University of Veterinary Medicine
Doctoral School of Veterinary Science

Novel viruses in enteral infections of dogs in Hungary

Summary of PhD thesis

dr. Mihalov-Kovács Eszter

2019
University of Veterinary Medicine

Doctoral School of Veterinary Science

Supervisor:

Dr. Bányai Krisztián, PhD

Institute for Veterinary Medical Research
Centre for Agricultural Research
Hungarian Academy of Sciences
Introduction

Dogs are the most popular pets worldwide, including this country. As the culture of dog keeping advances, owners pay more and more attention to protect the health of their pets, to maintain hygienic conditions and to prevent infectious diseases. On the other side, there is still a high number of stray dogs which end up in dog ponds or animal shelters.

A number of factors in these facilities (crowdedness, stress, vaccination program) promote and facilitate the propagation of viruses and the appearance of novel virus types.

Vaccination provides protection against long known enteral pathogens such as canine parvovirus 2, canine adenoavirus 1, or canine distemper virus. However, in spite of the immunity provided by vaccination, the emergence of diarrhoeal diseases seems to be unavoidable every now and then. The underlying causative agent in these cases is usually a pathogen which is substantially different from the vaccine virus or is a novel type of virus altogether. Therefore, more attention should be paid to investigate the diarrhoeal incidents of unknown causes and to detect the novel agents.

In the case of certain viruses such as canine rotavirus, canine astrovirus or canine caliciviruses, their presence is known for a while, and their pathogen nature is also suspected. However, our team was the first to successfully demonstrate the incidence of these viruses in dogs in Hungary.

Scientific knowledge about the variety and diversity of enteral canine viruses has been expanded rapidly with the use of polymerase chain reaction systems (PCR) built on general primers, matching the conservative sequences of virus genomes and of the deep sequencing technique following random DNA amplification gaining ground lately. We focused on the diagnostic potential of the NGS, and also mapping the viral diversity of canine faecal virome in a shelter. These methods were used among others for the identification of canine kobuviruses and canine vesiviruses from dogs in the *Picornaviridae* and *Caliciviridae* families, respectively.
Objectives

In the course of our current research faecal samples were collected at a shelter in Northern-Hungary, then the viruses detected with the combination of PCR systems and deep sequencing methods.

The main goal of the research was to detect enteral pathogen agents not yet determined from dogs using the PCR technique, and with the use of the new generation sequencing.

Another goal was to obtain a better understanding of the diversity seen among the viruses causing diarrhoea and mapping out the faecal virome of dogs at the same time, for which purpose the special micro-environment found at the dog shelter could provide an ideal background.

Analysis and mapping out of the phylogenetic connections of the viruses isolated from the faeces of dogs was another aim.
Materials and methods

Origin and processing of the faecal samples up to the extraction of the nucleic acid

Faecal samples originating from dogs were collected at an animal shelter situated in the north-east of Hungary (Miskolc Animal Shelter, M.Á.S.A.) in the period between January and September 2012. Samples were also collected from the dogs owned by a private individual who also keeps dogs received on a provisional basis beside the others. Less dogs occur at this location, typically only clinically healthy dogs live together here.

Sampling containers were used to collect a total of 76 faeces samples. Collection was made on a random basis, symptomless and diarrhoeal individuals were included in a mixed cohort, and collection was not influenced by age or sex. Samples were taken both individually and from kennels, and in the latter case it may have happened that multiple samples were taken into the same sampling container. Samples were kept at -20°C until transported or taken to the laboratory right away, on the same day.

Suspensions of 10-20% concentration levels were prepared from the faecal samples using PBS (phosphate buffered saline) solution. Subsequently the samples were stored at -65°C until used.

Next generation sequencing – sample preparation, genetic library and sequencing

RNA was extracted from the supernatant of the faecal samples for the purposes of the next generation sequencing (NGS) procedure. Reverse transcription (RT) was accomplished from the RNA samples in the presence of a random hexamer oligonucleotide, to the 3’ end of which a general oligonucleotide was attached. The cDNA product of the RT step was used to carry out PCR amplification and the product was run in agarose gel. The products falling within the 200 and 2000 nucleotide (nt) range were cut out and cleaned from the gel.

