Genetically highly divergent RNA virus with astrovirus-like (5'-end) and hepevirus-like (3'-end) genome organization in carnivorous birds, European roller *(Coracias garrulus)*

Péter Pankovics[a], Ákos Boros[a], Tamás Kiss[b], Péter Engelmann[c], Gábor Reuter^*

[a] *Department of Medical Microbiology and Immunology, Medical School, University of Pécs, Pécs, Hungary* [b] *Hungarian Ornithological and Nature Conservation Society, Budapest, Hungary*

[c] *Department of Immunology and Biotechnology, Clinical Center, Medical School, University of Pécs, Pécs, Hungary*

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

Keywords: Astrovirus Hepatitis E virus Rana hepevirus European roller

A B S T R A C T

Astroviruses (family *Astroviridae)* and hepeviruses (family *Hepeviridae)* are small, non-enveloped viruses with genetically diverse +ssRNA genome thought to be enteric pathogens infecting vertebrates including humans. Recently, many novel astro- and hepatitis E virus-like +ssRNA viruses have been described from lower vertebrate species. The non-structural proteins of astro- and hepeviruses are highly diverse, but the structural/capsid proteins represent a common phylogenetic position shed the light of their common origin by inter-viral recombination. In this study, a novel astrovirus/hepevirus-like virus with +ssRNA genome (Er/SZAL5/HUN/ 2011, MK450332) was serendipitously identified and characterized from 3 (8.5%) out of 35 European roller *(Coracias garrulus)* faecal samples by RT-PCR in Hungary. The complete genome of Er/SZAL5/HUN/2011 (MK450332) is 8402 nt-long and potentially composed three non-overlapping open reading frames (ORFs): ORF1a (4449nt/1482aa), ORF1b (1206nt/401aa) and ORF2 (1491 nt/496aa). The ORF1ab has an astrovirus-like genome organization containing the non-structural conserved elements (TM, CC, NLS, VPg) and enzyme residues (trypsine-like protease, RNA-dependent RNA-polymerase) with low amino acid sequence identity, 15% (ORF1a) and 44% (ORF1b), to astroviruses. Supposedly the ORF2 is a capsid protein but neither the astrovirus-like subgenomic RNA

promoter (sgRNA) nor the astrovirus-like capsid characteristics have been identifiable. However, the predicted capsid protein (ORF2) showed 26% identity to the corresponding protein of hepevirus-like novel Rana hepevirus (MH330682). This novel +ssRNA virus strain Er/SZAL5/HUN/2011 with astrovirus-like genome organization in the non-structural genome regions (ORF1a and ORF1b) and Rana hepevirus-related capsid (ORF2) protein represent a potentially recombinant virus species and supports the common origin hypothesis, although, the taxonomic position of the studied virus is still under discussion.

1. Introduction

The European roller *(Coracias garrulus)* is a small (average size 32 cm) long-distance migrant predator among wild birds breeding from South and Eastern Europe to Baltic but overwinters mainly in sub-Saharan Africa (del Hoyo et al., 1994; Fry et al., 2018; http://www. arkive.org/european-roller/coracias-garrulus/). They breed in roosts from the steppe, grasslands to parklands and in Hungary, the European roller is a frequent bird mainly in the territory Kiskunság and BorsodiMezőség (Boros et al., 2013; Kiss et al., 2014, 2016). The heavy-bodied European roller feeds on invertebrates, such as beetles, crickets, locusts, caterpillars, flies and spiders, however, they are also known to prey on weak, small birds and lower vertebrates, such as frogs, lizards and snakes (del Hoyo et al., 1994; Kiss et al., 2014, 2016).

Astroviruses (family *Astroviridae)* and hepatitis E viruses (family *Hepeviridae)* are positive sense, single-stranded RNA viruses, without an envelope, transmitted by the faecal-oral route. Astroviruses (AstVs) are classified into two major groups of viruses, genera *Mamastrovirus* and *Avastrovirus* (Guix et al., 2013; Pantin-Jackwood et al., 2013, Bosch et al., 2014), while hepatitis E viruses (HEVs) are currently divided into *Orthohepevirus* and *Pischihepevirus* genera (Smith et al., 2014). Both AstVs and HEVs have = 6.6-8.5 kb long genome flanked by viral protein genome-linked (VPg) and a cap structure at the 5'end, respectively, and a poly(A)-tail at the 3'end (Meng et al., 2012; Bosch et al., 2014). AstVs and HEVs have three/four open reading frames (ORFs): the ORF1a and ORF1b in AstVs and ORF1 in HEVs encode the non-structural proteins for virus RNA transcription and translation, the ORF2 encodes the protein generating the viral capsid. An alternate ORFx, ORF3 were described in some AstVs and HEVs species, respectively (Graff et al., 2006; Firth and Atkins, 2010; Meng et al., 2012).

In astroviruses, the capsid protein is translated from a subgenomic RNA (sgRNA) using the highly conservative sgRNA translation initiation signal (Monroe et al., 1993; Knowles et al., 2012; Méndez et al., 2013). Focusing the main feature of astrovirus capsid proteins, these proteins show higher genetic variability among astroviruses, the N-terminal half of the capsid protein in astroviruses containing the basic/ inner core/outer core elements are "conserved" with 84%-97% aa identities, while the C-terminal half containing the spike/acidic elements are "variable" with lower 39%-77% aa identities (Arias and DuBois,

2017). A conserved nucleotide stem-loop-II-motif (s2 m) was described in the ORF2/3'UTR junction of some astroviruses (Monroe et al., 1993; Monceyron et al., 1997).

In hepeviruses, the capsid coding genomic region shows considerably complex structure and function (Cao and Meng, 2012; Kelly et al., 2016). Both the capsid protein (ORF2) and the phosphoprotein (ORF3) are translated from a bicistronic subgenomic RNA (sgRNA) (Kelly et al., 2016) using a highly conserved RNA motif in the cis-reactive (CRE) element between the ORF1 and ORF2/ORF3 junction region (Graff et al., 2005, 2006; Cao et al., 2010; Ding et al., 2018). Additionally, two highly conserved internal stem-loops (ISL) are present in the middle of the capsid protein coding region for regulation of the hepatitis E virus replication mechanism (Emmerson et al., 2013). The second CRE structure is located at the 3'end of the genome overlapping the 3'end of the ORF2 and the 3' non-coding region (Emerson et al., 2001).

Based on phylogenetic analysis of the complete RNA genomes AstVs represent a clade among of the "Picorna-like" supergroup of viruses (Koonin et al., 2008), unlike HEVs, where the non-structural proteins are clustered together with the "Alpha-like" supergroup of viruses (Purdy et al., 1993; Kelly et al., 2016). In contrast, the capsid protein encoded by ORF2 of HEVs has no phylogenetic connection the members of the "Alpha-like" viruses, but are more related to viruses in family *Astroviridae* ("Picorna-like" supergroup) (Kelly et al., 2016). Recently, many novel astrovirus and hepatitis E virus-like RNA viruses have been described from lower vertebrate species, however, the phylogenetic positions of these viruses are still under discussion (Shi et al., 2018.). A recent study reports a novel Rana hepevirus from agile frog tadpoles, which has a hepatitis E virus-like non-structural genome organization and a capsid protein with unknown phylogenetic position (Reuter et al., 2018).

