Canine rotavirus C strain detected in Hungary shows marked genotype diversity

Szilvia Marton,1 Eszter Mihalov-Kovács,1 Renáta Dóró,1 Tünde Csata,1 Enikő Fehér,1 Miklós Oldal,2,3 Ferenc Jakab,2,3 Jelle Matthijnssens,4 Vito Martella5 and Krisztián Bánya1

1Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian Academy of Sciences, Hungária krt. 21, Budapest 1143, Hungary
2Virological Research Group, Szentágathai Research Centre, University of Pécs, Pécs, Hungary
3Institute of Biology, Faculty of Sciences, University of Pécs, Pécs, Hungary
4KU Leuven-University of Leuven, Department of Microbiology and Immunology, Laboratory for Clinical and Epidemiological Virology, Rega Institute for Medical Research, B-3000 Leuven, Belgium
5Department of Veterinary Public Health, University of Bari, S.p. per Casamassima km 3, 70010 Valenzano, Bari, Italy

Species C rotaviruses (RVC) have been identified in humans and animals, including pigs, cows and ferrets. In dogs, RVC strains have been reported anecdotal on the basis of visualization of rotavirus-like virions by electron microscopy combined with specific electrophoretic migration patterns of the genomic RNA segments. However, no further molecular characterization of these viruses was performed. Here, we report the detection of a canine RVC in the stool of a dog with enteritis. Analysis of the complete viral genome uncovered distinctive genetic features of the identified RVC strain. The genes encoding VP7, VP4 and VP6 were distantly related to those of other RVC strains and were putatively classified as G10, P8 and I8, respectively. The new strain was named RVC/Dog-wt/HUN/KE174/2012/G10P[8]. Phylogenetic analyses revealed that canine RVC was most closely related to bovine RVC strains with the exception of the NSP4 gene, which clustered together with porcine RVC strains. These findings provide further evidence for the genetic diversity of RVC strains.

Rotaviruses (genus Rotavirus, family Reoviridae) are classified into eight approved and one candidate species designated Rotavirus A–H and Rotavirus I, respectively, on the basis of serological and genetic features (Estes & Greenberg, 2013; Matthijnssens et al., 2012, Mihalov-Kovács et al., 2015). The rotavirus genome is composed of 11 segments of dsRNA, which encode six structural proteins (VP1–P4, VP6 and VP7) and five non-structural proteins (NSP1–NSP5). The primary coding potential for non-structural proteins may vary across rotavirus species and even within a species (Mihalov-Kovács et al., 2015).

Rotavirus infections affect mainly the young of mammalian and avian host species. From both public health and veterinary health perspectives, rotavirus A (RVA) is one of the most significant cause of acute dehydrating diarrhea. Much less is known about the epidemiology and disease burden associated with infection by non-species A rotaviruses. However, RVA, RVB, RVC, RVE, RVH and RVI have been detected in sporadic infections, as well as in endemic or epidemic settings of various mammalian species, whereas RVD, RVF and RVG are typically found in domestic poultry, such as chicken and turkey (Martella et al., 2010; Marthaler et al., 2014; Matthijnssens et al., 2010).

Rotavirus C (RVC) has been identified as a cause of diarrhoea in humans, pigs, cows, ferrets and dogs (Bánya et al., 2006; Bridger et al., 1986; Collins et al., 2008; Kim et al., 1999; Marthaler et al., 2013; Mawatari et al., 2004; Otto et al., 1999; Torres-Medina, 1987). In humans, up to 10% of community-acquired diarrhoea episodes have been attributed to RVC (Phan et al., 2004). Food and

Received 15 April 2015
Accepted 2 July 2015

The GenBank/EMBL/DDBJ accession numbers for the sequence of KE174/2002 determined in this study are: VP1, KP988013; VP2, KP988014; VP3, KP988015; VP4, KP988016; VP6, KP988017; VP7, KP988018; NSP1, KP988019; NSP2, KP988020; NSP3, KP988021; NSP4, KP988022; and NSP5, KP988023.