Enzymatic fragmentation was made during library compilation and bar coded adapters were associated with the fragments. After quantitative determination by fluorimeters, identical quantities were measured from the indexed libraries. Such library mixtures were used for the purposes of the emulsion PCR, with the aim to achieve clonal amplification of the library DNA. Sequencing was carried out on an Ion Torrent Personal Genome Machine®, using 316 and 318 type chips, respectively.
Conventional PCR screening – sample preparation, RT-PCR, real-time PCR

Combined viral DNA and RNA extraction was accomplished from the faeces supernatant in order to achieve conventional PCR based screening. Viruses were detected with the use of virus specific RT-PCR, in separate reaction tubes for each virus. One and two steps RT-PCR sessions and the heminested PCR served the detection of canine CAstV, CaKoV and CaCV. The oligonucleotides used for the work were chosen the literature and in the case of the CaKoV it was designed by the research team using matches from GenBank sequences. Bufavirus screening was accomplished in the light of preliminary metagenomic data, while the protocol used in the real-time PCR was developed in cooperation and made available by Vito Martella, University of Bari, Italy.

Following agarose gel electrophoresis the specific PCR products were purified from the gel slices. The extracted PCR products were used to direct sequence determination from both directions with our cooperating partners in the Sanger-type DNA sequencing (PTE TTK and Honvédkórház).

Genome ends of the single strand RNA viruses were amplified with the help of the 5’ and 3’ RACE methods. In the case of double strand RNA viruses determination of the segment ends was achieved by RNA ligation combined with 3’ RACE adapted in house.

Analysis of the sequencing data obtained from the next generation sequencing and the conventional Sanger-type DNA sequencing

The process of bioinformatics analysis included the following key steps: the raw sequence reads obtained from the NGS were subjected to quality control, and in the next step reference matching was carried out, disregarding the too short (≤35 nt) or poor quality nucleotide segments and by deleting the adapter/bar code regions, for which purpose the database consisting of virus sequences available in the GenBank was used.

Assembly and analysis of the viral genomes was completed by the research team. Data generated in the course of the sequencing process were arranged and analysed with the help of the CLC Genomics Workbench software programme (https://www.qiagenbioinformatics.com/products/clc-main-workbench). The same software package was used when the sequence reads were compared to virus sequences infecting Eukaryotes, which were downloaded from the GenBank. Matched individual sequences were checked and a single contig (a major section of the sequence consisting of overlapping reads) was compiled from the matches.

Editing and evaluation of the sequences obtained from the Sanger sequencing process was made using the GeneDoc and BioEdit software. Similarity was established
using the BLAST algorithm, through the web based search engine of the NCBI (National Center for Biotechnology Information).

Sequence particulars obtained from the conventional Sanger sequencing and the NGS sequencing particulars were analysed the same way from this point on. Open reading frames were determined with the web based open reading frame (ORF) search engine of the NCBI (https://www.ncbi.nlm.nih.gov/gorf/gorf.html). Multiple sequence matching was accomplished using the translatorX (http://translatorx.co.uk/) interactive translator matching program. Additional editing and evaluation of the sequences was made using the GeneDoc and BioEdit software. Phylogenetic analysis was completed using the MEGA6 software and the neighbour-joining or maximum likelihood methods, with the application of the substitution model determined as the best one by the program. Average genetic distances of the nucleic acids and amino acids were calculated with the p-distance method, also in the MEGA6 program.
Results and discussion

A total of 76 dog faeces samples were taken in the year of 2012 from the Miskolc site of M.Á.S.A., and from the dogs of a private individual. A total of 37 samples were tested using metagenomic methods. Thirty-three of the samples originated from the shelter and four of the samples were collected from animals kept by the private owner.

Virus-specific PCR and RT-PCR systems were developed based on the preliminary virus metagenomic data or adapted from the literary references. The detailed results of the virus metagenomic tests and of the PCR screening as well as the conclusions are presented with the discussions of each of the viruses, for didactic considerations.