To date, several novel viruses with RNA genome representing different viral families have been identified from European roller's faecal samples by viral metagenomic survey (Boros et al., 2013; Reuter et al., 2014; Pankovics et al., 2015a, 2015b) and this method gives us good opportunities to discover further interesting findings. This study reports the genome characterization and phylogenetic analysis of a novel + ssRNA virus strain Er/SZAL5/HUN/2011 with astrovirus-like genome organization in the non-structural genome regions (ORF1a and ORF1b) and Rana hepevirus-related capsid (ORF2) protein from European roller's faecal samples, in Hungary.

2. Materials and methods

In July 2011, faecal samples were collected from clinically healthy wild birds of prey (N = 19), European rollers *(Coracias garrulus)in* Dorozsma-Majsai homokhát (N = 15) and Borsodi Mezőség (N = 4) breeding territories of the Great Hungarian Plain, Hungary. Specimens were collected from natural nests during regular bird ringing by qualified ornithologists with valid permission (Permit No. NIFENW14/ 1368-5/2011). The main goal was to screen astroviruses by RT-PCR using common screening primer-pairs

targeting the RNA-dependent RNA polymerase region (Chu et al., 2008). The 443 bp-long PCRproducts were cloned with CloneJET PCR Cloning Kit (Thermo Scientific,) according to the manufacturer's sticky-end cloning protocol. Briefly, 10 |il 2 x reaction buffer, 3 |il non-purified PCR product (45.1 ng/|il), 1 |il DNA blunting enzyme and 4 |il nuclease-free water (Thermo Scientific). The mixture was incubated at 70 °C 5 min, then immediately placed on ice. 1 |il pJET1.2/blunt cloning vector (50 ng/|il) and 1 |il T4 DNA ligase enzyme were added to the reaction. Two different competent *E. coli* strains SURE2 and DH5a were used in the transformation reaction. In total 50 |il $CaCl_2$ transformed bacterial cells and 5 |il plasmid ligated PCR-product were gently rocked, then placed on ice 1 h and given a heat-shock at 42 °C 1 min. Afterwards, the mixture was put in 500 |il liquid Luria-Bertani (LB) Medium (Sambrook Handbook) and incubated at 22 °C 1 h. Finally, the bacterial solution was plated on ampicillin (100 |ig/ml) containing LB plates and incubated at 37 °C overnight. The individual clones were subjected to colony PCR using pJET1.2 forward and reverse primer sets included in the CloneJet Kit.

In addition, several different primers sets were designed based on astrovirus-like sequences determined from the colony PCR to obtain the complete viral genome using primer walking (Sverdlov and Azhikina, 2005), 5'/3' rapid amplification of cDNA ends, RACE (Boros et al., 2011) and TAIL-PCR (Liu and Chen, 2007) strategies (Suppl 1. Fig. 1). Briefly, a complementary DNA (cDNA) was generated using the universal OligodT anchor reverse primer (Boros et al., 2011) using RNA sample isolated from the original faecal specimen which transcribed all RNAs having a poly(A) end. Then, at least two sequence specific forward (SpecF) (5'-»3' for downstream reaction) primers were designed based on the available partial RdRp sequence. The first round PCR contained the SpecF1 and the PCR anchor (the reverse complement of the anchor part of OligodT anchor primer) and the second round PCR contained the SpecF2 and PCR anchor primer sets to amplify the questionable genomic region between the partial RdRp and the 3'-poly (A) tail. The generated PCR-product was sequenced by primer-walking method. Further sequence specific forward and reverse primers were designed for obtaining the 5' genome parts respectively by 5'/3' RACE (Boros et al., 2011) and thermal asymmetric interlaced PCR (Liu and Chen, 2007) methods. After every second/third sequencing reactions sequence specific reverse/forward oligonucleotides and long-PCR reactions (up to 2500 bp) were designed to control the single genome continuity. All PCR-products were directly sequenced and run on an automated sequencer (ABI Prism 310, Applied Biosystems, Stafford,
USA).

Faecal specimens from additional European rollers, and potential dietary sources such as agile frog tadpole *(Rana* dalmatina)(Reuter et al., 2018), common toad *(Bufo bufo)* (in this study), smooth newt *(Lissotriton* vulgaris)(Reuter et al., 2015; Pankovics et al., 2017) and pools (M1-M9) of ornithophil mosquitoes (Culex sp.) (Reuter et al., 2016) were tested by RT-PCR using sequence-specific screening primer pairs [SZAL5A-ORF2-KontrF: 5'-ATAGTCACTGCACTAACTGCT-3' corresponding nt positions 6234-6254 and SZAL5A-ORF2-KontrF: 5'-AAC TTGAGGGCGTGACATCAT-3' corresponding nt

positions 6741-6761 of the study strain Er/SZAL5/2011/HUN] designed for the putative capsid genome region (ORF2).

Representative and complete astrovirus nucleotide (nt) sequences (Suppl. 2 Table 1) were collected from GenBank database and ORF1a, ORF1b and ORF2 were predicted by the NCBI ORF finder with default settings (https://www.ncbi.nlm.nih.gov/orffinder). The amino acid (aa) sequences were aligned with Multiple Alignment using Fast Fourier Transform application of EMBL-EBI (https://www.ebi.ac.uk/Tools/msa/mafft/) using default settings and the alignments were used for phylogenetic and evolutionary analysis conducted by MEGA7 (Kumar, Stecher, and Tamura 2016). Dendrograms were conducted by Maximum-likelihood method based on the JTT with Freqs (+F) substitution model, G + I rate and use all sites chosen by the results of best aa model search. Bootstrap values were determined for 1000 replicates using the whole data subset and the tree was drawn to scale, with branch length measured in the number of substitutions per site.