Two supplementary figures are available with the online Supplementary Material.
water are commonly identified as the source of RVC-associated gastroenteritis outbreaks in all age groups (Otsu, 1998). Seroepidemiological studies indicate an increase in human seroprevalence by age group (Iturrizago´mar et al., 2004). In addition, the antibody prevalence in cattle and pigs (24–56 % and 58–100 %, respectively) shows that RVC infection may be very common in livestock herds in developed countries (Saif & Jiang, 1994). Unlike RVA, which is commonly described as a zoonotic virus (Martella et al., 2010), there is little evidence in support of RVC strains being transmitted from one host species to another. Examples include the identification of porcine RVC-derived genes detected in human and bovine RVC strains detected in Brazil and Korea, respectively (Gabbay et al., 2008; Jeong et al., 2009).

In dogs, RVC strains were first described in a study in Germany in the 1990s (Otto et al., 1999). By electron microscopy, rotavirus-like particles were detected in nine out of 26 faecal samples collected from dogs with enteric disease. Following PAGE, three samples showed an electropherotype similar to that of an RVC reference strain. However, the aetiological role of RVC in diarrhoea of dogs was formally not demonstrated. In addition, the diagnostic findings were not confirmed with specific molecular assays or with sequencing, thus hampering a precise characterization/classification of these viruses.

Whole-genome sequencing and phylogenetic analysis are becoming new standards in rotavirus strain characterization, enabling an in-depth understanding of the evolutionary history of rotaviruses in general and a novel RV strain in particular. Hundreds of RVA strains have had their whole genome sequenced during the past several years (Matthijnssens et al., 2011). In contrast, so far only one porcine RVC strain, eight bovine RVC strains and 10 human RVC strains have had their whole genome sequenced and described in independent studies (Chen et al., 2002; Mackow, 1995; Marton et al., 2015; Mawatari et al., 2014; Soma et al., 2013; Yamamoto et al., 2011). As cultivation of RVC strains is fastidious, molecular characterization remains the main approach to obtain insight into the biological features of RVC strains and determine their genetic diversity and evolution across different host species (Kusanagi et al., 1992).

In an effort to classify RVC strains into sequence-based genotypes, an approach that was developed to classify RVA strains was also recently applied to the RVC VP7, VP6 and VP4 genes, resulting in the description of nine G genotypes (G1–G9), seven P genotypes (P1–P7) and seven I genotypes (I1–I7) (Jeong et al., 2015; Jiang et al., 1999; Martella et al., 2007; Marthaler et al., 2013; Rahman et al., 2005; Stipp et al., 2015; Suzuki et al., 2014, 2015; Tsunemitsu et al., 1996). Attempts have also been made to extend this classification system to all 11 genes (Soma et al., 2013; Yamamoto et al., 2011), although this newly proposed all-gene-based genotyping system relied on very few sequences awaiting formal confirmation when larger numbers of sequence data will be available from a wide variety of RVC strains isolated from various host species.

In this study, we describe the detection and characterization of an RVC strain (denoted KE174/2012) in a stool sample collected from a 10-week-old diarrhoeic puppy in May 2012, using a viral metagenomics approach. As no sequence data for canine RVC strains was available in the databases, we attempted to determine the whole-genome sequence of the identified RVC strain. Our study demonstrated that this canine RVC strain was unique and highly distinct from human, porcine and bovine RVC strains.

The laboratory methods used for this study have been described elsewhere in detail (Mihalov-Kovács et al., 2015). In brief, 10 % faecal suspensions were prepared in PBS and centrifuged at 5000 × g for 10 min. Viral RNA was extracted using a Zymo Direct-zol kit (Zymo Research) combined with RiboZol RNA extraction reagent (Amresco), according to the protocol recommended by the manufacturer for biological liquids, although the DNase treatment was omitted from the workflow. The extracted nucleic acid specimen was subsequently denatured at 97 °C for 5 min in the presence of 10 μM random hexamers tailed by a common PCR primer sequence (Integrated DNA Technologies). Reverse transcription (RT) was performed with 1 U avian myeloblastosis virus reverse transcriptase (Promega), 400 μM dNTP mix and 1 × AMV RT buffer at 42 °C for 45 min, followed by a 5 min incubation at room temperature. Next, 5 μl cDNA from the RT reaction was added to 45 μl PCR mixture to obtain final concentrations of 500 μM PCR primers, 200 μM dNTP mix, 1.5 mM MgCl2, 1 × Taq DNA polymerase buffer and 0.5 U Taq DNA polymerase (Thermo Scientific). The PCR conditions consisted of an initial denaturation step at 95 °C for 3 min, followed by 40 cycles of amplification (95 °C for 30 s, 48 °C for 30 s and 72 °C for 2 min), with a final extension step at 72 °C for 8 min.