Canine astrovirus, CAstV

Astroviroses were detected in 21% of all samples by merging the results of pan-astrovirus RT-PCR and the metagenomic tests (16/76).

Four complete AstV genomes (HUN/2012/2, HUN/2012/115, HUN/2012/126, HUN/2012/135), and two partial genomes (HUN/2012/8, HUN/2012/6) were analysed in details. The structure of the encoding sections showed a picture typical for AstV. It contained three overlapping ORFs, which encode structural and non-structural proteins. The length of the encoding sections and the GC (guanine: cytosine) contents varied between 6441-6447 nt and 45.19-45.48%, respectively. The length of the genome, including the 5’ and 3’ UTR and excluding the poly(A) tail was between 6535 and 6587 nt. The 5’ UTR could be determined in the case of four strains (HUN/2012/2, HUN/2012/115, HUN/2012/126 and HUN/2012/135), their respective length varied between 42 and 60 nt. The 3’ UTR was sequenced in the case of all the six strains and their length varied between 49 and 100 nt.

The phylogenetic analysis of the entire ORF2 nucleotide sequence provides the basis for taxonomic classification and characterisation of AstVs. In this analysis all of the Chinese strains, the strain from the United Kingdom, GBR/2014/Lincoln, and the HUN/2012/2 as well as the HUN/2012/135 strains constitute a unique group, with an nt similarity level across the group of more than 91%. The second group contained the Italian strains (ITA/2005-3, ITA/2005-6 and ITA/2005-8), plus GBR/2014/Gillingham from Great Britain and HUN/2012/6 from Hungary, again with a nucleotide level similarity of more than 91%. The third group contains the virulent ITA/2010/Zoid and HUN/2012/126 strains, with a 97% similarity at the nucleotide level within the group. The other CAstV strains (GBR/2014/Braintree, GBR/2014/Huntingdon, HUN/2012/115) are situated between the strains referred to above and constitute separate sub-groups.

The examination of the ORFs of the AstVs sequenced in this project pointed out the diversity of the sequences. In the phylogenetic examinations five of the strains were
arranged together with the other canine astroviruses, and constituted several sub-groups on the basis of their capsid sequences. Human astroviruses are listed in the species MAstV 1, and are divided into eight different serotypes. It might be possible that the sub-groups observed at the human astroviruses are typical for canine astroviruses just as well. The Hungarian HUN/2012/8 CAstV strain showed only remote relationship with both the other Hungarian strains and the CAstV strains originating from other parts of the world. In the case of the HUN/2012/8 strain it is uncertain, whether this strain is a canine CastV per se, or was detected in dogs as the result of the cross species virus transfer.

Our knowledge is still very incomplete in terms of the CAstVs. Nevertheless, the rich diversity explored so far puts laboratory diagnostics to great challenges. The same diversity may represent an impediment to the design and implementation of future preventive strategies such as the development of a vaccine. Additional tests are necessary to understand the biology of the CAstV better, and surveillance programs are needed to assess the importance of the virus in terms of animal health implications.

**Canine calicivirus, CaCV**

Canine calicivirus RNA was detected in 3.9% of all samples (3/76) using the RT-PCR method, while the metagenomic analysis identified calicivirus (i.e. vesivirus) sequence reads in one case, from which the partial genome of the HUN/2012/528 strain could be successfully assembled. Having sequenced the PCR products of three samples positive for calicivirus, one of the samples (HUN/2012/528) was identified as vesivirus, while the other two (HUN/2012/7 and HUN/2012/64) were noroviruses.

The length of the partial encoding sequence of the HUN/2012/528 vesivirus strain was 8270 nucleotides. The full 3’ UTR was 162 nucleotides long, not counting the poly(A) tail. GC content of the sequence was 47.56%. The RNA genome contained three ORFs.