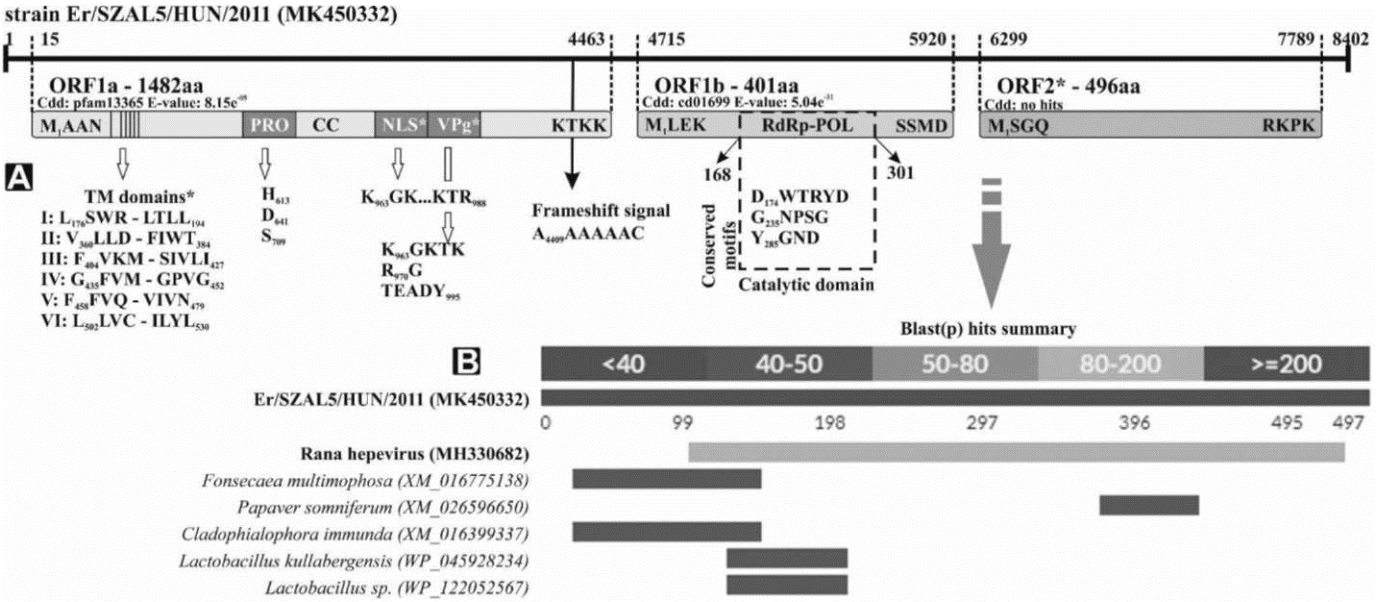

Fig. 1. A) The genome map of the novel virus strain Er/SZAL5/HUN/2011 (MK450332) from European roller. The thin line represents the nt sequence of the study strain, while the boxes below shows the potential viral polyproteins included the length of the proteins. The conserved nt aa motifs were indicated to the schematic genome organization of each ORFs. The similarity analysis of the putative ORF1a, ORF1b and ORF2 proteins of the study strain are calculated by conserved domain search and Blast(p) and the results (if available) are indicated above each ORF boxes. B) The blast(p) search summary of the study strain's putative capsid protein. The significant hit marked. TM: transmembrane; PRO: viral protease; CC: coiled-coiled elements; NLS: nuclear localization signal; VPg: viral protein genome-linked; RdRp-POL: RNA-dependent RNA polymerase; *: The asterisk represents the link to the analysis found in the supplementary 1 material (Suppl. 1 Fig. 2-5, 7).

The pairwise nt/aa sequence comparisons and identity calculations were conducted by Needleman-Wunsch Global Alignment method (https:// blast.ncbi.nlm.nih.gov/Blast.cgi) (Suppl. 3 Table 1). The closest relative search of the study strain was done using NCBI Blast(p) (https://blast. ncbi.nlm.nih.gov/Blast.cgi). Because of the basic Blast search did not return any specific hits, the parameters were specified into *Astroviridae* (taxid: 39733); unclassified *Astroviridae* (taxid: 352926); *Hepeviridae* (taxid: 291484) and unclassified *Hepeviridae* (taxid: 1009842) using DELTA-BLAST algorithm and BLOSUM90 scoring matrix. The aa

sequence analysis was carried out using profile hidden Markov Model (HMMER) (https://www.ebi.ac.uk/Tools/hmmer/;Potter et al., 2018). The prediction of transmembrane domains was analysed by Protter (http://wlab.ethz.ch/protter/start/) web services (Omasits et al., 2014) (Suppl. 1 Fig. 2) and the nuclear localization signals searched by NLStradamus web service (http://www.moseslab.csb.utoronto.ca/ NLStradamus/;Ba et al., 2009) (Suppl. 1 Fig. 4). The secondary RNA structures were predicted by RNAstructure (http://rna.urmc.rochester.edu/RNAstructureWeb/index.html;Bellaousov et al., 2018.), the protein folding prediction was carried out using Foldlndex (https://fold. weizmann.ac.il/fldbin/findex;Prilusky et al., 2005) (Suppl. 1 Fig. 3). GeneDoc and Simplot softwares were used to edit sequences and recombination analyses (bootscanning), respectively (Nicolas and Nicolas, 1997; Lole et al., 1999).

The complete genome sequence of the astrovirus-like study strain Er/SZAL5/2011/HUN has been deposited in GenBank under accession
number MK450332.

3. Results

From the total of 19 analysed apparently healthy wild European roller faecal samples 3 (15.8%) were positive (samples: SZAL5, SZAL6 and BMTK-529) using PanAstV screening primers (Chu et al., 2008). Using the direct Sanger sequencing method, the PCR product of SZAL6 originated from a novel mamastrovirus in avian species (Pankovics et al., 2015a, b), but in case of samples SZAL5 and BMTK-529 no reliable sequences could be generated using direct sequencing method.

Therefore, these two unidentified PCR-products were cloned and se-quenced. BMTK-529 sample revealed a novel avastrovirus with 45% aa identity (Blast(x) query cover: 98%, E-value: $1e^{-29}$) to the RdRp gene (ORF1b) of Avastrovirus 3 strain MPJ0110 (JX985677), while the SZAL5 sample revealed a novel mamastrovirus with 48% aa identity (Blast(x) query cover: 84%, E-value: $3e^{-26}$) to the RdRp gene (ORF1b) of feline astrovirus (mamastrovirus 2) strain 443/07A (GU980969) as the closest relatives, respectively. This study focuses on the complete genome acquisition and genetic characteristics of the latest, presumably novel mamastrovirus from sample SZAL5.

From the 443 bp-long partial RdRp gene sequence acquired from the colony PCR, the complete, 8402 nt-long (excluding the poly(A) tail) viral RNA genome of Er/SZAL5/HUN/2011 was characterized (Fig. 1). The ORF1a, ORF1b and ORF2 regions of Er/SZAL5/HUN/2011 are
4449 nt, 1206 nt and 1491 nt long and encode three (1482 aa, 401 aa and 496 aa) potential protein precursors, respectively (Fig. 1). The hypothetical ORFX or other potential ORFs were not found (Firth and Atkins, 2010). The 5'UTR is 14 nt long and no conserved promoter sequence was observed. The putative in-frame AUG initiation codon of ORF1a starts at nt position 15 (GUU**A_{15}UG**GC, start codon was bolded).

The putative aa sequence of strain Er/SZAL5/HUN/2011 ORF1a/ ORF1b/ORF2 showed low identity and similarity scores to other representative astrovirus strains by using multiple sequences and comparative sequence-to-sequence (Global) analysis (Fig. 2 and Suppl. 3 Table 1). Using the Blast(p) and HMMER searches the ORF1a protein of the study strain showed the closest 41% aa identity to feline astrovirus D1 (NC_024701) with E-value $7e^{-8}$. However, the aa sequence identity searches were significant with E-value better than the threshold, but low sequence query coverage was revealed. Comparing the study strain to feline astrovirus D1 (NC_024701) by Needleman-Wunsch calculation, the query sequence had only 15% identity to the target (NC_024701) protein sequence. The ORF1a showed the closest match to trypsin-like peptidase domain by conserved domain search (CDS) (Fig. 1) (Marchler-Bauer et al., 2017). The catalytic triad of the viral serine protease is composed by His(H)-Asp(D)-Ser(S) amino acids (Speroni et al., 2009) and these conserved aa residues the H_{613},D_{641} and S_{709} were identified in strain Er/SZAL5/HUN/2011. In total, six potential transmembrane domains (TM) (aa 176-194, 360-384, 404-127, 435-452, 458-479, 502-530) were predicted in the N-terminal half of Er/SZAL5/HUN/2011 ORF1a (Suppl. 1 Fig. 2).

