Diverse picornaviruses are prevalent among free-living and laboratory rats *(Rattus norvegicus)* in Hungary and can cause disseminated infections

Ákos Boros[a], Katalin Orlovácz[a], Péter Pankovics[a], Sándor Szekeres[b], Gábor Földvári[b'c], Elizabeth Fahsbender[d], Eric Delwart[d'e], Gábor Reuter[a'*]

ᵃ Department of Medical Microbiology and Immunology, Medical School, University of Pécs Pécs, Hungary ᵇ Department of Parasitology and Zoology, University of Veterinary Medicine, Budapest, Hungary ᶜ Evolutionary Systems Research Group MTA Centre for Ecological Research, Tihany, Hungary ᵈ Vitalant Research Institute, San Francisco, CA, USA ᵉ University of California, San Francisco, CA, USA

* Corresponding author at: Department of Medical Microbiology and Immunology, University of Pécs, Szigeti út 12, Pécs H-7624, Hungary. *E-mail address:* reuter.gabor@gmail.com (G. Reuter).

A B S T R A C T

In this study, the full length genomes of three phylogenetically distant picornaviruses (family *Picornaviridae)* belonging to the genus *Rosavirus* (rat08/rRoB/HUN, MN116648), *Kobuvirus* (rat08/rAiA/HUN, MN116647), and *Cardiovirus* (rat08/rCaB/HUN, MN116646) were obtained from a single faecal sample of a free-living Norway rat *(Rattus norvegicus)* in Hungary using viral metagenomics and RT-PCR/Sanger sequencing. The acquired complete genomes were in silico analyzed in detail revealing the presence of a second minor open reading frame encoding an alternative Leader peptide (L*) in rat08/rCaB/HUN and a ca. 222 nt-long sequence repeat with compact secondary RNA structure in the 3' UTR of rat08/rRoB/HUN. The studied rat picornaviruses were frequently detectable by RT-PCR with relatively high viral loads ranged between $8.99E + 02$ and $1.29E + 06$ copies/ml in rat faecal samples collected from five geographically distant locations throughout Hungary. The VP1 sequence-based phylogenetic analyses show the presence of multiple, mostly location-specific lineages for all three picornaviruses. Rat rosavirus and rat cardiovirus were identified in spleen while rat cardiovirus was also detected in liver, muscle and kidney samples with variable copy numbers ($6.42E + 01$-$1.90E + 05$ copies/tg total RNA) suggesting extra-intestinal dissemination. Both viruses were also prevalent (70.0% and 18.2%)

among two populations of laboratory rats ("Wistar-type" and "hooded-type") held in different, isolated laboratory animal units.

1. Introduction

Members of family *Picornaviridae* have positive sense, single-stranded RNA genomes ranging between 7.2 and 9.7 kb (Palmenberg et al., 2010). Picornaviruses (PVs) generally have a single major open reading frame (ORF) flanked by 5' and 3' untranslated regions (UTRs), although there are certain PVs such as TO subtypes of Theiler's encephalomyelitis virus (TMEV) of genus *Cardiovirus* where an additional minor ORF encodes an alternative Leader (L) peptide was also present (Palmenberg et al., 2010; Takano-Maruyama et al., 2006). The genome layouts of PVs include the 5'UTR- ORF: [L (not always present)-VP0 (could be cleaved into VP4 and VP2)-VP3-VP1-2A-2B-2C-3A-3B-3C-3D]-3'UTR-polyA-tail (Palmenberg et al., 2010). The highly structured 5' UTRs contain predominantly one of the five types of internal ribo-somal entry sites (IRES) while the 3'UTRs could contain conserved RNA motifs like s2 m, "barbell-like" structures or multiple repeated sequences (Palmenberg et al., 2010; Kofstad and Jonassen, 2011; Boros et al., 2014). The family *Picornaviridae* currently consists of 110 species grouped into 47 genera and several unassigned PVs (Zell et al., 2017; www.picornaviridae.com). Due to the species richness (ca. 2200 known rodent species) and close contact to humans rodents have one of the highest numbers of discovered viruses including PVs (Chen et al., 2017). The majority of the currently known rodent PVs belong to 11 of the 47 picornavirus genera including *Cardiovirus, Kobuvirus* and the recently established *Rosavirus* (Chen et al., 2017; www.picornaviridae. com).

The novel genus *Rosavirus* currently consist of three species *(Rosavirus A-C)* mainly of rat and mouse viruses (Phan et al., 2011; Lau et al., 2016). Novel Rosavirus A2 (species *Rosavirus A)* sequences were also identified from pediatric diarrheic faecal specimens in Gambia (Lim et al., 2014). The genus *Cardiovirus* currently contains several well-known PVs including the Encephalomyocarditis virus (EMCV, *Cardiovirus* A), which can infect a wide range of animals, the TMEV of mice, the Vilyuisk human encephalomyelitis virus (VHEV) and human Saffold viruses of species *Cardiovirus B,* the baboon cardioviruses of *Cardiovirus C* and several currently unpublished cardiovirus sequences from different rat species (Zell et al., 2017; www.picornaviridae.com). From the six official kobuvirus species *(Aichivirus A-F)* only the *Aichi-virus A* contains rodent viruses (Murine kobuvirus-1, MuKV-1) together with different genotypes identified from humans, cats, dogs and birds as well (Zell et al., 2017). Currently, only partial VP1 sequences and a single unpublished complete genome of MuKV-1-related rat kobuvirus (rat/Wencheng-Rt386-2/China/2012, MF352432) has been known (Firth et al., 2014; www.picornaviridae.com). With a few exceptions (the pathogenesis of Rosavirus C or the genomic

diversity of rodent kobuviruses), the prevalence, genomic diversity and extraintestinal presence or viral load of these viruses in rats are not investigated before (Lau et al., 2016; Lu et al., 2018).

Rodents, including different types of rats such as black rat *(Rattus rattus)* or brown rat/Norway rat *(Rattus norvegicus),* may serve as potential reservoirs of various known and novel pathogens including (picorna)viruses and due to the overlapping habitats and frequent interactions with humans these animals could be major sources of zoonotic infections (Chen et al., 2017; Strand and Lundkvist, 2019). Furthermore, besides the free-living rodent populations, various types of Norway rats (e.g. "Wistar-type" or "hooded-type") are extensively used as experimental animal models for diverse scientific research projects (Baker et al., 2013). Although separate populations of these animals are usually bred and held in special laboratory animal facilities/units with high hygienic standards little information is available about the viruses circulating in such experimental animal populations.

In this study, the complete genomes of three phylogenetically distant PVs belonging to the genus *Cardiovirus, Kobuvirus* and *Aichivirus* were characterized by viral metagenomic and RT-PCR from a single faecal sample of a free-living Norway rat in Hungary and analyzed in detail. The prevalence, genomic diversity of these PVs and virus RNA detection (RT-PCR/qPCR) from extra-intestinal tissues (including absolute copy numbers) of these PVs were also investigated. The study viruses were prevalent among free-living animals originated from geographically distant locations in Hungary and in a subset of laboratory rats ("Wistar-type" and "hooded-type") held in isolated laboratory animal units.

2. Materials and methods

2.1. Background data of the sampled animals

Total of 10 (N = 3 juvenile, *N* = 7 adult, animal IDs: rat-01 - rat-10) free-living Norway rats *(Rattus norvegicus)* with unknown health status were trapped by mechanic rat-traps from backyards of small livestock farms from five different geographical locations in Hungary between July 2017 and April 2018 (Fig. 1A). Faecal samples from 9 of the 10 trapped animals as well as various tissue specimens from all carcasses were collected for further analyses (Table 1).

An additional 21 faecal samples (N = 10 adult "Wistar-type", *N* =11 "hooded-type") were also collected from Norway rats bred and held in two separate laboratory animal units in May 2019 in University of Pécs, Pécs, Hungary (Fig. 1A). All of the apparently healthy sampled animals were held in separate cages. All faecal and tissue specimens were stored in -80 °C until further processing.

2.2. Viral metagenomics, next-generation sequencing and bioinformatics analyses

A single pooled specimen which contained four randomly selected faecal samples of free-living Norway rats (rat-01; -04; -08; -09) collected from geographically distant locations in Hungary (Budapest,

Hajdúböszörmény, Rácalmás and Zimány) (Fig. 1A) were subjected to viral metagenomic analysis using random (RT-)PCR amplification of viral-particle protected nucleic acids as previously described (Phan et al., 2013b). Briefly, 200 ml of the supernatant of centrifuged (15,000 g for 10 min) faecal suspension was filtered through a 0.45-mm filter (Millipore) and then treated with a mixture of DNases (Turbo DNase from Ambion, Baseline-ZERO from Epicentre, and Benzonase from Novagen) and RNase (Fermentas) to digest unprotected nucleic acids. Viral cDNA library was constructed by ScriptSeqTM v2 RNA-Seq Library Preparation Kit (Epicentre) and sequenced using the MiSeq Illumina platform. The generated sequence reads (singletons) were assembled to contigs and sequences (reads and contigs) longer than 100-bp were compared to the protein database of GenBank using BLASTx. The Expected value (E) cut-off was set up to 10^{-5} for the categorization of the reads. Only sequences with better matches to known viruses were analyzed further.

2.3.   Genome acquisition and RT-PCR screening reactions

The total RNA was extracted from 150 ul of faecal suspensions (35-40 v/v% in sterile 0.1 M PBS) or 50-100 mg of tissue samples using TRI Reagent (MRC, USA) by the manufacturer's instructions. Prior to RNA extraction, the tissue samples were measured and homogenized mechanically using manual Potter-Elvehjem tissue grinder (Sigma, Germany). The amount of extracted total RNA was measured using NanoDrop 2000 spectrophotometer (Thermo Fisher, Waltham, MA, USA). The reaction conditions and reagents applied in the RT and PCR reactions of screening or genome acquisitions were the same as described previously with some modifications (Boros et al., 2011). Briefly, in the RT reactions, 1 ug of total RNA was applied in a final volume of 25 ul and the all RT volume was used in the PCR round in a final volume of 50 ul in a PCR thermal program contains 39 cycles. For complete genome acquisition reactions, the sequence reads and contigs with the highest match to different picornaviruses are used as targets for virus-specific oligonucleotide primer design. The designed primers were used to amplify the genomic regions not covered (or poorly covered) by metagenomic sequences including the 5' and 3' ends in different RT-PCR and 5'/3' rapid amplification of cDNA ends (RACE) reactions (Boros et al., 2011, 2012). The primer pairs used for the complete genome determinations were available on request. For screening reactions, picornavirus-specific primer pairs designed for either the conserved 3D RNA-dependent RNA polymerase ($3D^{RdRp}$) or the VP3-VP1 capsid regions of the study viruses were used (Table S1). The generated PCR-products were sequenced directly in both directions using BigDye Terminator v1.1 Cycle Sequencing Kit (Thermo Fisher) and run on an automated sequencer ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Stafford, Texas, USA).