cDNA (100 ng) obtained by random PCR was subjected to enzymatic fragmentation and adaptor ligation (NEBNext Fast DNA Fragmentation & Library Prep Set for Ion Torrent kit; New England Biolabs). The barcoded adaptors were retrieved from the Ion Xpress Barcode Adapters (Life Technologies). The resulting cDNA libraries were measured on a Qubit 2.0 equipment using a Qubit dsDNA BR Assay kit (Invitrogen). An emulsion PCR was carried out according to the manufacturer’s protocol using an Ion PGM Template kit on a OneTouch v.2 instrument. Enrichment of the templated beads (on an Ion One Touch ES machine) and further steps of the pre-sequencing set-up were performed according to the 200 bp protocol of the manufacturer. The sequencing protocol recommended for the Ion PGM Sequencing kit on a 316 chip was strictly followed.

Processing of Ion Torrent data was carried out using the CLC Genomic Workbench v.7 (http://www.clcbio.com/). Further analysis consisted of the mapping of reads >40 bases against ~1.7 million viral sequences downloaded
from GenBank by applying moderately rigorous mapping parameters (length fraction 0.6; similarity fraction 0.8). To construct the genome sequence of the identified RVC strain, raw sequence reads were subjected to quality trimming to include sequence reads with lengths between 40 and 200 nt. Trimmed sequence reads were mapped onto reference RVC sequences obtained from GenBank. After visual inspection of the sequence alignments and remapping onto the obtained consensus sequence, a single contig was finalized for each genome segment. The genome sequence of KE174/2002 was deposited in GenBank. Multiple sequence alignments including RVC reference strains were constructed and manually adjusted with GeneDoc software (Nicholas et al., 1997). Phylogenetic analysis was conducted using the MEGA6 package (Tamura et al., 2013). Best-fit substitution models were selected for each dataset based on the Bayesian information criterion. Subsequently, maximum-likelihood trees were generated and bootstrap analysis was performed with 500 replications.

In the metagenomic assembly, 30 and 383 sequence reads were mapped onto a canine parvovirus and the RVC reference genomes. Subsequently, the almost full-length genome of the canine RVC strain was assembled from a total of 11,826 sequence reads obtained by independent sequencing runs. The 3’ ends of VP3 and VP4 genes and the 5’ ends of VP1 and VP3 genes could be partially determined. These genomic regions were determined by RNA ligation coupled with traditional sequencing methods (Mihalov-Kovács et al., 2015). For the obtained rotavirus genomic sequence, the coverage was extremely high (range, 65 × for VP6 and 194 × for NSP3).

The sizes of the 11 RNA segments, and the ORFs they contained, were compared with reference RVC strains from humans, pigs and cattle (Table 1). The length of the deduced amino acid sequences of the VP7 gene was slightly longer than that seen in the human, bovine and porcine reference RVC strains genomes, whereas VP6, NSP2 and NSP3 were of identical size compared with their counterparts in heterologous RVC strains. The ranges of nucleotide sequence identities between the canine RVC and the human, porcine and bovine RVC strains and relationship of these values in terms of genotype specificity are summarized in Table 2. In brief, the nucleotide sequence identities within any gene to the corresponding gene of any heterologous RVC ranged approximately from 67 to 84 %. The lowest sequence identities were seen in the NSP1 (range 67–78 %) and NSP4 (67–77 %) genes, whereas the highest similarities were seen in the NSP2 gene (79–84 %).