The amino acid sequence (as) of the capsid protein and similar sequence characteristics thereof provide the basis for classification of the virus strains. The HUN/2012/528 vesivirus strain showed 84.8-91.5% amino acid similarity with the cell line contaminant vesivirus strains (87.4%-Allston/2009/USA, 88.3%-Allston/2008/USA, 91.2%-Geel/2008/BEL, 86.5-2117/DEU), with canine vesivirus strains classified in Type II by Binn et al. (3-68, W191R), and with the CU/296 strain (similarity level of 86.9-92.3%). At the same time, in the comparative analysis carried out using the 48 canine vesivirus reference strains only 70.3% of similarity could be detected. Analysing our strain in accordance with the criteria listed in the ninth ICTV report it seems that the HUN/2012/528 strain might be a novel representative of the *Vesivirus* genus, together with the cell line contaminant vesivirus strains, and the vesivirus strains described recently in the United States and in Italy.
Prevalence specified in this paper (3.9%) falls in line with other values mentioned in the references (1.1-64.8%), but it was found lower than the level of prevalence detected in kennels and dog shelters (9.7%, 64.8%). In a unique way, norovirus and vesivirus were both detected from a site in the course of the work. The norovirus strains could be detected using the RT-PCR methodology, while the vesivirus strain was identified with the general calicivirus primer, and metagenomic methods.

**Canine kobuvirus, CaKoV**

A total of 11 samples of the 76 faecal samples tested (14.47%) proved positive for kobuvirus using the RT-PCR and/or metagenomic method. The RT-PCR system is our own development, it is designed with two primer pairs matching the RdRp and VP1 genome regions each. Virus metagenomic methods were used to detect and identify kobuvirus sequence reads in six samples, and one of them (HUN/2012/2) was used to determine a 8279 nt long section from the entire genome by consolidating the data generated in parallel runs, and extending consensus ends. The encoding section showed a great degree of similarity with the CaKoV 1 sequences found in the GenBank (92.2-94.8% at the nt level, 97-99.4% at the as level). The genome structure corresponded to that of the other members of the genus: VPg+5'UTR[L/P1(VP0-VP3-VP1)-P2(2A<sup>Hbox/NC-2B-2C<sup>Hel</sup>)-P3(3A-3B<sup>VPg-3C<sup>Pro</sup>-3D<sup>pol</sup>)-3'UTR].

Phylogenetic calculations were carried out to determine the position of the CaKoV strain among kobuviruses. The basis for the calculation was provided by the as sequence of the polyprotein P1 region. The Hungarian CaKoV was grouped on the phylogenetic tree together with the other CaKoV strains, creating a monophyletic branch. Aichivirus A was situated closest to the CaKoVa from among the Aichi virus species approved by the ICTV. The HUN/2012/2 CaKoV is clearly separated from the other members of the *Kobuvirus* genus, i.e. from the Aichivirus B, C, D, E, F species.

Eight and three of the kobuvirus positive samples tested originated from dogs under the age of one year and three adult animals, respectively, six showing symptoms of diarrhoea and five being symptomless. CAstV and CCoV were present in the HUN/2012/2 CaKoV strain analysed in details as co-infective agents. These data taken all around with the knowledge from the previous studies shows, that kobuvirus infection might be common among dogs, but it usually does not show clinical symptoms, or other known pathogens contribute to the development of the diarrhoeal symptoms.
**Canine rotavirus C, CRVC**

In the course of the virus-metagenomic tests sequence reads akin to *Rotavirus C* could be isolated from the faeces sample (HUN/174/2012) of a ten weeks old diarrhoeal puppy. No canine RVC strain sequence was uploaded to the online databases so far, thus an attempt was made to assemble the entire encoding sequence. The sequence of the strain identified was unique, and it distinctly differed from the human, porcine and bovine RVC strains. The genome of the RVC strain tested encoded 11 proteins, six of them structural and five non-structural proteins. Comparing the human, porcine and bovine reference strains with the RVC determined by the research team, the nucleotide sequence similarity levels varied among the genes of the RVC strains in the range of 67-84%. The least similar sequence compared to the reference strains was observed in the case of NSP1 and NSP4 (67-87%, and 67-77%, respectively), while the highest level of sequence similarity was found with NSP2 (79-84%). As for the genes VP7, VP4, VP6, our canine RVC strain can be classified in the G10–VP7, P8–VP4, and I8–VP6 genotypes with a great degree of certainty.