### 2.4.   Absolute quantification of viral copy numbers using RT-qPCR

For the absolute quantification of viral RNA in tissue samples 1 ug, and in case of faecal samples 1 ul of total RNA was transcribed using random hexamer (500 ng/reaction), oligo dT (250 ng/reaction) primers and M-MLV-RT enzyme (Promega, USA) in a final volume of 10 ul. For quantification PCR reactions 1-1 ul of RT products and picornavirus-specific qPCR primer pairs (Table S1) designed for the conserved $3D^{RdRp}$ regions of the study viruses were used in separate SYBR Green-based real-time PCR assays (Maxima SYBR Green qPCR Master Mix, Thermo Scientific, USA) (Table S1). The quantification steps of the qPCR reactions were always followed by a dissociation assay and both run in on a Rotor-Gene Q MDx 5plex thermal cycler (Qiagen, Hilden, Germany). For the generation of assay-specific standard curves ten-fold dilution series of purified and spectrophotometrically quantified study picornavirus $3D^{RdRp}$ RT-PCR products were used (Table S1). All of the qPCR assays contained three technical repeats of every sample and standard. The slopes of the standard curves were -3.60486 (rCaB-qPCR), -3.67803 (rRoB-qPCR) and -3.54215 (rAiA-qPCR). The calculated PCR-efficiencies were 99.818% (rCaB-qPCR), 99.729% (rRoB-qPCR) and 99.882% (rAiA-qPCR). The heat map generated from the absolute viral copy numbers measured in the faecal samples of free-living rats was created by the Heatmapper software (Babicki et al., 2016).
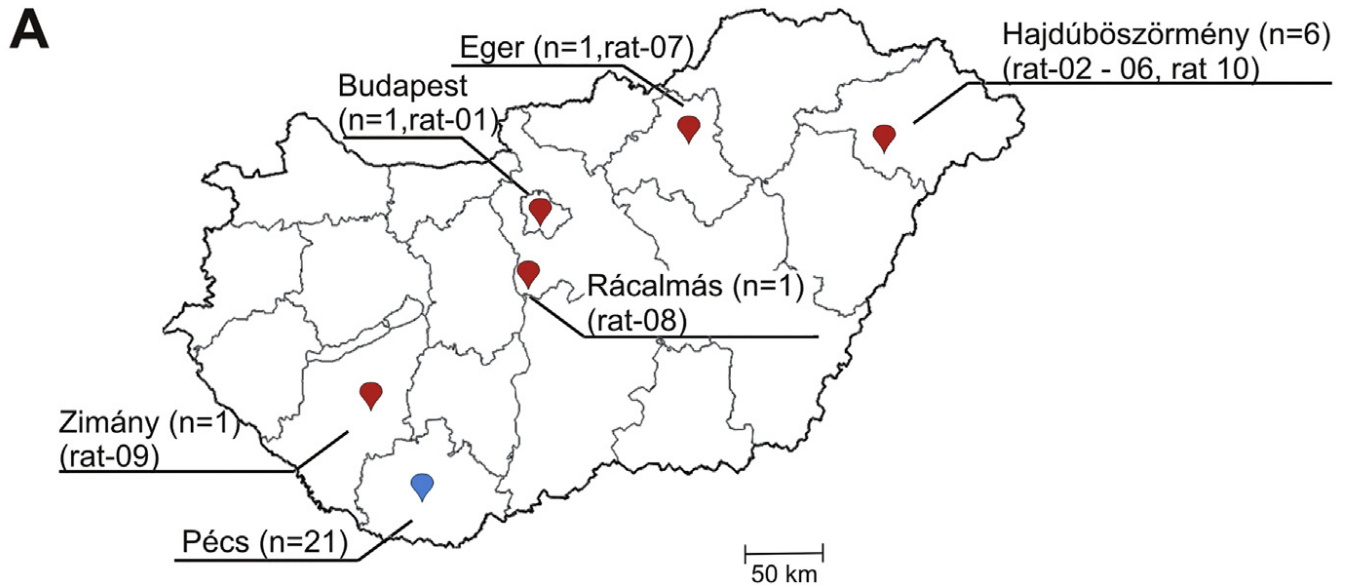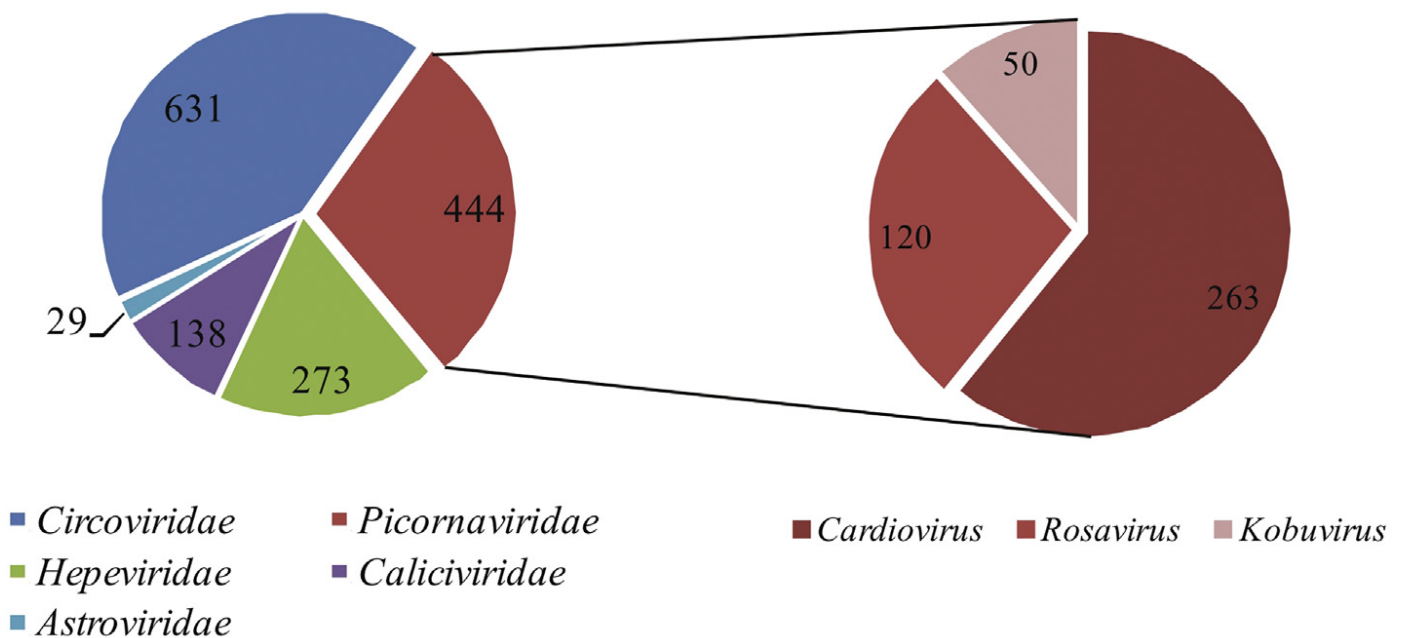
Fig. 1. (A) Geographical locations of free-living (red markers) and laboratory-types ("Wistar" and "hooded") of Norway rats (blue marker) in the map of Hungary with numbers of collected faecal samples and animal IDs in parenthesis. **(B)** Numbers of individual metagenomic reads belong to various vertebrate virus families (left side) as well as the total reads belong to different genera of family *Picornaviridae* (right side) based on the best BLASTx-scores (E-value $<10^{-5}$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.5. Sequence and phylogenetic analyses

The picornavirus sequences were aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE) (https://www.ebi.ac.uk/ Tools/msa/muscle/). The pairwise nucleotide (nt) and amino acid (aa) identity calculations of the aligned sequences were performed using GeneDoc v. 2.7 (Nicholas and Nicholas, 1997). Phylogenetic trees were constructed using MEGA software ver. 7.026 (Kumar et al., 2016). The Neighbor-Joining model with the number of differences method was used to construct the 3CD amino acid tree, while the Jukes-Cantor model was used to create the phylogenetic tree of VP1 nucleotide sequences. Bootstrap values were set up to 1000 replicates and > 50% were indicated in all phylogenetic trees. The possible proteolytic cleavage sites were predicted primarily based on the aa alignments of the closest relatives. The secondary RNA structures of the 5' and 3' UTR were predicted by the Mfold program (Zuker, 2003) and the 2D models were visualized and annotated using VARNA and the CorelDraw Graphics Suite v. 12 (Darty et al., 2009).

The VP1 sequences (MN116627-MN116645) and complete genomes of rat08/rRoB/HUN (MN116648), rat08/rAiA/HUN, (MN116647) and rat08/rCaB/HUN (MN116646) were deposited to the GenBank database.

3. Results

3.1. Metagenomic overview

A single pooled specimen which contained faecal samples of four randomly selected rats (rat-01; -04; -08; -09) collected from geographically distant locations in Hungary (Budapest, Hajdúböszörmény, Rácalmás and Zimány) was subjected to viral metagenomics investigation (Fig. 1A). A total of 19,181 viral reads were identified by bioinformatic pipeline of which 12,396 reads were likely originated from viruses infecting bacteria *(Microviridae: N =* 5663, *Picobirnavir-idae: N =* 1092 and *Siphoviridae: N =* 632...), plants *(Phycodnaviridae: N =* 46, *Secoviridae: N =*17....) and various unassigned viruses *(N =* 3842). The remaining 6785 viral reads were related to various invertebrate *(Nodaviridae: N =* 89, *Iflaviridae: N =* 51, *Dicistroviridae: N =*25...) and vertebrate virus families including sequences of family *Picornaviridae* (Fig. 1B). The detailed BLAST analyses of the picorna-virus-related sequences suggested the presence of at least three different viruses related to the members of genera *Rosavirus, Kobuvirus* and *Cardiovirus* (Fig. 1B). This study is focusing on complete genome characterization, detailed genome analysis, extraintestinal presence and small scale epidemiological investigation of these three PVs.