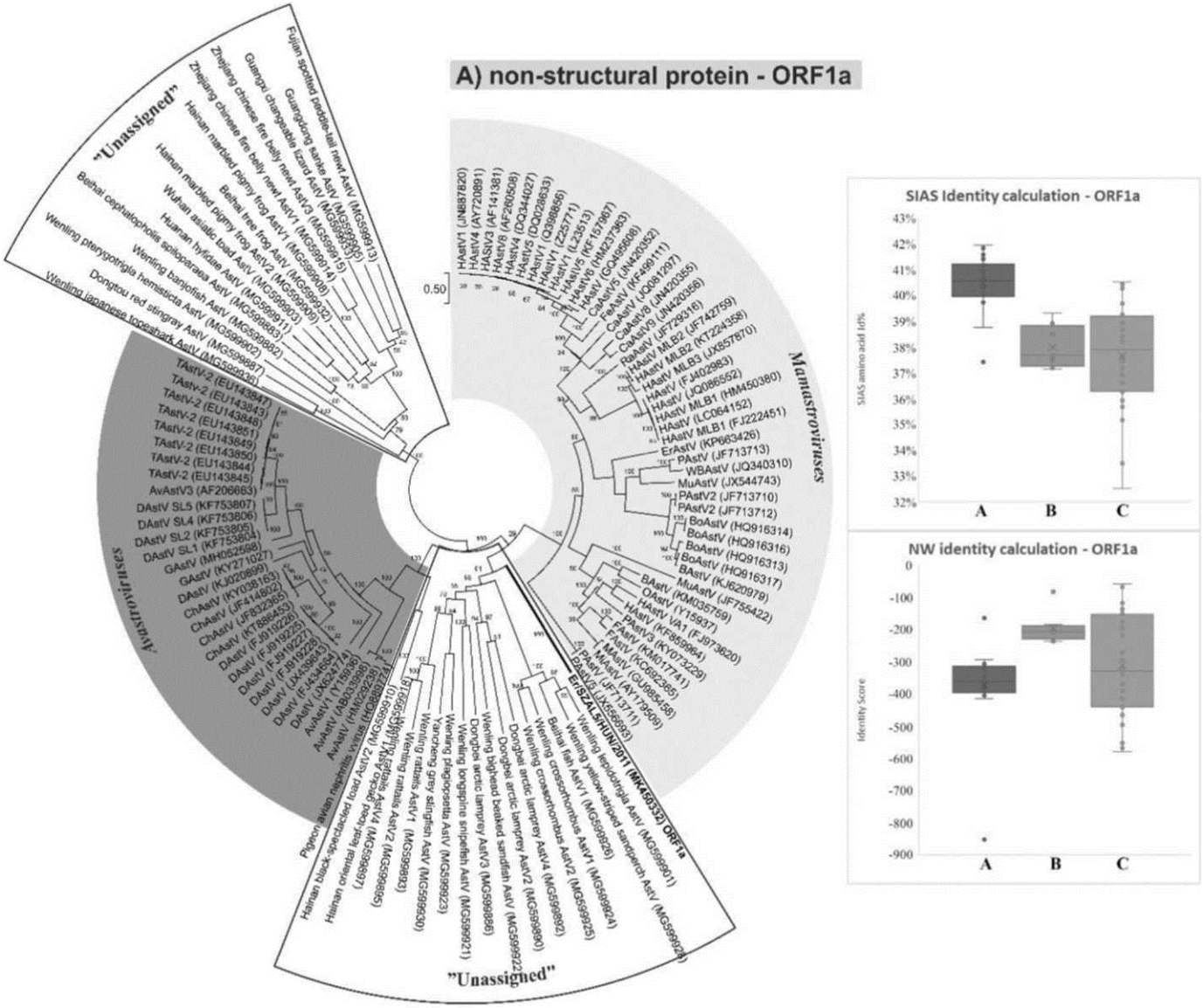

Fig. 2. Unrooted phylogenetic trees based on the complete aa sequences of A) ORF1a and B) ORF1b non-structural proteins of astroviruses and C) ORF2 capsid proteins of astroviruses and representative members of hepatitis E virus strains with the corresponding protein of the study strain Er/SZAL5/HUN/2011 (MK450332) (bolded). Species within the family *Astroviridae* including the genus *Mamastrovirus, Avastrovirus,* unclassified astroviruses (Shi et al., 2018) and species from family *Hepeviridae* (Reuter et al., 2018) are involved to the analysis. Dendrograms were constructed by the Maximum Likelihood method using MEGA7 based on the JTT with Freqs (+ F) model (Jones et al., 1992; Kumar et al., 2016) with bootstrap values and 1000 replicates. The box-plots have represented the summary of SIAS and NW identity calculations. Estimates of percentage (%) and identity scores of amino acids between astrovirus strain Er/SZAL5/HUN/2011 (MK450332) and representative members of astroviruses and hepeviruses (Suppl. 3 Table 1). The highest amino acid percentage (%) and positive scores mean the closest relative. A: *Mamastroviruses;B: Avastroviruses;* C: unclassified astroviruses; D: hepeviruses; E: unclassified hepeviruses.