It was demonstrated during the phylogenetic analysis of the genes VP4, VP6 and VP7, which could be suspected from the similarity assessment, that the KE174/2012 strain represents a separate group compared to the genotypes described so far.

Metagenomic methods were used to identify a genetically heterogeneous *Rotavirus C* strain, which showed a novel genotype in most, if not all of its genes. Additional identification of canine RVC strains might be instrumental in revealing the potential genetic diversity among RV strains isolated from the host species. As the number of RVC strains described increases, it is expected to manage the development of a reliable classification system.

**Canine rotavirus I, CRVI**

Unusual rotavirus sequence reads were discovered during metagenomic testing in the case of two samples (KE135/2012; KE528/2012) in the preliminary analyses. However, since the virus was present in a low titre only in the sample (i.e. even silver staining carried out in PAGE could not detect the genome RNA), therefore the amount of sample sequences was drastically increased during deep sequencing. As a result of the running session, all structural and non-structural genes were obtained. These genes differed substantially from the rotavirus A-H reference sequences known so far.

The nucleotide sequence of segment ends was well preserved in both unusual rotaviruses (5’ and GGC/TA; 3’ end AACCC); and in the event of most genes the similarity of the sequence between the two was great (for instance VP2: 88% nt, 95% as similarity; NSP4 99% nt, 99% as similarity). In the case of the VP7 gene, however, the similarity of the sequences was extremely low (53% nt and 38% as).
Sequence and phylogenetic analyses together illustrate a moderate genetic similarity with the representatives of the rotavirus A-H species, substantiating that the representative of a new species, *Rotavirus I* were actually found. According to the most recent recommendations, the first representatives of the new rotavirus species were named RVI/Dog-wt/HUN/KE135/2012 and RVI/Dog-wt/HUN/KE528/2012.

Tell-tale signs alluding to the presence of viruses other than the Rotavirus species known so far were found earlier on in the faecal virome of cats and Californian sea lions, while RVI could be identified from the faeces sample of a diarrhoeal cat recently. These studies, concurring with our paper, confirm the assumption that certain carnivores are being infected with a novel type of rotavirus.

**Canine bufavirus, CBuV**

A total of 26 of the 70 faecal samples tested was positive for bufavirus with the real-time PCR method, while the metagenomic method identified sequence reads alluding to a novel type of parvoviruses in two samples (HUN/2012/22 and HUN/2012/126).

Nearly the entire length of the genome sequence for two new protoparvovirus strains could be determined with the metagenomic methods (HUN/2012/2 and HUN/2012/126). The encoding sequence was 4219 nt long in both cases, and if the partial UTRs were also included, 4463 nt and 4308 nt long sequences were obtained for the HUN/2012/22 and HUN/2012/126 samples, respectively.

Two ORFs were identified in the encoding sequence, the 1917 nucleotides long ORF1 on the left hand side encoded the non-structural protein (NS1, 638 as), while the 2316 nucleotides long ORF2 provided the code for the two protein molecules the capsid consisted of (VP1, 710 as; VP2, 568 as). The sequence match analysis for the entire encoding segment revealed that these new bufaviruses differed clearly from the known canine parvoviruses (<58% nt similarity) and grouped with other bufaviruses. CBuV showed greater similarity with bufaviruses detected from primate (61.6-63.2% nt), and porcine (59.6% nt) samples, while it reflected a 45% nucleotide level similarity with the CPV-2 strains.

During the phylogenetic analysis the canine bufaviruses set up a clearly distinguished group, with rat, bat, porcine, and primate bufaviruses situated around. Closest to our strains in the group were human bufaviruses, and bufaviruses detected from rhesus monkeys and pigs.

The new type of canine protoparvovirus was identified and analysed in faecal virome of dogs, and the samples were screened with a real-time PCR method to reveal potential etiological relationships. The identified virus, since its closest relatives belong to bufaviruses, was called canine bufavirus, CBuV. It showed low level of similarity at both the nt level (24.1-
69.4%), and as level (19.3-51.4%) with the other members of the *Protoparvovirus* genus in terms of the NS1 gene. The protoparvoviruses could be classified in the same species, when their NS1 protein as sequence is identical in at least 85% (criterion specified by the ICTV). Based on this feature, the CBuV strains determined with the metagenomic methods could be classified as a new PV species.