In order to find the exact faecal sample(s) from the pool which originally contained these PVs individual picornavirus read-specific primer-pairs and RT-PCR screening reactions were used. All the three PVs were

detectable from a single faecal sample (rat-08), therefore this sample was used for RT-PCR-based complete genome determination reactions.

Table 1

Results of screening RT-PCR and RT-qPCR (in brackets) reactions of various tissue and faecal samples of free-living Norway rats (rat 01 -10) using generic primerpairs targeting the conserved 3D^{RdRp} genome regions of the three different picornaviruses (PV column). rCaB: rat cardiovirus B, rAiA: rat Aichivirus A, rRoB: rat rosavirus B. +: RT-PCR-positive sample, —: RT-PCR-negative sample, n.a.: no available sample, qPCR —: negative sample by RT-qPCR. Note that in the case of the RT-qPCR positive tissue samples the absolute viral copy numbers/tg total RNA are given.

Sample type	PV	rat 01	rat 02	rat 03	rat 04	rat 05	rat 06	rat 07	rat 08	rat 09	rat 10
Muscle	rCaB	–	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	+ (6.42E+01)	n.a.	n.a.
	rAiA	–	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.	n.a.
	rRoB	–	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.	n.a.
Brain	rCaB	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.	n.a.
	rAiA	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.	n.a.
	rRoB	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.	n.a.
Liver	rCaB	–	–	+ (qPCR –)	–	–	–	–	+ (1.19E+02)	–	–
	rAiA	–	–	–	–	–	–	–	–	–	–
	rRoB	–	–	–	–	–	–	–	–	–	–
Spleen	rCaB	–	–	n.a.	–	n.a.	n.a.	–	+ (1.74E+02)	–	+ (5.77E+01)
	rAiA	–	–	n.a.	–	n.a.	n.a.	–	–	–	–
	rRoB	–	–	n.a.	–	n.a.	n.a.	–	–	–	+ (1.90E+05)
Kidney	rCaB	–	–	n.a.	–	n.a.	–	–	+ (qPCR –)	–	–
	rAiA	–	–	n.a.	–	n.a.	–	–	–	–	–
	rRoB	–	–	n.a.	–	n.a.	–	–	–	–	–
Lung	rCaB	–	–	n.a.	n.a.	n.a.	n.a.	–	–	–	–
	rAiA	–	–	n.a.	n.a.	n.a.	n.a.	–	–	–	–
	rRoB	–	–	n.a.	n.a.	n.a.	n.a.	–	–	–	–
Heart	rCaB	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	–	n.a.
	rAiA	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	–	n.a.
	rRoB	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	–	n.a.
Omasum	rCaB	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.
	rAiA	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.
	rRoB	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	–	n.a.
Faeces	rCaB	–	–	–	–	+	+	–	+	–	n.a.
	rAiA	–	+	–	–	+	+	–	+	+	n.a.
	rRoB	–	–	+	+	+	+	+	+	+	n.a.

3.2. Genome analysis of rat Rosavirus B

The rosavirus-related sequence reads *(n = 120)* could be aligned with the best coverage to the strain RNCW0602091R (KX783423) of species *Rosavirus B* (RoB) genus *Rosavirus* identified recently from Norway rats *(Rattus norvegicus)* in China (Lau et al., 2016). The aligned reads covered =42% of the 8923-nt-long genome of strain RNCW0602091R with variable depth of coverage up to 21 (Fig. 2A).

The acquired 8731-nt-long (excluding the poly(A)-tail) complete genome of strain rat08/rRoB/HUN (MN116648) is 192 nt shorter than the strain RNCW0602091R (KX783423) as the closest match identified by BLASTn search. The two virus sequences overall show 83% nt and 92% aa identities. The 566 nt-long 5'UTR contains a 61-nt long stretch at the 5' end predicted to form a single stem-loop (data not shown),

which is missing from the 5'UTR of strain RNCW0602091R as well as from the other currently known RoB sequences (KX783421-23). The 5'UTRs of the study virus and the other RoB (excluding the 61 nt stretch) shows 90-92% nt pairwise identities. The 5'UTR of rat08/rRoB/ HUN is predicted to contain a type II-like internal ribosomal entry site (IRES) (data not shown) similar to that found in the other rosaviruses (Phan et al., 2013a; Lau et al., 2016).

The 7470 nt-long single open reading frame (ORF) starts with a start codon found in an optimal Kozak context (AuuA**567**UGG, start codon is underlined, conserved nts are capitalized). The single ORF encodes a 2489-aa-long viral polyprotein which shares the same organization of 4-3-4, and contains highly similar conserved aa motifs of 2A$^{Hbox/NC}$,2C helicase (2C^{Hel}), 3C protease (3C^{Pro}) and 3D^{RdRp} and similar predicted cleavage sites as the other rosaviruses (Phan et al., 2013a; Lau et al., 2016)(Fig. 2A). The aa sequence identities were ranged between 86% (2A) and 97% (3CD) of the study and RNCW0602091R strains, which also indicates a close sequence relationship between these viruses (Fig. 2A).

The 695-nt-long 3'UTR study strain of rat08/rRoB/HUN possess a 248 nt deletion compared to the 943-nt-long 3'UTR of strain RNCW0602091R (Fig. 3). The 3'UTRs of rosaviruses contain a c.a. 222 nt-long repeated sequence ("Rep") based on our manual sequence analysis (Fig. 3). The reference strain RNCW0602091R has three Reps and the sequences have 74-91% nt pairwise identities and each repeat contains the conserved "q motif" identified in the 3'UTRs of rosa-, and cadiciviruses (Lim et al., 2014)(Fig. 3). The study strain rat08/rRoB/ HUN also contains copies of the c.a. 222-nt-long repeated sequences (Reps), but interestingly the shorter 3'UTR sequence of the study strain has only two Reps with high (94-95% nt) identities to the corresponding regions of the reference strain (Fig. 3). The secondary structure analysis of the "Rep" sequences revealed essentially similar, conserved RNA structure with multiple stem-loops, bulges and internal loops (Fig. 3).

The rat08/rRoB/HUN clustered together with the other rosaviruses in the 3D^{RdRp} phylogenetic tree where the sequences of *Rosavirus B* species are the closest relatives (Fig. 4).
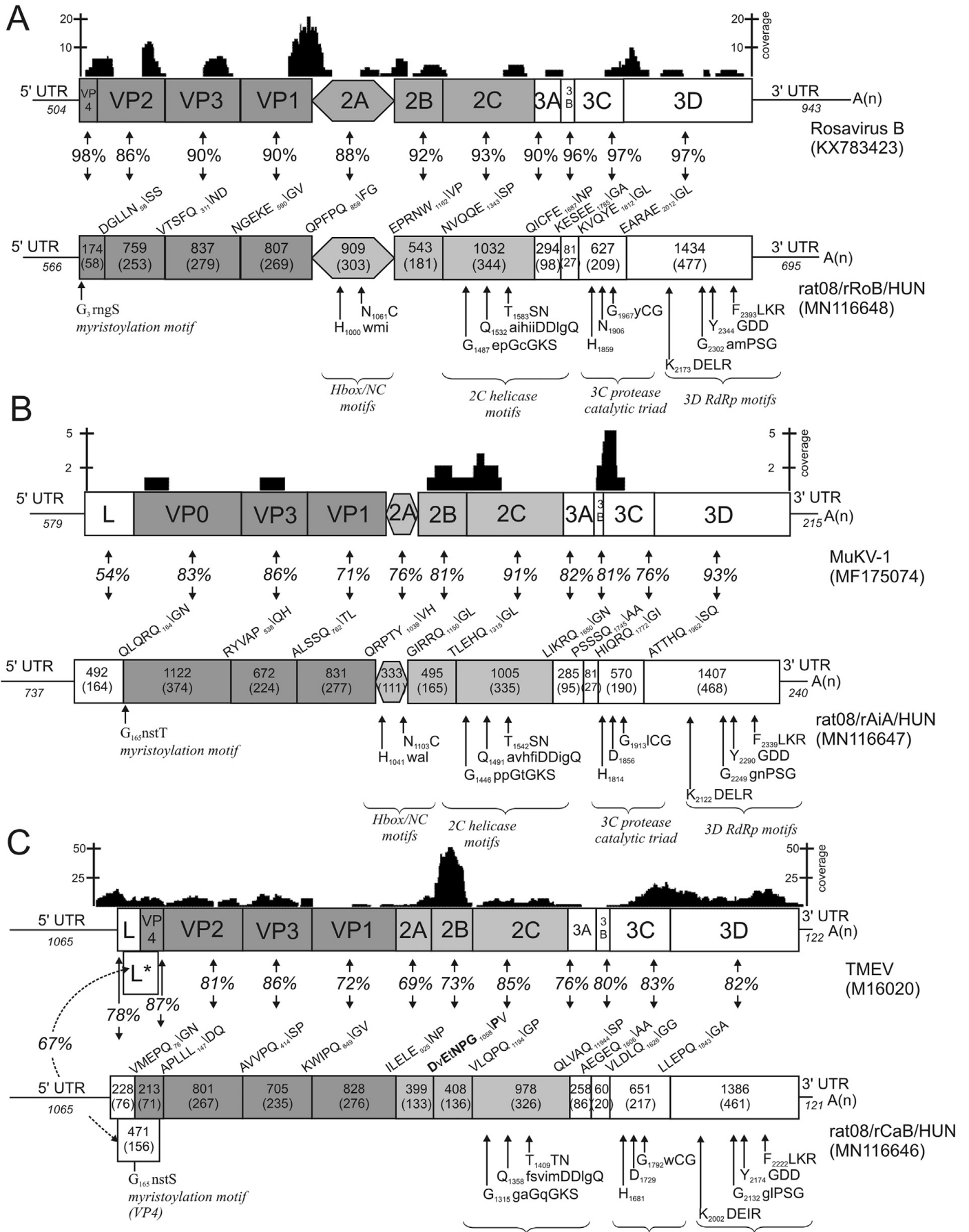
Fig. 2. Genome maps with conserved amino acid motifs and presumed cleavage sites (P5\P2′) of the study picornaviruses (bottom side of the panels) and their closest relatives (top side). The gene boxes corresponding to the P1 (viral capsid proteins) and P2 (non-structural proteins) are depicted with different