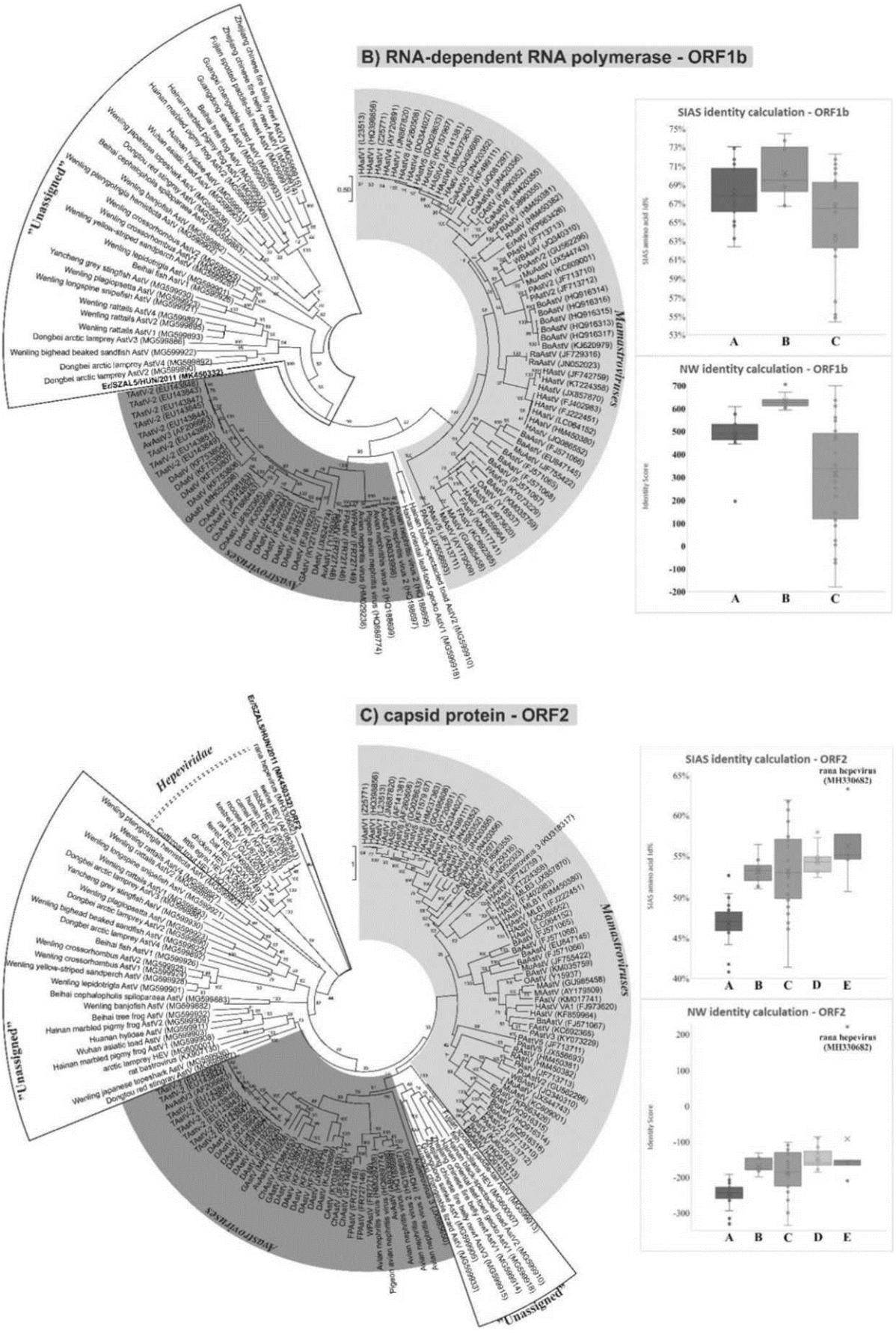

Fig. 2. *(continued)*

The disorder prediction of the ORF1a polyprotein showed 20 disordered regions, but the most probable/longest disordered region was 171aa-long at aa 946 to 1116 by FoldIndex (Suppl. 1 Fig. 3). The possible nuclear localization signal (K_{963}GKTKKGRGRKHGNRPLTVKNVKKTR) (Suppl. 1 Fig. 4) and the putative VPg polyprotein were also identifiable (Fig. 1, Suppl. 1 Fig. 3). The possible N-terminal cleavage site of the predicted VPg was at aa E**960**/E**961,** but the C-terminal cleavage could not be reliably detected based on the multiple sequence alignment. There were seven Y (Tyr995, 1026, 1040, 1053, 1058, 1135, 1154) residues altogether, but based on the nt alignment and similarity comparison, the Tyr995 was the most probable as the part of the highly conserved (D/E)EY aa motif (Fig. 1, Suppl. 1 Fig. 5). The conserved aa motifs of oligomerization, polymerase interacting domains and the potential hypervariable region (HVR) could not be identified because of the very low sequence identity (Guix et al., 2008; Fuentes et al., 2011). Altogether five coiled-coil structures were detected at aa 821-841, 923-943, 981-1001, 1179-1206 and 1221 -1241 positions. The conserved heptamer frame-shifting signal plays an essential role in switching frames +1/ —1 in protein translation (Jiang et al., 1993). Usually, the frameshift signal was followed by a stem-loop secondary RNA structure in mamas-troviruses then the ORF1a stop codon is placed, finally the ORF1b initiation codon (Lewis and Matsui, 1995, 1997; Guix et al., 2013). In contrast, the pattern is modified in avastroviruses, the frameshift signal was followed by the ORF1a stop codon, then the stem-loop secondary RNA structure, finally the ORF1b initiation codon (Suppl. 1 Fig. 6) (Lewis and Matsui, 1995, 1997; Guix et al., 2013). The potential ribo-somal frameshift heptamer sequence (AAAAAAC) was observed at nt 4409 positions (Fig. 1)(Jiang et al., 1993), but searching for one of the patterns mentioned before were not successful.

The ORF1b is out-of-frame and has no overlapping regions with ORF1a. The putative AUG initiation codon of ORF1b is at nt position 4715-4717 (AAA**A$_{4715}$UG**C). The ORF1b protein of strain Er/SZAL5/HUN/2011 showed 44% aa identity to the corresponding protein of turkey astrovirus (EU143849) with E-values $5e^{[-87-100]}$. Because of the high (95%) query coverage, Needleman-Wunsch calculation was not applied. The ORF1b encoded polyprotein analysed by CBS, contains a putative RdRp with conserved aa motifs: F_{95}PKQEL, D_{174}WTRYD, G_{235}PSG and Y_{285}GND (Fig. 1)(Koonin, 1991; Méndez et al., 2013; Marchler-Bauer et al., 2017).

The astrovirus ORF2 capsid proteins are translated from a sub-genomic RNA (sgRNA) using a promoter RNA sequence UUUGGAGN GGNGGACCNAAN$_{4-11}$AUGNC, (N: any of four nucleotides, underlined: ORF2 initiation codon) (Jonassen et al., 2003). This promoter sequence could not be identified in the study virus. The putative ORF2 protein starts at nt position 6299-6301 (AAA**A$_{6299}$UG**U) and no overlap region detected with ORF1b. Analysing the putative ORF2 capsid protein by conserved domain search no significant result could be identifiable and the protein does not show the characteristics of astrovirus capsid protein (Arias and DuBois, 2017). But using blast(p) algorithm it showed 26% aa sequence identity (E-value: $3e^{-8}$, query cover: 80%) to the corresponding protein of novel Rana hepevirus (MH330682) as

single significant sequence hit in GenBank database (Suppl. 1 Fig. 7) (Reuter et al., 2018). Analysing of these two aa sequences with the Needleman-Wunsch method, the study strain shares 23% identity in total (Suppl. 3 Table 1, Suppl. 1 Fig. 7). Based on this similarity, the highly conserved RNA motif (AUG**AAUAAC**AUG) of the cis-reactive (CRE) element functioning as the sgRNA promoter found in hepeviruses (Graff et al., 2005; Cao et al., 2010; Ding et al., 2018) was not identifiable in the corresponding region of the study strain. Searching for the second CRE in the junction at the 3'end of the genome was unsuccessful. However, the potential two internal stem-loops (ISL1 & ISL2) are found based on the alignment and secondary RNA structure comparison of the study strain and the corresponding nucleotide sequence of hepatitis E virus strain SAR-55 (M80581) (Suppl. 1 Fig. 8) (Emerson et al., 2013).

