457 Sano, K., Okazaki, S., Katayama, Y., Oba, M., Furuya, T., Shirai, J., Mizutani, T., 2015.
458 Full genome analysis of bovine astrovirus from fecal samples of cattle in Japan:
459 identification of possible interspecies transmission of bovine astrovirus. Arch. Virol.
460 160, 2491–2501. doi: 10.1007/s00705-015-2543-7.
461 Pankovics, P., Boros, Á., Kiss, T., Delwart, E., Reuter, G., 2015. Detection of a mammalian-like
462 astrovirus in bird, European roller (Coracias garrulus). Infect. Genet. Evol. 34, 114–121.
463 doi: 10.1016/j.meegid.2015.06.020.
464 Schultz-Cherry, S., 2013. Astrovirus Research. Essential Ideas, Everyday Impacts, Future
465 Directions. Springer New York, New York.
466 Sun, N., Yang, Y., Wang, G.S., Shao, X.Q., Zhang, S.Q., Wang, F.X., Tan, B., Tian, F.L., Cheng,
467 S.P., Wen, Y.J., 2014. Detection and characterization of avastrovirus associated with
468 diarrhea isolated from minks in China. Food. Environ. Virology. 6, 169–174. doi:
469 10.1007/s12560-014-9155-3.

470 Takano, T., Takashina, M., Doki, T., Hohdatsu, T., 2015. Detection of canine astrovirus in dogs
471 with diarrhea in Japan. Arch. Virol. 160, 1549–1553. doi: 10.1007/s00705-015-2405-3.
472 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular
473 Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729. doi:
474 10.1093/molbev/mst197.
475 Toffan, A., Jonassen, C.M., De Battisti, C., Schiavon, E., Kofstad, T., Capua, I., Cattoli, G.,
476 2009. Genetic characterization of a new astrovirus detected in dogs suffering from
477 diarrhoea. Vet. Microbiol. 139, 147–152. doi: 10.1016/j.vetmic.2009.04.031.
478 Vieler, E., Herbst, W., 1995. Electron microscopic demonstration of viruses in feces of dogs with
479 diarrhea. Tierärztl. Prax. 23, 66–69.
480 Walter, J.E., Mitchell, D.K., 2003. Astrovirus infection in children. Curr. Opin. Infect. Dis. 16,
481 247–253. doi: 10.1097/01.qco.0000073775.11390.60.
482 Williams, F.P. Jr., 1980. Astrovirus-like, coronavirus-like, and parvovirus-like particles detected
483 in the diarrheal stools of beagle pups. Arch. Virol. 66, 215–226.
484 Zhu, A.L., Zhao, W., Yin, H., Shan, T.L., Zhu, C.X., Yang, X., Hua, X.G., Cui, L., 2011.
485 Isolation and characterization of canine astrovirus in China. Arch. Virol. 156, 1671–
486 1675. doi: 10.1007/s00705-011-1022-z.
487

488 **Figure legends and tables**

489

490 **Figure 1.** Distribution of eukaryotic viral sequences in 26 fecal specimens collected from
491 sheltered dogs.

492

493 **Figure 2.** Schematic representation of the genome of canine AstV strain HUN/2012/115. The
494 upper line represents the complete nt length of the genome from the 5’ untranslated region (UTR)
495 through the 3’ UTR and the poly-A tail (pA). The relative position of the slippery sequence
496 AAAAAAC and downstream stem loop structure is indicated. Also, the position of the
497 subgenomic RNA (sgRNA) promoter and of the stem loop-II-like (s2m) structure are indicated.
498 Rectangles represent the three possible open reading frames (ORFs) (light grey) and the alternate
499 ORF1b and ORF2 products due to in-frame start codons (dark grey). The starting aa motifs of the
500 various ORFs and the ending aa motif of ORF2 are marked. The potential trans-membrane
501 regions (TM), the predicted protease motif (PRO), the genome-linked protein (VPg) and the
502 polymerase region (POL) are also shown.

503

504 **Figure 3.** Phylogenetic trees based on the full-length ORF1a (3A), ORF1b (3B) and ORF2 (3C)
505 of canine AstVs with a selection of mammalian AstVs. The tree was generated using the
506 neighbor-joining method, with the Jukes Cantor algorithm of distance correction and
507 bootstrapping over 1,000 replicates. The scale bar indicates the number of nt substitutions per
508 site. Black circles indicate the AstV strains detected in this study.