shades of grey. The nucleotide (upper numbers) and amino acid (lower numbers in brackets) lengths of the corresponding genomic regions of the study viruses are shown in each gene box. The coverage graphs with the locations of metagenomic reads above the genome maps of each virus were generated by the UGENE software (Okonechnikov et al., 2012). The positions and sequences conserved amino acid motifs of 2C Helicase (Gorbalenya et al., 1990), 3C Protease (Gorbalenya et al., 1989), and 3D RdRp (RNA-dependent RNA polymerase; Argos et al., 1984) are shown under each map of the study viruses. Conserved and variable amino acids are represented by uppercase and lowercase letters respectively. The pairwise amino acid identity values (%) of each genomic region are found between the genome maps.
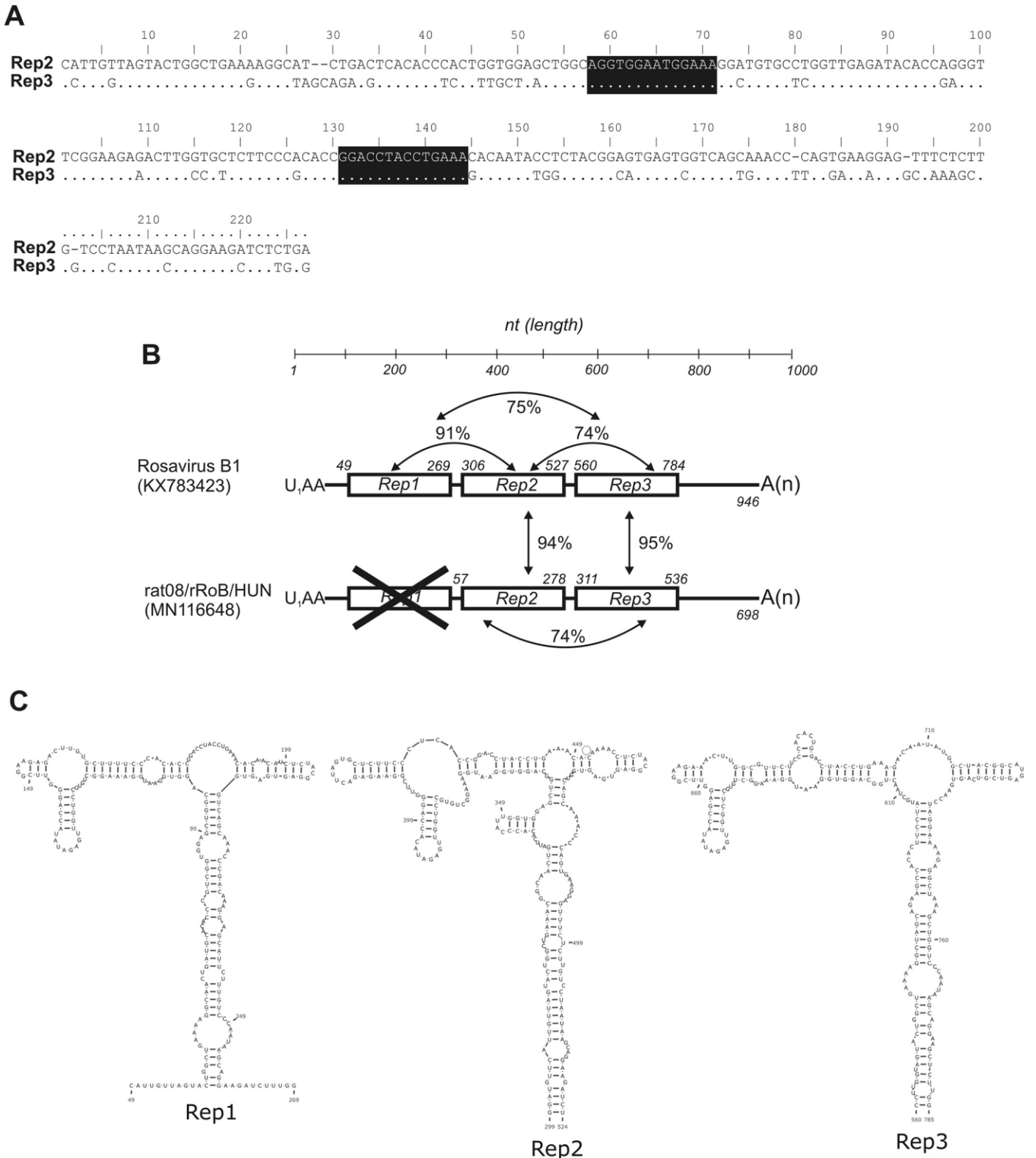
Fig. 3. (A) Nucleotide sequence alignment of the sequence repeats ("Rep2 and Rep3") of the 3′UTR of rat08/rRoB/HUN (MN116648). The localization in the alignment of conserved "q motif" identified in the 3′UTRs of rosa-, and cadiciviruses are marked with black background (Lim et al., 2014). (B) Schematic organization scheme of the 3′UTRs of RoB strain RNCW0602091R (KX783423, top) and of rat08/rRoB/HUN (bottom) including the pairwise sequence identities (%) and localizations of repeated sequences ("Rep1–3"). Crossed lines indicate the absence of "Rep1" in the study sequence of rat08/rRoB/HUN. (C) The predicted secondary RNA structure of the ca. 222 nt-long 3′UTR sequence

repeats ("Rep1-3 of Rosavirus B strain RNCW0602091R (KX783423). Note that the secondary RNA structure of "Rep-2" and "Rep-3" of rat08/rRoB/HUN is essentially the same as found in Rosavirus B strain RNCW0602091R.

3.3. Genome analysis of rat Aichivirus A

The sequence reads of a kobuvirus-related picornavirus *(N =* 50) could be aligned with the best coverage to the currently unpublished strain Wencheng-Rt386-2 (MF352432) of species *Aichivirus A* (AiA) in genus *Kobuvirus* identified from a faecal sample of an Asian house rat *(Rattus tanezumi)* in China. The aligned reads covered only = 19% of the 8269-nucleotide (nt)-long genome of Wencheng-Rt386-2 with low depth of coverage (Fig. 2B).

The determined 8268-nt-long (excluding the polyA-tail) complete genome of strain rat08/rAiA/HUN (MN116647) shows the highest sequence identity (92% nt) to AiA strain Wencheng-Rt386-2 (MF352432) as the closest match identified by BLASTn search.

The 735-nt-long 5'UTR of strain rat08/rAiA/HUN shows the highest (94% and 82%) nt identities to the corresponding genome parts of AiA strain Wencheng-Rt386-2 (MF352432) and human Aichivirus A1 strain A846/88 (AB010145). Secondary RNA structure analysis of the 5'UTR of strain rat08/rAiA/HUN revealed the presence of 5'terminal conserved hairpins (hairpin A-C) of *Ori* region and a type V-like IRES and (data not shown) similar as found in the members of species *Aichivirus A* (Sweeney et al., 2012).

The start codon of the 7293 nt-long single ORF located in an optimal Kozak context (AucA$_{736}$UGG, start codon is underlined, conserved nts are capitalized). The 2430-aa-long viral polyprotein of the study strain shows an overall 98% aa identity to a currently unpublished sequence AiA strain Wencheng-Rt386-2. Hence, further analysis of the viral polypeptide of strain rat08/rAiA/HUN was carried out the comparison of the previously described and analyzed Murine kobuvirus 1 (MuKV-1) isolate MKV1/NYC/2014/M014/0146 (MF175074) of species *Aichivirus A* (Williams et al., 2018) which is also closely related to the study strain. Based on the predicted cleavage sites the strain rat08/rAiA/HUN has a typical kobuvirus genome organization pattern of L-3-3-4 (Fig. 2.). The aa sequence identities were ranged between 54% (L) and 93% (3D) of the study strain and MuKV-1, which suggest a close sequence relationship between these viruses (Fig. 2). Based on the presence of conserved aa motifs the N-terminal part of VP0 could be myristoylated and 2A peptide could belong to the Hbox/NC peptide group (Hughes and Stanway, 2000)(Fig. 2). The functional groups of 2C^{Hel},3C^{Pro} and 3D^{RdRp} can also be clearly identifiable in the genome of rat08/rAiA/HUN (Fig. 2B).

The 240-nt-long 3' UTR of rat08/rAiA/HUN shows the highest sequence identity (95%) to AiA strain Wencheng-Rt386-2 and both sequences contain the characteristic "barbell-like" structure with conserved

motifs of GAUAUAAAG and CCCUAA (Fig. S1) which is also identifiable in several different PVs including other kobuviruses as well (Boros et al., 2014).