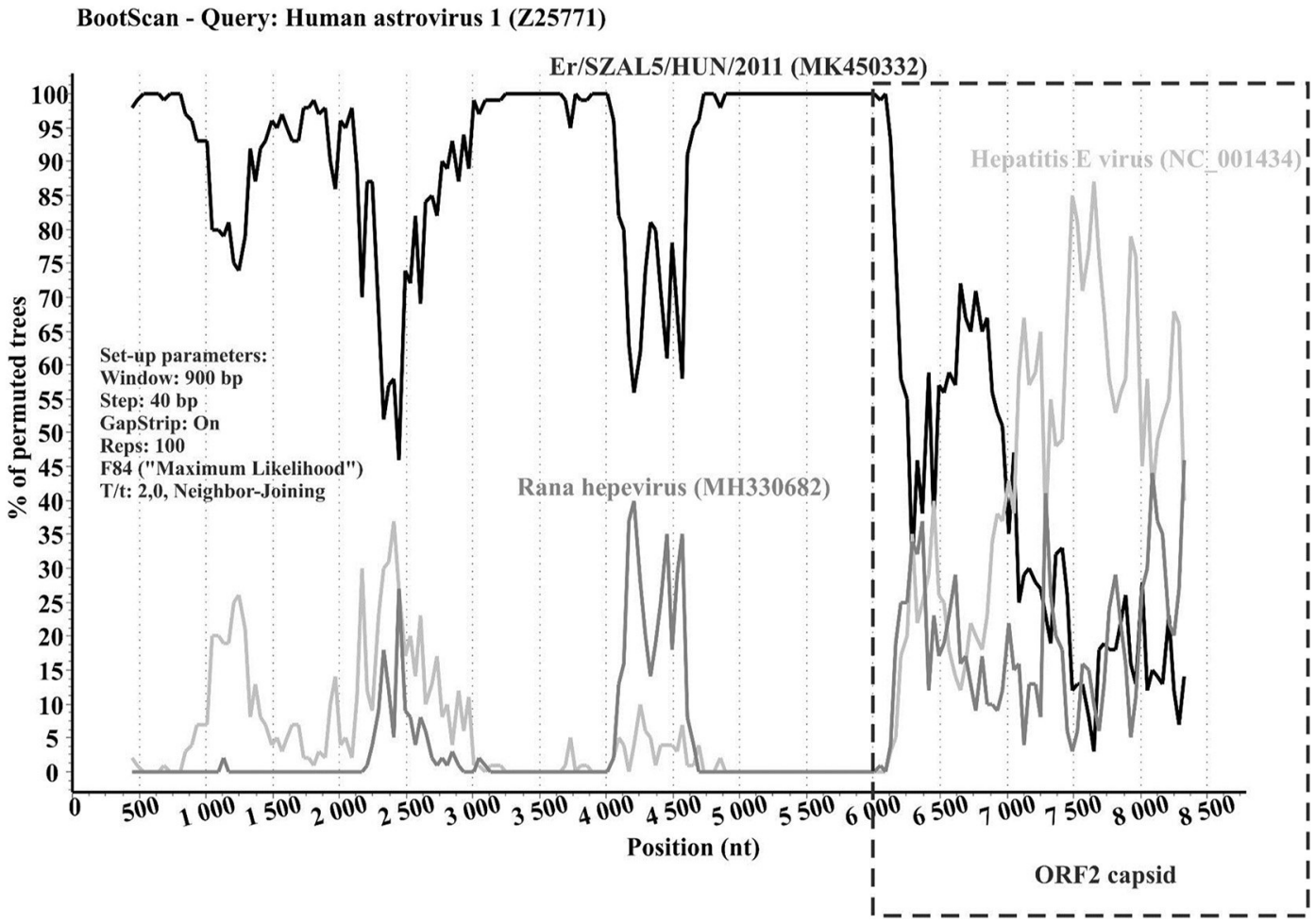

Fig. 3. The bootscan analysis of the study strain Er/SZAL5/HUN/2011 (MK450332) by SimPlot (Loleetal., 1999). The human astrovirus 1 (Z25771) is selected as the query, the hepatitis E virus (NC001434) and rana hepevirus (MH330682) are selected as reference sequences in the analysis. The striped line highlights the ORF2 capsid protein region where the potential recombination event has occurred.

The s2 m (stem-loop II) mobile genetic element in the 3'UTR was originally identified in some members of astroviruses, but lately this conserved structure was identified four other different virus families (Jonassen et al., 1998; Kofstad and Jonassen, 2011; Tengs et al., 2013). However, this conserved genetic element present in viruses representing different virus families, was neither identifiable in the study strain nor the closely related Rana hepevirus (MH330682) strain (data not shown).

Phylogenetic analysis based on the complete aa sequences of the ORF1a and ORF1b proteins showed that Er/SZAL5/HUN/2011 was separated from the known astroviruses and distantly clustered with novel astro-like viruses identified from amphibian and fish hosts. According to the result of ORF2 aa sequence phylogeny, the study strain clustered with the recently discovered novel Rana hepevirus presumably as a novel species in family *Hepeviridae* (Fig. 2; Suppl. 3 Table 1). The bootscan analysis of the study strain shows that the strain Er/SZAL5/HUN/2011 (MK450332) has an astrovirus characteristic in the ORF1a and ORF1b region and has Rana hepevirus characteristic in the ORF2 capsid region (Fig. 3, Suppl. 3 Table 1).

Using the specific screening primer pairs 3 (8.5%) out of 35 European roller faecal samples were RT-PCR-positive for the study strain. All amplicons were confirmed by nt sequencing and showed more than 91% nt identities. Four aa changes were confirmed sequence identity to each other in the amplified partial (527 nt long) ORF2 capsid region. Searching for the study strain in agile frog tadpole *(Rana dal-matind),* common toad *(Bufo bufo),* smooth newt *(Lissotriton vulgaris)* faecal samples and mosquito pools was done, but no specific PCR-product was obtained.

4. Discussion

The members of bird family *Coraciiformes* are one of the most colourful birds in Europe living in open fields, shrubby or gallery forests (Kiss et al., 2014, 2016). Based on our previous studies the European roller *(Coracias garrulus)* is prominent virus "collector and spreader" species presumably due to its predatory lifestyle. This species eats mostly insects, but consumes every small vertebrate, like amphibians, reptilians or even fish (del Hoyo et al., 1994), hence, many types of viruses are able to replicate in different target cells simultaneously, supporting the chance of the virus-host co-divergence, virus evolution or "multiple" virus-virus recombination between homologue RNA virus genome (Ghedin et al., 2009; Shi et al., 2018). Recently, several novel astrovirus-like and hepatitis E virus-like RNA virus has been discovered from lower class vertebrates mostly fish and amphibian species (Shi et al., 2018; Reuter et al., 2018). This raises the possibility of the dietary-originated virus detection hypothesis. To prove this presumption, the study strain was searched for in selected mosquito pools, common toad, agile frog and smooth newt faecal samples, but unfortunately, no specific hits were found.

The novel +ssRNA virus genome, which was serendipitous identified in three European rollers, is related to astroviruses (family *Astroviridae)* in the non-structural genome regions (ORF1a and ORF1b) and related

to an unclassified hepevirus (family *Hepeviridae)* in the capsid (ORF2) protein representing a novel, potentially recombinant virus species. In spite of the study strain Er/SZAL5/HUN/2011 has astrovirus-like genome characteristics based on the comparative analysis of the non-structural polyprotein (ORF1ab) using extensive bioinformatics methods, but genetically far distant from both *Avastrovirus* and *Mamastrovirus* species. Unfortunately, neither the aa identity scores nor the phylogenetic analysis gives convincing evidence about the closest relative(s) of the novel astrovirus strain Er/SZAL5/
HUN/2011.