The CBuV DNA had a high level of prevalence in the samples, but the difference between the prevalence of healthy and diseased animals was not significant. This might refer to the fact that the CBuV was a frequent constituent of the faecal virome in dogs.
**New scientific results**

1. It was the first time ever to have a metagenomic study carried out in Hungary among dogs living in a shelter.
2. It was the first time in Hungary to identify astroviruses in dogs and to study the genomic structure of these viruses.
3. It was the first time in Hungary to isolate a canine calicivirus, which was proven as a canine vesivirus. The detailed genome of this strain was determined and analysed.
4. It was the first time in Hungary to isolate a canine kobuvirus using metagenomic and RT-PCR methods.
5. It was the first time in the world that the genome of a canine RVC was analysed.
6. A so far unknown rotavirus was described in dogs (*Rotavirus I*), adopted by ICTV as a new *Rotavirus* species since.
7. It was the first time in the world that a new bufalike parvovirus was isolated from dogs and the protein encoding genome sequence of these viruses was determined.
Publications derived from the results of the doctoral research work

1. Publications in peer-reviewed scientific journals with an impact factor


   IF: 7,422


   IF: 2,484


   IF: 3,192


   IF: 6,994


   IF: 0,212


   IF: 0,185


   IF: 0,364
2. Books, book chapters


3. Abstract in international conference proceedings


Magyar Mikrobiológiai Társaság Nagygyűlése 2014. október 16-18.: Mihalov-Kovács E., Marton Sz., Farkas Sz. L., Martella, V., Bányai K.: Új rotavírus faj kimutatása kutya fekáls viromjában

International Meeting on Emerging Diseases and Surveillance 2014. október 31- november 3.: Kovács E., Marton Sz., Tuboly T., Martella, V., Bányai K.: Genetic diversity and possible recombination events in canine astroviruses

4. Scientific publications not related to the doctoral research


Acknowledgements

Before all, I’d like to thank to my supervisors, Dr. Krisztián Bányai and Dr. Tamás Tuboly the privilege to be able to carry out my doctoral studies under their guidance. Tamás made me love laboratory and research work during my undergraduate years and through his help I could become a member of the Novel pathogen discovery thematic group led by Krisztián. The death of Tamás has left an enormous void which cannot be overcome. I am grateful to Krisztián for casting his trust in me, I personally developed both in human and professional terms during the years spent in his laboratory. I am also thankful to all members of the Novel pathogen discovery thematic groups for the huge amount of help in theory and practice as well as the happy atmosphere.

During my PhD work dr. Angéla Gulya, a veterinarian provided a lot of help in the sampling phase, permitting to collect samples at the Miskolc Animal Shelter. My thanks go to all employees at the shelter for the help in the sampling procedure.

My gratitude to dr. Vito Martella and prof. Canio Bounavoglia for providing assistance in processing the astrovirus, bufavirus and RVI samples, as well as to dr. Gianvito Lanave, who supervised, organised and helped my work during my stay in Italy (Dipartimento di Medicina Veterinaria, Università Aldo Moro di Bari, Italy).

I am very grateful to my family, who ensured a calm and undisturbed family background and supported, encouraged me all the time, which was particularly indispensable in the last phase of the PhD work. Thanks for the support to my partner for life, Csabi. And many thanks to my friends, Bari, Sziszkó, Eni, Reni, Katóka, and Little Reni for the lot of help. My mother, my colleague, dr. Andrea Vass, and dr. Ákos Kenéz deserve special thanks for the accurate and considerate reading of the manuscript and the useful advices. And last, but not least, I have to express my thanks to all employees at the Népkert Animal Clinic for the encouragement, and the flexibility, which allowed me to complete the writing of the doctoral dissertation.

Financial support for the work was provided by the Lendület program.