In the 3D^RdRp phylogenetic tree, rat08/rAiA/HUN was located at the same branch as the members of genus *Kobuvirus,* where phylogeneti-cally most closely related to the members of *AichivirusA* species (Fig. 4).
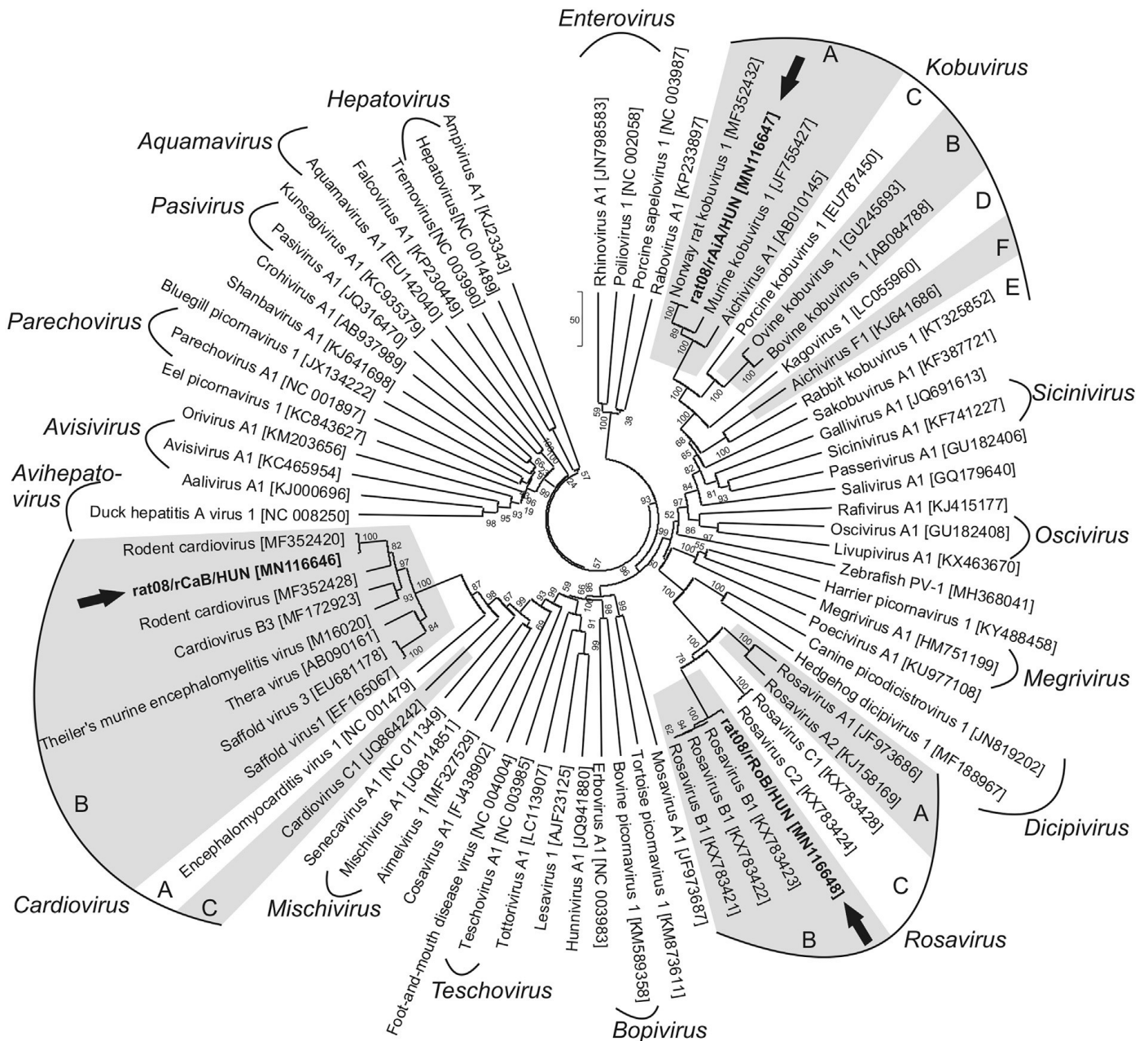

Fig. 4. Phylogenetic analysis of the study picornaviruses (in bold and indicated with black arrows). The Neighbor-Joining tree was constructed from the alignment of 3CD (helicase and polymerase) amino acid sequences of the study viruses as well as the representative members of family Picornaviridae. Representative members of different species of genus Rosavirus (A-C), Kobuvirus (A-F) and Cardiovirus (A-C) are marked with letters and with grey or white backgrounds). Representative picornavirus genera are shown with italics to help the orientation.

3.4. *Genome analysis of rat Cardiovirus B*

The remaining picornavirus reads (N = 263) showed the closest sequence relationship to a currently unpublished sequence of rodent cardiovirus strain Ruian-Rn93-1 (MF352420) which was identified from a faecal sample of a Norway rat *(Rattus norvegicus)* in China. The aligned reads covered 78% of the 7927-nt-long genome of Ruian-Rn93-1 with variable depth of coverage ranged between 1 and 52 (Fig. 2C).

The determined 8101-nt-long (excluding the polyA-tail) complete genome of rat08/rCaB/HUN (MN116646) is 174 nt longer than the genome of Ruian-Rn93-1 as the closest relative identified by BLASTn and the two sequences show overall 88% nt identity.

The 1065-nt-long 5' UTR of strain rat08/rCaB/HUN contains a 171 nt long stretch at the 5' end which is missing from the 5'UTR of Ruian-Rn93-1. Instead of a polyC tract identified in the 5'UTR of Theiler's encephalomyelitis virus (TMEV; Pevear et al., 1987) of species *Cardiovirus B* (CaB) of genus *Cardiovirus* a polyGC tract was present in the 5' terminus (5'-G_8GGGGCCCGGGGGG-3') of rat08/rCaB/HUN. The two 5'UTRs (excluding the 171 nt stretch) show 94% nt identity. The 5'UTR of rat08/rCaB/HUN predicted to contain a type II-like IRES (data not shown) similar as found in the other cardioviruses (Duke et al., 1992).

The presumed start codon of the 6915 nt-long major ORF can be found in an optimal Kozak context (AcgA$_{1066}$UGG, start codon is underlined, conserved nts are capitalized). Interestingly, a second minor ORF with an alternative translation initiation site (AacA$_{1079}$UGG) was found in a different reading frame that encodes a 156 aa long peptide with unknown function (Fig. 2C). This peptide shows the highest (78%) sequence identity to the L* protein of Theiler's-like virus of rats (BAC58036) by BLASTp. The 2304-aa-long viral polyprotein encoded by the major ORF shows an overall 95% aa identity to Ruian-Rn93-1. Because Ruian-Rn93-1 and closely related rat cardioviruses of Long-quan-Aa3-1 (MF352428), Ruian-Rn93-3 (MF352411) are currently unpublished further analysis of the viral polypeptide of strain rat08/ rCaB/HUN was based on the comparison of the previously described and analyzed Theiler's encephalomyelitis virus (TMEV, M16020) of species *Cardiovirus B* (Pevear et al., 1987) which also closely related to the study sequence.

The rat08/rCaB/HUN shows genome similar architecture of L-(L*)-4-3-4 as the TMEV. The aa sequence identities were ranged between 67% (L*) and 86% (VP3) of the study strain and TMEV (Fig. 2C). The 2A of rat08/rCaB/HUN belongs to the DxExNPG|P (x = variable aa, conserved aas are capitalized; |= cleavage site) motif-containing aphthovirus-like 2A group (Fig. 2). The presumed cleavage sites as well as the conserved aa motifs of 2C^{Hel},3C^{Pro} and 3D^{RdRp} can be found in Fig. 2C. The 125-nt-long 3' UTR (excluding the polyA-tail) of rat08/rCaB/HUN shows the highest sequence identity (98%) to Ruian-Rn93-1 and none of the previously identified conserved motifs like "barbell-like" structure (Boros et al., 2014)or "s2 m" (Kofstad and Jonassen, 2011)is detectable.

The rat08/rCaB/HUN was clustered together with the other cardi-oviruses in the $3D^{RdRp}$ phylogenetic tree where the sequences of rodent cardioviruses of *Cardiovirus B* species are the closest relatives (Fig. 4).