The low level of identity in the conserved non-structural (ORF1a-ORF1b) proteins reduced the likelihood of finding any significant match with the characteristics of astrovirus-(like) capsid proteins. Unfortunately, the sgRNA promoter signal could not be identified upstream the ORF2 initiation codon and neither the conserved nor the variable elements could be identified in the capsid protein and the study strain has no stem-loop-II-motif. In addition, there were no significant hit in the GenBank database blasting the putative capsid protein until the novel Rana hepevirus (MH330682) sequence has been released (Reuter et al., 2018). The homology between the study strain and Rana hepevirus could suggest an inter-viral family recombination event at the junction of non-structural (ORF1b) and structural (ORF2) genome regions and if the Rana hepevirus was considered as a species of family *Hepeviridae,* this result further strengthening the hypothesis that capsid proteins of family *Hepeviridae* partially originated from "Picorna-like" supergroup of viruses (Dryden et al., 2012.; Kelly et al., 2016). Although neither the capsid proteins of Rana hepevirus (MH330682) nor the novel astro-like virus strain Er/SZAL5/HUN/2011
have known relatives in GenBank database or close contact to species in family *Hepeviridae* (Reuter et al., 2018). Even though the several structural similarities *(cis*-reactive sites, internal stem-loops) and identical RNA motifs are present in hepevirus capsid proteins the percent identity are changing from 49% aa to 57% aa among species *Orthohe-peviridae* (Kelly et al., 2016). Kelly and colleagues have reported that the non-hepevirus, chicken astrovirus capsid (JN582319) protein with 25% aa identity (over 278 amino acids) was the most significant homologue sequence to the prototype HEV GI capsid protein (NP_056788). According to us this identity level is needed to be decreased to 18% aa identity based on the global Needleman-Wunsch alignment method. Comparing the identity level of the study strain with the corresponding protein of HEV GI capsid the identity was only 16% in aa level. The potential ISL1 and ISL2 RNA motifs identity and secondary RNA structure similarity in the study strain suggest that this protein is a viral capsid with closer evolutionary contact with hepe-viruses than astroviruses, hence, we have no information about the origin of this putative capsid protein, yet. Further experimental study is needed to locate the potential sgRNA promoter site and clarify the characteristics of the capsid protein of the novel RNA virus strain Er/SZAL5/HUN/2011.

The genome sequence of Er/SZAL5/HUN/2011 is a potential recombinant chimera and its taxonomic position in family *Astroviridae* and/or *Hepeviridae* needs further discussions. Considering the phylo-genetic

position of the study strain and the novel astrovirus-like viruses published by Shi and his co-workers and hepevirus-like viruses published by Reuter and his co-workers, could evolve one or more phylo-genetic lineages, and forming novel taxa containing "herpeto" astro-viruses and a novel lineage in hepeviruses, but the classification of these viruses is still under discussion and the privilege of the ICTV taxonomic study group.

5. Conclusion

Until now, several novel RNA viruses have been found in faeces of the predator European roller; however, the dietary or host species origins are remains unknown. This study reports the complete genome characterization of a unique +ssRNA virus with ambivalent genome organization in the non-structural and capsid protein regions in faecal samples of European roller. This work supports further information to understand the genetic diversity of intermediates and genetic recombination between astroviruses and hepatitis E viruses, and might establish an understandable phylogenetic reconstruction of these highly diverse RNA viruses.

Acknowledgements

This work was supported by a grant from the Hungarian Scientific Research Fund (OTKA/NKFIH K111615) and by NHLBI R01-HL105770.

We thank Attila Hunor Török, Béla Tokody (supported by the Conservation Management and Animal Health Monitoring of Natura 2000 Bird Species HU-SRB IPA CBS PROJECT) and Edit Pollák for the collection of faecal samples. P.P. and Á.B. were supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https:// doi.org/10.1016/j.meegid.2019.04.003.

References

Arias, C.F., DuBois, R.M., 2017. The astrovirus capsid: a review. Viruses 9, 15. https://doi.org/10.3390/v9010015.

Bellaousov, S., Kayedkhordeh, M., Peterson, R.J., Mathews, D.H., 2018. Accelerated RNA secondary structure design using preselected sequences for helices and loops. RNA 24, 1555-1567. https://doi.org/10.1261/rna.066324.118.

Boros, Á., Pankovics, P., Simmonds, P., Reuter, G., 2011. Novel positive-sense, single-stranded RNA (+ssRNA) virus with dicistronic genome from intestinal content of freshwater carp *(Cyprinus carpio).* PLoS One, e29145. https://doi.org/10.1371/journal.pone.0029145.

Boros, Á., Kiss, T., Kiss, O., Pankovics, P., Kapusinszky, B., Delwart, E., Reuter, G., 2013. Genetic characterization of a novel picornavirus distantly related to the marine mammal-infecting aquamaviruses in a long-distance migrant bird species, European roller *(Coracias garrulus).* J. Gen. Virol. 94, 2029-2035. https://doi.org/10.1099/vir.0.054676-0.

Bosch, A., Pintó, R.M., Guix, S., 2014. Human astroviruses. Clin. Microbiol. Rev. 27, 1048-1074. https://doi.org/10.1128/CMR.00013-14. Cao, D., Meng, X.J., 2012. Molecular biology and replication of hepatitis E virus. Emerg Microbes and Infect 1, e17. https://doi.org/10.1038/emi.2012.7.

Cao, D., Huang, Y.W., Meng, X.J., 2010. The nucleotides on the stem-loop RNA structure in the junction region of the hepatitis E virus genome are critical for virus replication. J. Virol. 84, 13040-13044. https://doi.org/10.1128/JVI.01475-10.

Chu, D.K.W., Poon, L.L.M., Guan, Y., Peiris, J.S.M., 2008. Novel astroviruses in insectivorous bats. J. Virol. 82, 9107-9114. https://doi.org/10.1128/JVI.00857-08.

del Hoyo, J., Elliott, A., Sargatal, J., 1994. Handbook of the birds of the world. In: Mousebirds to Hornbills. vol. 6 Lynx Edicions, Barcelona.

Ding, Q., Nimgaonkar, I., Archer, N.F., Bram, Y., Heller, B., Schwartz, R.E., Ploss, A., 2018. Identification of the intragenomic promoter controlling hepatitis E virus sub-genomic RNA transcription. MBio 9https://doi.org/10.1128/mBio.00769-18.e00769-18.

Dryden, K.A., Tihova, M., Nowothny, N., Matsui, S.M., Méndez, E., Yeager, M., 2012. Immature and mature human astrovirus: structure, conformational changes, and similarities to hepatitis E virus. J. Mol. Biol. 422, 650-658. https://doi.org/10.1016/]'. jmb.2012.06.029.

Emerson, S.U., Zhang, M., Meng, X.J., Nguyen, H., St. Claire, M., Govindarajan, S., Huang, Y.K., Purcell, R.H., 2001. Recombinant hepatitis E virus genomes infectious for primates: importance of capping and discovery of a cis-reactive element. Proc. Natl. Acad. Sci. U. S. A. 98, 15270-15275. https://doi.org/10.1073/pnas.251555098.

Emmerson, S.U., Nguyen, H.T., Torian, U., Mather, K., Firth, A.E., 2013. An essential RNA element resides in a central region of hepatitis E virus ORF2. J Gen. Virol. 94, 1468-1476. https://doi.org/10.1099/vir.0.051870-0.

Firth, A.E., Atkins, J.F., 2010. Candidates in astroviruses, seadornaviruses, cytorhabdo-viruses and coronaviruses for +1 frame overlapping genes accessed by leaky scanning. Virology 7, 17. https://doi.org/10.1186/1743-422X-7-17.

Fry, H., Boesman, P., Kirwan, G.M., Sharpe, C.J., 2018. European roller *(Coracias gar-rulus).* In: del Hoyo, J., Elliott, A., Sargatal, J., Christie, D.A., de Juana, E. (Eds.), Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona.

Fuentes, C., Guix, S., Bosch, A., Pintó, R.M., 2011. The C-terminal nsp1a protein of human astrovirus is a phosphoprotein that interacts with the viral polymerase. J. Virol. 85, 4470-4479. https://doi.org/10.1128/JVI.01515-10.