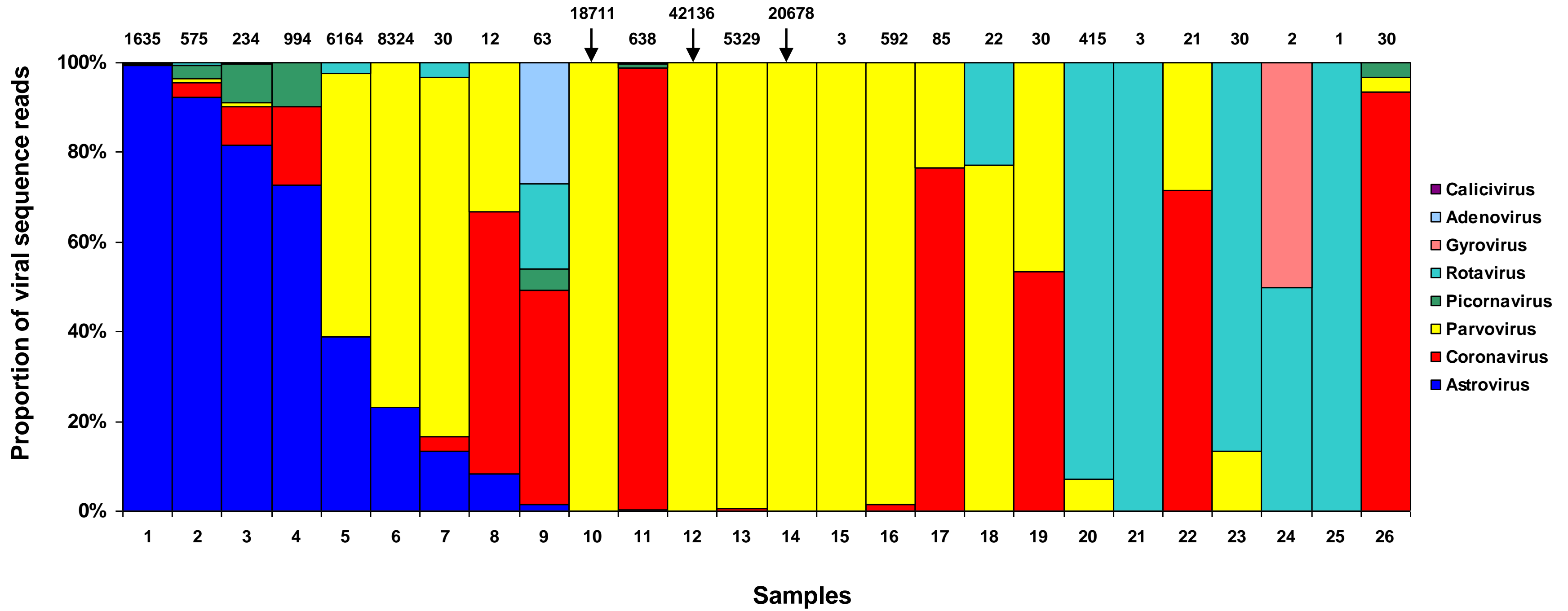
509

510

Table 1

Origin and information on canine stool samples positive for AstV by broad-range RT-PCR and viral metagenomics

Sample name	Date of sampling	Age of dog	Diarrheic symptoms	Screening by pan-astrovirus PCR	AstV reads by metagenomics
HUN/2012/2	20/01/2012	5 weeks	Yes	Positive	723
HUN/2012/4	26/01/2012	5 weeks	Yes	Positive	0
HUN/2012/5	30/01/2012	3 months	Yes	Positive	0
HUN/2012/6	03/02/2012	8 years	No	Negative	191
HUN/2012/8	10/02/2012	4 months	Yes	Positive	0
HUN/2012/22	17/02/2012	3 months	No	Negative	4
HUN/2012/14	09/03/2012	Pool of 3 adults	No	Positive	0
HUN/2012/16	09/03/2012	Pool of 2 adults	No	Positive	0
HUN/2012/114	15/04/2012	6 years	No	Negative	1
HUN/2012/115	20/04/2012	6 months	Yes	Negative	1922
HUN/2012/116	20/04/2012	1 year	Yes	Negative	1
HUN/2012/126	27/04/2012	under one year	No	Negative	1627
HUN/2012/135	09/05/2012	3 weeks	No	Negative	2398
HUN/2012/175	05/06/2012	6 weeks	No	Positive	531
HUN/2012/527	23/08/2012	1 year	Yes	Negative	1
HUN/2012/528	07/09/2012	2 year	Yes	Negative	1



1635
575
234
994
6164
8324
30
12
63
18711
638
42136
5329
20678
3
592
85
22
30
415
3
21
30
2
1
30
100%
80%
60%
40%
20%
0%
Proportion of viral sequence reads
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
Samples
Calicivirus
Adenovirus
Gyrovirus
Rotavirus
Picornavirus
Parvovirus
Coronavirus
Astrovirus