3.5. *Prevalence and faecal shedding of rat picornaviruses*

To investigate the epidemiological distribution of the study PVs, among different rat populations in Hungary beside rat-08 additional faecal samples of trapped free-living (N = 8 from five geographically distant locations) and laboratory-bread (N =10 "Wistar-type and N = 11" hooded-type") Norway rats from two separate laboratory animal units were tested. For the screening experiments, separate RT-PCR reactions and generic primer-pairs targeting the conserved $3D^{RdRp}$ genome regions of the three different PVs (Fig. 1A, Table 1 and Table S1) were used. The absolute copy numbers of the viruses from the RT-PCR positive faecal samples were also determined using SYBR-Green-based RT-qPCR method (Fig. 5B). Rat Rosavirus B (rRoB), rat Aichivirus A (rAiA) and rat Cardiovirus B (rCaB) were detectable in 7/9 (77.7%), 5/9 (55.5%) and 3/9 (37.5%) of the analyzed faecal samples of trapped wild animals (Fig. 5A). The absolute copy numbers of rCaB and rRoB were measured in the same range between = 1E + 03 and = 1E + 06 copies/ ml faeces while the copy number of rAiA were slightly lower ranging between 8.99E + 02 and 5.60E + 05 (Fig. 5B). rAiA was un-detectable by RT-PCR in the analyzed faecal samples of both types of laboratory rats. However rRoB viruses were detectable in both Wistar-type (7/10, 70.0%) and "hooded-type" (1/11, 9.1%) laboratory rats while rCaB viruses were present only among "hooded"-rats (2/11, 18.2%) (Fig. 5A).
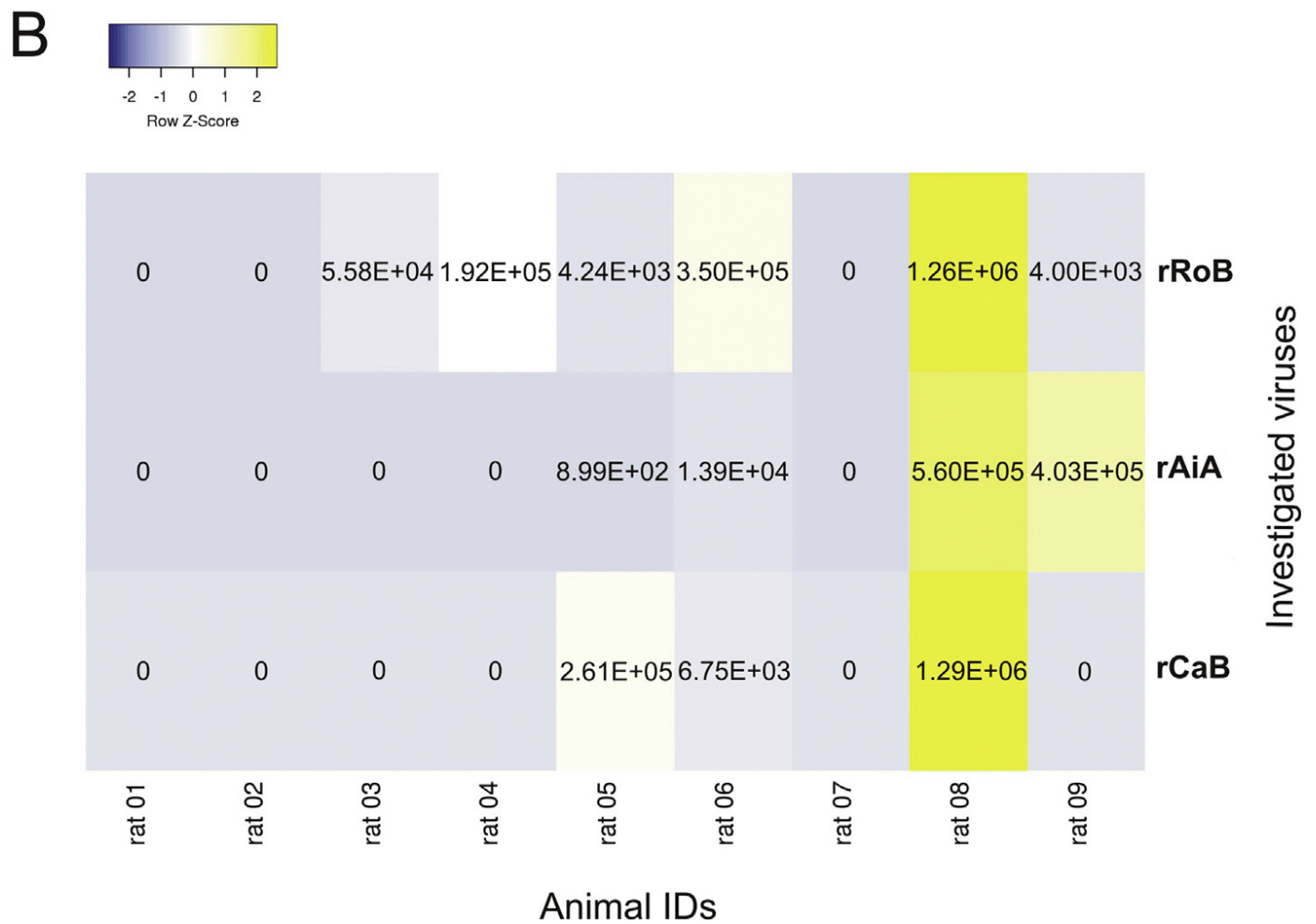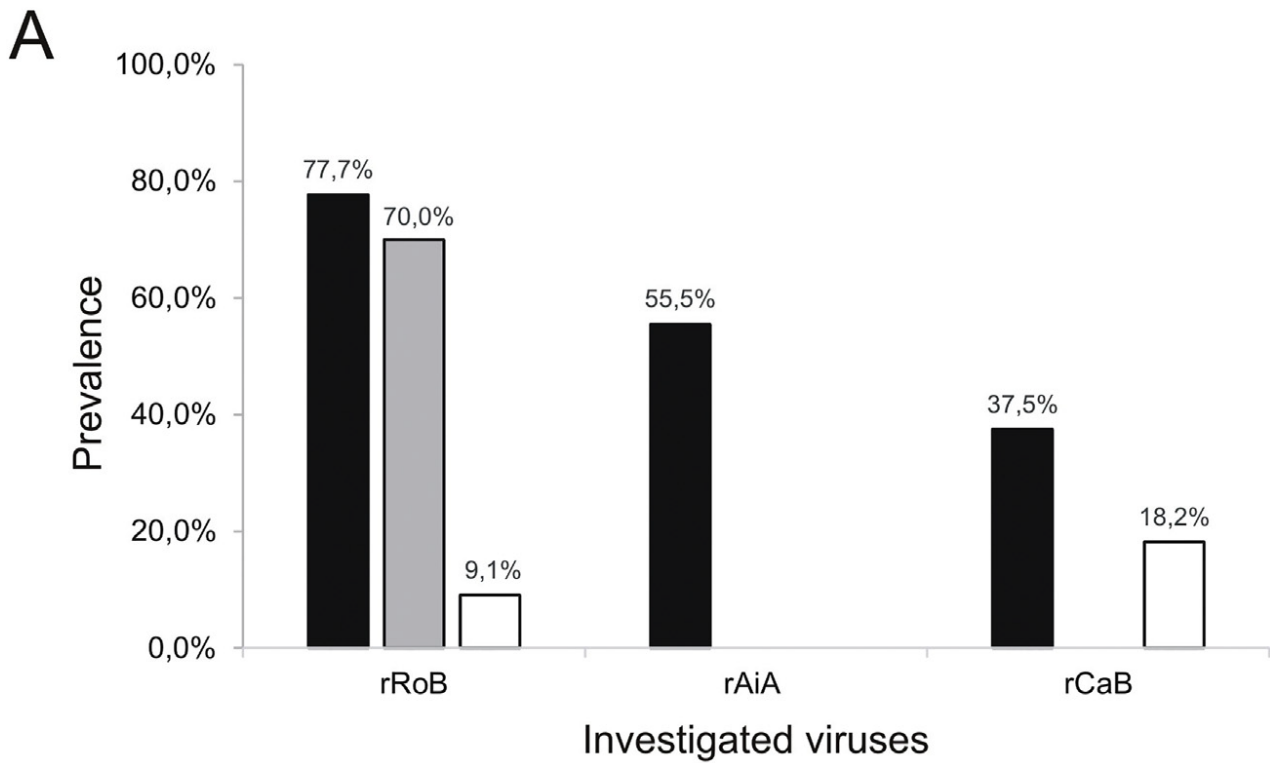
Fig. 5. (A) Prevalence bar plot of the study viruses (rRoB, rAiA and rCaB) based on the results of RTPCR screening of faecal samples collected from freeliving (black bars), laboratory-bred Whistar-type (grey bars) and "hooded-type" (empty bars) Norway rats. (B) Heat map of the investigated viruses (rRoB, rAiA and

rCaB) in faecal samples of free living Norway rats (rat 01 – 09) by RT-qPCR. The yellow and blue colors indicate high and low (or undetectable) copy numbers, respectively. The numbers in the boxes are the measured absolute viral copy numbers/ ml faeces. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.6. Extraintestinal virus detection and quantification

In order to investigate the possible extra-intestinal presence of the study PVs multiple tissue samples were collected from the trapped free-living Norway rats *(N =* 10) from which the virus-positive faecal samples originated. Note that no faecal sample from rat-10 was available for testing (Table 1). The collected tissue samples were tested by the same generic primer-pairs and RT-PCR screening reactions used of the screening of faecal specimens. The absolute copy numbers of the viruses from the RT-PCR-positive tissue samples were also determined using the same SYBR-Green based RT-qPCR method as used for the faecal virus copy number investigations. Only rRoB and rCaB positive RT-PCR tissue samples were found (Table 1). rRoB was detectable only in 1/7 spleen samples while rCaB was detectable in 2/10 liver and 2/7 spleen samples as well as further 1/2 muscle and 1/8 kidney samples (Table 1). Interestingly the rCaB positive liver, spleen, muscle, and kidney specimens were taken from the same animal (rat-08) as the faecal sample which contained all three PVs (in case of rCaB with the highest viral load). Besides the positive tissue samples of rat-08, a spleen sample of rat-10 and a liver sample of rat-03 were also rCaB RT-PCR-positive while the faecal sample of rat-03 was negative for rCaB (Table 1). Only the liver, spleen and muscle samples of rat-08 and the spleen sample of rat-10 were rCaB RT-qPCR-positive. The highest viral loads of rCaB were measured in the spleen (1.74E + 02 copies/|ig total RNA) and liver (1.19E + 02 copies/^g total RNA) followed by the muscle (6.42E + 01 copies/|ig total RNA) samples of rat-08. The spleen sample of rat-10 contained only 5.77E + 01 copies/|ig total RNA. The single rRoB positive spleen sample from rat-10 contained relatively high copy number (1.90E + 05 copies/total RNA).

3.7. VP1 capsid sequence-based phylogenetic analysis of rat picornaviruses

In order to gain more detailed insight into the genetic variance of the three PVs, the VP1 sequence-parts of the previously identified strains were determined and analyzed.

From all of the rRoB positive faecal *(N =* 15) and tissue samples (N = 1) identified by previous screening RT-PCR reactions, a total of 13 VP1 sequences could be determined (Fig. 6A). The partial 579-nt-long rRoB VP1 sequences from the free-living rats show 74-91% nt and 79-96% aa pairwise identities while the VP1 sequences from the laboratory rats show 99% nt and 100% aa identities. The phylogenetic analysis of the VP1 nt sequences shows the presence of multiple different lineages of species *Rosavirus B,* which could be circulating among the different populations of free-living Norway rats in Hungary (Fig. 6A). Interestingly, there are certain rRoB strains like strains from rat-04 (Hajdúböszörmény)

and rat-07 (Eger) or strains from rat-03 (Hajdúböszörmény) and rat-09 (Zimány) that belong to same lineages but originated from geographically distant locations (Fig. 1A; 6A). The rRoB strains identified from the two different types ("Wistar" and "hooded") of laboratory Norway rats belong to a single lineage different from the lineages of the free-living rats suggesting a continuous circulation of this rRoB variant in the two separated rat populations located in two different laboratory animal units (Fig. 6A).

From the rCaB positive faecal *(N = 5)* and tissue samples *(N=6)* identified by previous screening RT-PCR reactions total of four VP1 sequences could be determined (Fig. 6B). Unfortunately, we are unable to acquire VP1 sequences from the two rCaB screening RT-PCR-positive samples of "hooded-type" laboratory rats. The complete 828-nt-long VP1 sequences of free-living rats show 79-99% nt and 88-99% aa pairwise identities to each other. The VP1 sequences identified from free-living rats from Hajdúböszörmény are clustered together and from a single lineage separate from the rCaB strain rat08/rCaB/HUN from Rácalmás (Fig. 6B).