Using a set of human-, porcine- and bovine-origin RVC gene sequences, Soma et al. (2013) proposed a classification scheme assigning particular gene specific cut-off values to demarcate individual genotypes. This approach was similar to that established for RVA by Matthijnssens et al. (2008), but the number of RVC-derived genes used in the analysis was considerably lower and the sequences showed host species-specific bias towards human-origin RVC gene sequences. Based on the criteria in this system, the VP1–VP7 and NSP2 genes of the canine RVC strain could be assigned into novel genotypes, whereas the position of the NSP1 and NSP3–NSP5 genes remained uncertain. Different cut-off values were used for the VP7, VP4 and VP6 genes by independent research groups (Jeong et al., 2015; Marthaler et al., 2013; Suzuki et al., 2014, 2015), analyzing a greater number of sequences and realizing higher cut-off values for genotype demarcation. For these three genes, it seems to be safe to classify the canine RVC strain into novel genotypes, putatively: G10, P[8] and I8. Concerning the other genes, we felt it was more

### Table 1. Comparison of the coding potential of the canine RVC strain, KE174/2012, with reference porcine (Cowden), bovine (Shintoku) and human (Bristol) RVC strains

<table>
<thead>
<tr>
<th>Genome segment</th>
<th>RVC/Cowden</th>
<th>RVC/Shintoku</th>
<th>RVC/Bristol</th>
<th>Rotavirus C, KE174/2012</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Segment size (nt)</td>
<td>ORF (nt)</td>
<td>aa</td>
<td>Segment size (nt)</td>
</tr>
<tr>
<td>VP7</td>
<td>1063</td>
<td>999</td>
<td>332</td>
<td>1063</td>
</tr>
<tr>
<td>VP4</td>
<td>2246</td>
<td>2211</td>
<td>736</td>
<td>2253</td>
</tr>
<tr>
<td>VP6</td>
<td>1352</td>
<td>1188</td>
<td>395</td>
<td>1352</td>
</tr>
<tr>
<td>VP1</td>
<td>3290</td>
<td>3249</td>
<td>1082</td>
<td>3309</td>
</tr>
<tr>
<td>VP2</td>
<td>2736</td>
<td>2655</td>
<td>884</td>
<td>2727</td>
</tr>
<tr>
<td>VP3</td>
<td>2145</td>
<td>2079</td>
<td>692</td>
<td>2166</td>
</tr>
<tr>
<td>NSP1</td>
<td>1235</td>
<td>1182</td>
<td>393</td>
<td>1273</td>
</tr>
<tr>
<td>NSP2</td>
<td>995</td>
<td>939</td>
<td>312</td>
<td>1037</td>
</tr>
<tr>
<td>NSP3</td>
<td>1348</td>
<td>1209</td>
<td>402</td>
<td>1350</td>
</tr>
<tr>
<td>NSP4</td>
<td>613</td>
<td>453</td>
<td>150</td>
<td>Partial</td>
</tr>
<tr>
<td>NSP5</td>
<td>693</td>
<td>633</td>
<td>210</td>
<td>719</td>
</tr>
</tbody>
</table>
appropriate to wait until additional sequence data become publicly available to establish reliable and robust genotype demarcation rules. Using the genotype information obtained for VP7 and VP4 and following current rotavirus nomenclature scheme, the name of the canine RVC strain was formulated as RVC/Dog-wt/HUN/KE174/2012/G10P[8].

Deduced amino acid sequences revealed sequence divergence values ranging from 4.8 to 42.8 % when compared with cognate gene products of heterologous RVC strains (VP1, 7.2–10.5 %; VP2, 6.4–10 %; VP3, 14–18.8 %; VP4, 18.7–27.2 %; VP6, 4.8–9.4 %; VP7, 16.9–27.1 %; NSP1, 22.1–36.1 %; NSP2, 12.2–15.1 %; NSP3, 14.7–21.1 %; NSP4, 25.3–42.8 %; NSP5, 18.1–30.5 %; data not shown). Conserved protein regions and motifs described and analysed in detail in earlier studies were readily identified (Bremont et al., 1992; Fielding et al., 1994; James et al., 1999; Jiang et al., 1992; Luchs & Timenetsky, 2014; Marthaler et al., 2013; Soma et al., 2013; Suzuki et al., 2012; Tsunemitsu et al., 1996; Yamamoto et al., 2011) (data not shown). When analysing the deduced amino acid sequence of the major neutralization antigens in detail, VP7 of the novel canine RVC strain was found to be 1–5 aa longer than that seen in other RVC strains, and one major insertion was localized in the variable region 8 (VR8). Sequence alignment of the VP7 protein also revealed great divergence in the other variable regions (Tsunemitsu et al., 1996), and comparison of the concatenated sequence of the VR1–8 regions showed 29.4–48.6 % sequence divergence between canine and heterologous RVC strains, suggesting marked serological differences among RVC strains of heterologous hosts. Similarly low sequence similarities were found when the deduced VP4 amino acid sequences were compared. The canine RVC amino acid similarity values of the putative VP8* and VP5* fragments ranged from 28 to 39.7 % and from 13.8 to 22.6 %, respectively, with cognate protein regions of human, bovine and porcine RVC strains. This was consistent with the greater divergence observed in the VP8* region of VP4 of RVA strains (Hoshino & Kapikian, 1994).