VP1 sequences could be determined from all of the rAiA positive faecal (N = 5) samples (Fig. 6C). The partial 803-nt-long VP1 sequences show 85-90% nt and 91-99% aa pairwise identities to each other. The phylogenetic analysis based on the VP1 nucleotide sequences shows that the rAiA strains from the same location are clustered together and formed geographic origin-specific lineages (Fig. 6C). All Hungarian rAiA strains were located in the same main branch as other AiA strains identified from different rat species (Firth et al., 2014; Lu et al., 2018)(Fig. 6C).
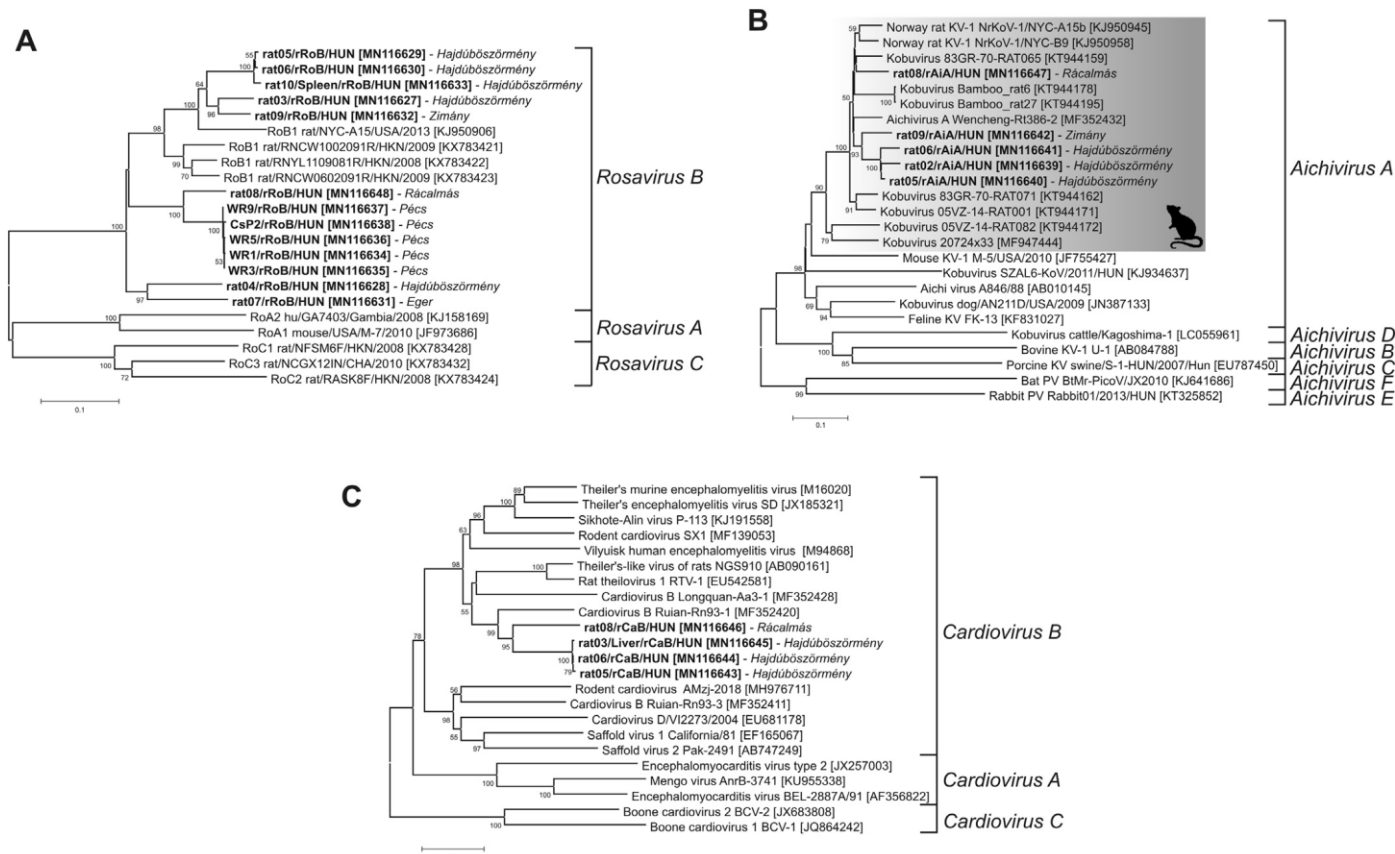
Fig. 6. Phylogenetic analysis of the partial VP1 sequences of the study viruses (in bold) with sample collection sites (in italics) as well as the closest relatives of rat08/rRoB/HUN and other rosaviruses (A) rat08/rAiA/HUN and other kobuviruses (B) and rat08/rCaB/HUN and other cardioviruses (C). The Neighbor-Joining tree was constructed from the alignments of the VP1 nucleotide sequences. Representative members of all species in genus Rosavirus (Rosavirus A-C), genus Kobuvirus (Aichivirus A-F) and genus Cardiovirus (Cardiovirus A-C) are marked. Grey background in panel B indicates the kobuvirus sequences with rat origin. Abbreviations: RoA-C: Rosavirus A-C; WR: "Wistar-type" rat; CsP: "hooded-type" rat; KV: kobuvirus; PV: picornavirus.

4. Discussion

In this study, the complete genomes of three phylogenetically distant PVs from a single faecal sample of a free-living Norway rat (rat-08) in Hungary were determined using viral metagenomics next-generation sequencing and different RT-PCR methods. Beside the analyzed three PVs, the same faecal sample of rat-08 also contains hepatitis E virus-related *(Hepeviridae)* and norovirus-related *(Caliciviridae)* sequence reads (unpublished data) reflecting the simultaneous infection of five different RNA viruses of the same animal.

Based on the results of phylogenetic analyses and genome comparisons the closest relatives of rat08/rRoB/HUN are the well known Rosavirus B1 viruses (species *Rosavirus B* of genus *Rosavirus)* of Norway rats. Therefore rat08/rRoB/HUN appears to be a distinct genotype of *Rosavirus B* (Firth et al., 2014; Lau et al., 2016). Noteworthy, rosaviruses have one of the longest 3'UTRs among PVs ranged between 696 nt (Rosavirus A2, KJ158169) and 943 nt (Rosavirus B1, KX783423) (Lim et al., 2014; Lau et al., 2016). Interestingly, a 248 nt long deletion is observed at the beginning of 3'UTR of rat08/rRoB/HUN resulting in the shortest 3'UTR sequence among rosaviruses. The manual sequence-repeat search revealed a previously unknown 222-nt-long sequence repeat ("Rep", Fig. 3) with a compact, conserved secondary RNA structure in the 3'UTR of the study sequence as well as in the other rosavirus 3'UTRs which suggest a currently unknown but most likely pivotal role for these structures in genome replication. The functional roles of conserved RNA secondary structures and repeated sequence motifs of the 3'UTRs in the genome circularization and replication processes of various +ssRNA viruses have been recently shown (Garcia-Moreno et al., 2016; Kloc et al., 2018).

In the cases of rat08/rAiA/HUN and rat08/rCaB/HUN, there are only unassigned rat PV sequences as closest relatives in GenBank and none of them is currently analyzed in detail. For further genome analyses, the most closely assigned and published PVs (MuKV-1 and TMEV) were therefore used. Based on the results of phylogenetic analyses as well as the presence of similar type of IRES (type-V), identical genome architecture (L-3-3-4), high sequence identities between the corresponding genome parts (up to 93%) and the general presence of "barbell-like" structure in the 3'UTRs of rat08/rAiA/HUN and MuKV-1

could suggest that rat08/rAiA/HUN belongs to the same species of *Aichivirus A* of genus *Kobuvirus* as MuKV-1. The rat08/rCaB/HUN has a type II-like IRES, a genome layout of L-(L*)-4-3-4 including the presence of an alternative Leader (L*) encoded by a minor ORF which is a characteristic genome feature of TO subgroup of TMEV of species *Cardiovirus B* (Takano-Maruyama et al., 2006). These genomic properties, as well as the results of phylogenetic analyses, indicate that /rCaB/HUN could be a distinct genotype of species *Cardiovirus B*.

Small scale epidemiological investigation of the study PVs revealed high prevalence (77.7% of rRoB, 55.5% of rAiA and 37.5% of rCaB) in the analyzed faecal samples of free-living Norway rats collected from various locations in Hungary suggesting wide distributions among different wild rat populations.

The phylogenetic analyses of the VP1 capsid protein sequences show the presence of multiple, mostly geographical location-specific lineages of all three examined rat PVs. However, the presence of phy-logenetically closely related RoB viruses like strains from rat03 and rat09 originated from geographically distant locations (Hajdúböszörmény and Zimány, c.a. 400 km distance apart) suggests continuous virus-exchange between distant wild rat populations (Fig. 1A; 6A). Mostly due to a low number of known VP1 sequences little information is available about the genomic diversity of rat *Rosavirus B* and rat *Cardiovirus B* viruses while multiple kobuvirus lineages were described recently which could be co-circulating among different rodent populations (Lu et al., 2018).

The measured highest copy numbers of the examined PVs with the magnitude of $= 10^5$ and 10^6 (Table 1) could indicates the replication capability and the faecal-oral route of transmission for these viruses among Norway rats as natural hosts.

Due to the high detection rates of PVs in the faeces of analyzed free-living rats and because different types of Norway rats (i.e. "Wistar" and "hooded" types) served as an important experimental animal models for various scientific research projects (Baker et al., 2013) the presence of these viruses were also examined among laboratory rats originated from two separated rat laboratory populations ("Wistar-type" and "hooded-type" rats) held in separate laboratory animal units in Pécs, Hungary. While rAiA was not detected, both rRoB and rCaB viruses were detectable in the faeces of some laboratory rats. Surprisingly, while rCaB was detectable only among "hooded-type" of rats with lower abundance (18.2%) than found in free-living rats (37.5%) the detection rate of rRoB was similarly high (70.0%) in the Wistar-type laboratory rats than among wild animals (77.7%). The high abundance of rRoB and rCaB in adult Norway rats indicates the continuous circulations and likely chronic, persistent type of infections of both viruses. Persistent infection was described in the case of certain TMEV (species *Cardiovirus B)* types among mice (Trottier et al., 2001). To our knowledge, the abundance of these viruses was not previously investigated among laboratory rats. Whether these viruses are also present in other national or foreign laboratory animal facilities worldwide remain an outstanding question. This is not the first report of viruses prevalent in widely used experimental animals. Similarly, highly prevalent murine astrovirus *(Astroviridae),* rat

polyomavirus *(Polyomaviridae)* and zebra-fish picornavirus were identified recently in laboratory mice, rats and laboratory-raised zebrafish as well (Ng et al., 2013; Besch-Williford et al., 2017; Altan et al., 2019).

Although the analyzed adult laboratory rats show no observable clinical symptoms the possible impact of these infections on the results of scientific experiments conducted on these types of rats as well as the potential risks of zoonotic infections affecting the research and animal care personnel are currently unknown. However the recent detection of Rosavirus A viruses in human infections (Lim et al., 2014) as well as the wide host species spectra and high risk of human transmissibility of Encephalomyocarditis virus of species *Cardiovirus A* (Carocci and Bakkali-Kassimi, 2012; Walker et al., 2018) could increase human health concerns with regard to identified rat rosa-, and cardioviruses.

Tissue distributions of the study viruses using RT-PCR/qPCR were also investigated. Beside the high abundance of Rosavirus B found in faecal samples this virus was detectable only in a single tissue sample (a spleen specimen) from a free-living Norway rat with a relatively high copy number (1.90E + 05 copies/ug total RNA), which suggest the extra-intestinal replication capability of certain rRoB strains in a permissive host. Although to our knowledge the tissue tropism of RoB viruses has not been investigated before the previous detection of these viruses in respiratory samples of Norway rats also suggests the extraintestinal spread of the virus (Lau et al., 2016). While MuKV can be rarely detected in liver samples of mice (Williams et al., 2018) the RT-PCR screening of rAiA showed no positivity in any of the analyzed tissues of rats including liver likely due to the low number of available samples.

Interestingly, rCaB was detectable in the liver, spleen, muscle and kidney of the same animal (rat-08) of which faecal sample contained all three PVs. The highest viral loads were measured in the liver and spleen (Table 1). The virus was also detectable from the additional liver and spleen samples from different rats as well which suggests the capacity for the extra-intestinal spread of rCaB that is also supported by the detection of the virus in a liver sample while it was undetectable in the faeces of the same animal (rat-3, Table 1). Although due to the low copy numbers in the organs and high viral loads of rCaB measured in the faeces there is still no indication whether it can replicate efficiently outside the intestinal tract. The extra-intestinal spread of other cardi-oviruses was previously reported. The rCaB-related TMEV was reported to infect neurons of the CNS and cardiomyocytes of the heart and can also be detectable in the liver and spleen samples of mice (Ha-Lee et al., 1995; Omura et al., 2018).

Limitations of this study are the relatively small number of available samples for investigation and the unknown health status of the wild animals sampled. Therefore more detailed analyses including in vivo experimental infection studies are necessary to explore the true prevalence, pathogenesis and genomic diversity of these PVs.

Supplementary data to this article can be found online at https:// doi.org/10.1016/j.meegid.2019.103988.

Funding

Declaration Competing Interest

None.

Acknowledgements

References

Altan, E., Kubiski, S.V., Boros, A., Reuter, G., Sadeghi, M., Deng, X., ... Delwart, E., 2019. A highly divergent Picornavirus infecting the gut epithelia of Zebrafish (Danio rerio) in research institutions worldwide. Zebrafish 16 (3), 291-299.

Argos, P., Kamer, G., Nicklin, M.J., Wimmer, E., 1984. Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. Nucleic Acids Res. 25, 7251-7267.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., Wishart, D.S., 2016. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 44 (W1), W147-W153.

Baker, H.J., Lindsey, J.R., Wesibroth, S.H., 2013. The Laboratory Rat: Biology and Diseases (Vol. 1). Elsevier.

Besch-Williford, C., Pesavento, P., Hamilton, S., Bauer, B., Kapusinszky, B., Phan, T., Levin, S., 2017. A naturally transmitted epitheliotropic polyomavirus pathogenic in immunodeficient rats: characterization, transmission, and preliminary epidemiologic studies. Toxicol. Pathol. 45 (5), 593-603.

Boros, A., Pankovics, P., Simmonds, P., Reuter, G., 2011. Novel positive-sense, single-stranded RNA (+ssRNA) virus with di-cistronic genome from intestinal content of freshwater carp *(Cyprinus carpio)*. PLoS One 6, e29145.

Boros, A., Pankovics, P., Knowles, N.J., Reuter, G., 2012. Natural interspecies recombinant bovine/porcine enterovirus in sheep. J. Gen. Virol. 93, 1941-1951. Boros, Á., Pankovics, P., Reuter, G., 2014. Avian picornaviruses: molecular evolution, genome diversity and unusual genome features of a rapidly

expanding group of viruses in birds. Infect. Genet. Evol. 28, 151-166. Carocci, M., Bakkali-Kassimi, L., 2012. The encephalomyocarditis virus. Virulence 3 (4), 351-367.

Chen, L., Liu, B., Wu, Z., Jin, Q., Yang, J., 2017. DRodVir: a resource for exploring the virome diversity in rodents. J. of Genet. and Genomics 44 (5), 259-264. Darty, K., Denise, A., Ponty, Y., 2009. VARNA: interactive drawing and editing of the RNA secondary structure. Bioinformatics 25 (15), 1974.

Duke, G.M., Hoffman, M.A., Palmenberg, A.C., 1992. Sequence and structural elements that contribute to efficient encephalomyocarditis virus RNA translation. J. Virol. 66 (3), 1602-1609.

Firth, C., Bhat, M., Firth, M.A., Williams, S.H., Frye, M.J., Simmonds, P., ... Lee, B., 2014. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal Rattus norvegicus in New York City. MBio 5 (5) (e01933-14).

Garcia-Moreno, M., Sanz, M.A., Carrasco, L., 2016. A viral mRNA motif at the 3'-untranslated region that confers translatability in a cell-specific manner. Implications for virus evolution. Sci. Report. 6, 19217.

Gorbalenya, A.E., Donchenko, A.P., Blinov, V.M., Koonin, E.V., 1989. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. FEBS Lett. 243, 103-114.

Gorbalenya, A.E., Koonin, E.V., Wolf, Y.I., 1990. A new superfamily of putative NTP-binding domains encoded by genomes of small DNA and RNA viruses. FEBS Lett. 262, 145-148.

Ha-Lee, Y.M., Dillon, K., Kosaras, B., Sidman, R., Revell, P., Fujinami, R., Chow, M., 1995. Mode of spread to and within the central nervous system after oral infection of neonatal mice with the DA strain of Theiler's murine encephalomyelitis virus. J. Virol. 69 (11), 7354-7361.

Hughes, P.J., Stanway, G., 2000. The 2A proteins of three diverse picornaviruses are related to each other and to the H-rev107 family of proteins involved in the control of cell proliferation. J. Gen. Virol. 81 (1), 201-207.

Kloc, A., Rai, D.K., Rieder, E., 2018. The roles of picornavirus untranslated regions in infection and innate immunity. Front. Microbiol. 9, 485.

Kofstad, T., Jonassen, C.M., 2011. Screening of feral and wood pigeons for viruses harbouring a conserved mobile viral element: characterization of novel Astroviruses and Picornaviruses. PLoS One 6 (10), e25964.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33 (7), 1870-1874.

Lau, S.K., Woo, P.C., Li, K.S., Zhang, H.J., Fan, R.Y., Zhang, A.J., ... Chan, K.H., 2016. Identification of novel Rosavirus species that infects diverse rodent species and causes multisystemic dissemination in mouse model. PLoS Pathog. 12 (10), e1005911.

Lim, E.S., Cao, S., Holtz, L.R., Antonio, M., Stine, O.C., Wang, D., 2014. Discovery of rosavirus 2, a novel variant of a rodent-associated picornavirus, in children from the Gambia. Virology 454, 25-33.

Lu, L., Van Dung, N., Ivens, A., Bogaardt, C., O'Toole, A., Bryant, J.E., ... Tue, N.T., 2018. Genetic diversity and cross-species transmission of kobuviruses in Vietnam. Virus Evol. 4 (1), vey002.

Ng, T.F.F., Kondov, N.O., Hayashimoto, N., Uchida, R., Cha, Y., Beyer, A.I., ... Delwart, E., 2013. Identification of an astrovirus commonly infecting laboratory mice in the US and Japan. PLoS One 8 (6), e66937.

Nicholas, K.B., Nicholas, H.B., 1997. GeneDoc: a tool for editing and annotating multiple sequence alignments. In: National Resource for Biomedical Supercomputing, http://iubio.bio.indiana.edu/soft/molbio/ibmpc/genedoc-readme.html, Accessed date: 11 May 2019.

Okonechnikov, K., Golosova, O., Fursov, M., the UGENE team, 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28, 1166-1167.

Omura, S., Kawai, E., Sato, F., Martinez, N.E., Minagar, A., Al-Kofahi, M., ... Tsunoda, I., 2018. Theiler's Virus-Mediated Immunopathology in the CNS and Heart: Roles of Organ-Specific Cytokine and Lymphatic Responses. Front. Immunol. 9.

Palmenberg, A., Neubauer, D., Skern, T., 2010. Chapter 1: Genome organization and encoded proteins. In: Ehrenfeld, E., Domingo, E., Roos, R.P. (Eds.), The Picornaviruses. ASM Press, Washington, DC, pp. 3-17.

Pevear, D.C., Calenoff, M., Rozhon, E., Lipton, H.L., 1987. Analysis of the complete nu-cleotide sequence of the picornavirus Theiler's murine encephalomyelitis virus indicates that it is closely related to cardioviruses. J. Virol. 61 (5), 1507-1516.

Phan, T.G., Kapusinszky, B., Wang, C., Rose, R.K., Lipton, H.L., Delwart, E.L., 2011. The fecal viral flora of wild rodents. PLoS Pathog. 7 (9), e1002218.

Phan, T.G., Vo, N.P., Simmonds, P., Samayoa, E., Naccache, S., Chiu, C.Y., Delwart, E., 2013a. Rosavirus: the prototype of a proposed new genus of the Picornaviridae family. Virus Genes 47 (3), 556-558.

Phan, T.G., Vo, N.P., Boros, A., Pankovics, P., Reuter, G., Li, O.T., Wang, C., Deng, X., Poon, L.L., Delwart, E., 2013b. The viruses of wild pigeon droppings. PLoS One 8, e72787.

Strand, T.M., Lundkvist, A., 2019. Rat-borne diseases at the horizon. A systematic review on infectious agents carried by rats in Europe 1995-2016. Infect. Ecol. & Epidemiol. 9 (1), 1553461.

Sweeney, T.R., Dhote, V., Yu, Y., Hellen, C.U., 2012. A distinct class of internal ribosomal entry site in members of the Kobuvirus and proposed Salivirus and Paraturdivirus genera of the Picornaviridae. J. Virol. 86 (3), 1468-1486.

Takano-Maruyama, M., Ohara, Y., Asakura, K., Okuwa, T., 2006. Leader (L) and L proteins of Theiler's murine encephalomyelitis virus (TMEV) and their regulation of the virus' biological activities. J. Neuroinflammation 3 (1), 19.

Trottier, M., Kallio, P., Wang, W., Lipton, H.L., 2001. High numbers of viral RNA copies in the central nervous system of mice during persistent infection with Theiler's virus. J. Virol. 75 (16), 7420-7428.

Walker, J.W., Han, B.A., Ott, I.M., Drake, J.M., 2018. Transmissibility of emerging viral zoonoses. PLoS One 13 (11), e0206926.

Williams, S.H., Che, X., Garcia, J.A., Klena, J.D., Lee, B., Muller, D., ... Lipkin, W.I., 2018. Viral diversity of house mice in New York City. Mbio 9 (2) (e01354-17).

Zell, R., Delwart, E., Gorbalenya, A.E., Hovi, T., King, A.M.Q., Knowles, N.J., Simmonds, P., 2017. ICTV virus taxonomy profile: Picornaviridae. The J. of Gen. Virol. 98 (10), 2421.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406-3415.