Phylogenetic analyses of the 11 RNA segments were carried out to compare strain KE174/2012 with multiple human, bovine and porcine RVC strains (Fig. 1, Figs S1 and S2, available in the online Supplementary Material). As could be expected from the similarity calculations, the VP4, VP6 and VP7 genes of KE174/2012 clustered distinctly from other established genotypes. However, for all three trees, the canine RVC strain was most closely related to bovine RVC strains. As for the remaining eight genes, human, bovine and porcine strains formed clearly defined clusters. The Hungarian canine RVC strain formed distinct clusters, which were generally most closely related to the bovine subcluster. The only exception was in the NSP4 gene phylogenetic tree, where the canine RVC strain was closely related to RVC strains of porcine origin. These findings could suggest a past reassortment event between porcine- and bovine-origin RVC strains either before or during the diversification of the canine RVC into a new phylogenetic lineage.

In conclusion, our study demonstrated that the canine RVC detected by random primed RT-PCR and high-throughput sequencing is a genetically heterogeneous member of the species Rotavirus C, probably representing novel genotypes in most, if not all, genome segments. Studies on RVC strains so far have identified little variation

<table>
<thead>
<tr>
<th>Gene segment</th>
<th>Nucleotide sequence identity to heterologous RVC strains (%)</th>
<th>Cut-off values (%)</th>
<th>Putative genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Human</td>
<td>Porcine</td>
<td>Bovine</td>
</tr>
<tr>
<td>VP7</td>
<td>76.2–76.9</td>
<td>73.0–78.3</td>
<td>76.7</td>
</tr>
<tr>
<td>VP4</td>
<td>71.2–71.9</td>
<td>71.9–75.3</td>
<td>76.0–76.3</td>
</tr>
<tr>
<td>VP6</td>
<td>80.1–80.8</td>
<td>79.9–83.2</td>
<td>84.1–84.3</td>
</tr>
<tr>
<td>VP1</td>
<td>80.0–80.3</td>
<td>80.3</td>
<td>81.6–81.9</td>
</tr>
<tr>
<td>VP2</td>
<td>78.1–78.9</td>
<td>79.8</td>
<td>81.6–81.9</td>
</tr>
<tr>
<td>VP3</td>
<td>76.2–78.8</td>
<td>76.6</td>
<td>79.9–81.0</td>
</tr>
<tr>
<td>NSP1</td>
<td>67.2–68.1</td>
<td>74.5</td>
<td>76.5–78.3</td>
</tr>
<tr>
<td>NSP2</td>
<td>79.1–80.9</td>
<td>81.6</td>
<td>83.2–84.2</td>
</tr>
<tr>
<td>NSP3</td>
<td>78.3–79.7</td>
<td>78.7</td>
<td>81.4–82.2</td>
</tr>
<tr>
<td>NSP4</td>
<td>67.2–70.2</td>
<td>75.5–77.3</td>
<td>67.4–68.3</td>
</tr>
<tr>
<td>NSP5</td>
<td>76.6–77.1</td>
<td>80.9</td>
<td>80.2–82.9</td>
</tr>
</tbody>
</table>

Cut-off values assigning various genotype specificities were adapted from the following publications: *Marthaler et al. (2013); †Soma et al. (2013); ‡Jeong et al. (2015); §Suzuki et al. (2014).
in the genotype constellation per host species, although a marked diversity was seen in the neutralization antigens of porcine RVC strains (Jeong et al., 2015; Marthaler et al., 2013; Soma et al., 2013). Future studies may uncover the genetic diversity, if any, among canine RVC strains. With the addition of further genome sequences of RVC strains from heterologous host species, a more robust classification scheme can be expected to be constructed.
Acknowledgements

The study was supported by the Hungarian Academy of Sciences (Momentum program). SM and EF were supported by the Bolyai Scholarship Programme of the Hungarian Academy of Sciences.

References