Ghedin, E., Fitch, A., Boyne, A., Griesemer, S., DePasse, J., Bera, J., Zhang, X., Halpin, R.A., Smit, M., Jennings, L., George, K.St., Holmes, E.C., Spiro, D.J., 2009. Mixed infection and the genesis of influenza virus diversity. J. Virol. 83, 8832-8841. https://doi.org/10.1128/JVI.00773-09.

Graff, J., Nguyen, H., Yu, C., Elkins, W.R., St Claire, M., Purcell, R.H., Emerson, S.U., 2005. The open reading frame 3 gene of hepatitis E virus contains a cis-reactive element and encodes a protein required for infection of macaques. J. Virol. 79, 6680-6689. https://doi.org/10.1128/JVI.79.11.6680-6689.2005.

Graff, J., Torian, U., Nguyen, H., Emerson, S.U., 2006. A Bicistronic mRNA encodes both the ORF2 and ORF3 proteins of hepatitis E virus. J. Virol. 80, 5919-5926.

Guix, S., Caballero, S., Fuentes, C., Bosch, A., Pintó, R.M., 2008. Genetic analysis of the hypervariable region of the human astrovirus nsp1a coding region: design of a new RFLP typing method. J.Med. Virol. 80, 306-315. https://doi.org/10.1002/jmv. 21058.

Guix, S., Bosch, A., Pintó, R.M., 2013. Chapter 6. Astrovirus taxonomy. In: Schultz-Cherry, S. (Ed.), Astrovirus Research: Essential Ideas, Everyday Impacts, Future Directions. Springer Science + Business Media, New York, pp. 97-119. ISBN: 978-14614-4734-4. https://doi.org/10.1007/978-1-4614-4735-1.

Jiang, B., Monroe, S.S., Koonin, E.V., Stine, S.E., Glass, R.I., 1993. RNA sequence of astrovirus: distinctive genomic organization and a putative retrovirus-like ribosomal frameshifting signal that directs the viral replicase synthesis. Proc. Natl. Acad. Sci. U. S. A. 90, 10539-10543. https://doi.org/10.1073/pnas.90.22.10539.

Jonassen, C.M., Jonassen, T.O., Grinde, B., 1998. A common RNA motif in the 3' end of the genomes of astroviruses, avian infectious bronchitis virus and an equine rhino-virus. J. Gen. Virol. 79, 715-718. https://doi.org/10.1099/0022-1317-79-4-715.

Jonassen, C.M., Jonassen, T.0., Sveen, T.M., Grinde, B., 2003. Complete genomic sequences of astroviruses from sheep and turkey: comparison with related viruses. Virus Res. 91, 195-201. https://doi.org/10.1016/S0168-1702(02)00269-1.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275-282. https://doi.org/ 10.1093/bioinformatics/8.3.275.

Kelly, A.G., Netzler, N.E., White, P.A., 2016. Ancient recombination events and the origins of hepatitis E virus. BMC Evol. Biol. 16 (210). https://doi.org/10.1186/s12862-016-0785-y.

Kiss, O., Elek, Z., Moskát, Cs, 2014. High breeding performance of European rollers Coracias garrulus in heterogeneous farmland habitat in southern Hungary. Bird Study 61, 496-505. https://doi.org/10.1080/00063657.2014.969191.

Kiss, O., Tokody, B., Deák, B., Moskát, Cs, 2016. Increased landscape heterogeneity supports the conservation of European rollers *(Coracias garrulus)* in southern Hungary. J. Nature Conserv. 29, 97-104. https://doi.org/10.1016/j.jnc.2015.12.003.

Knowles, N.J., Hovi, T., Hyypä, T., King, A.M.Q., Lindberg, A.M., Pallansch, M.A., Palmenberg, A.C., Simmonds, P., Skern, T., et al., 2012. Family - astroviridae. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, San Diego, pp. 953-959. ISBN 9780123846846. https://doi.org/10.1016/B978-0-12-384684-6.00081-1.

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. https://doi.org/10.1371/journal.pone. 0025964.

Koonin, E.V., 1991. The phylogeny of RNA-dependent RNA polymerases of positivestrand RNA viruses. J. Gen. Virol. 72, 2197-2206. https://doi.org/10.1099/0022-1317-72-9-2197.

Koonin, E.V., Wolf, Y.I., Nagasaki, K., Dolja, V.V., 2008. The big bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. Nature Reviews Microbiol 6, 925-939. https://doi.org/10.1038/nrmicro2030.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870-1874. https://doi.org/10.1093/molbev/msw054.

Lewis, T.L., Matsui, S.M., 1995. An astrovirus frameshift signal induces ribosomal frameshifting in vitro. Arch. Virol. 140, 1127-1135. https://doi.org/10.1007/BF01315421.

Lewis, T.L., Matsui, S.M., 1997. Studies of the astrovirus signal that induces (—1) ribo-somal frameshifting. Adv. Exp. Med. Biol. 12, 323-330. https://doi.org/10.1007/ 978-1-4899-1828-4_53.

Liu, Y.G., Chen, Y., 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. BioTechniques 43, 649-650. https://doi. org/10.2144/000112601.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. 73, 152-160.

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y., Bryant, S.H., 2017. CDD/SPARCLE: functional

classification of proteins via subfamily domain architectures. Nucleic Acids Res. 4, D200-D203. https://doi.org/10.1093/ nar/gkw1129.

Méndez, E., Murillo, A., Velázquez, R., Burnham, A., Arias, C.F., 2013. Chapter 2. Replication cycle of astroviruses. In: Schultz-Cherry, S. (Ed.), Astrovirus Research, Essential Ideas, Everyday Impacts, Future Directions. Springer Science + Business Media, New York, pp. 19-47. ISBN: 978-1-4614-4734-4. https://doi.org/10.1007/ 978-1-4614-4735-1.

Meng, X.J., Anderson, D.A., Arankelle, V.A., Emerson, S.U., Harrison, T.J., Jameel, S., Okamoto, H., King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., 2012. Family - *Hepeviridae.* Virus taxonomy, classification and nomenclature of viruses. In: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, pp. 1021-1028 (ISBN 9780123846846; 10.1016/B978-0-12-384684-6.00087-2).

Monceyron, C., Grinde, B., Jonassen, T.O., 1997. Molecular characterization of the 3'-end of the astrovirus genome. Arch. Virol. 142, 699-706. https://doi.org/10.1007/ s007050050112.

Monroe, S.S., Jiang, B., Stine, S.E., Koopmans, M., Glass, R.I., 1993. Subgenomic RNA sequence of human astrovirus supports classification of Astroviridae as a new family of RNA viruses. J. Virol. 67, 3611-3614.

Nguyen, Ba, Pogoutse, A., Provart, N., Moses, A.M., 2009. NLStradamus: a simple hidden Markov model for nuclear localization signal prediction. BMC Bioinformatics 29, 202.

Nicolas, K.B., Nicolas, H.B., 1997. GeneDoc: A Tool for Editing and Annotating Multiple Sequence Alignments. www.psc.edu/biomed/genedoc.

Omasits, U., Ahrens, C.H., Müller, S.,Wollscheid, B., 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30, 884-886. https://doi.org/10.1093/bioinformatics/btt607.162.

Pankovics, P., Boros, Á., Kiss, T., Reuter, G., 2015a. Identification and complete genome analysis of kobuvirus in faecal samples of European roller *(Coracias garrulus):* for the first time in a bird. Arch. Virol. 160, 345-351. https://doi.org/10.1007/s00705-014-2228-7.

Pankovics, P., Boros, Á., Kiss, T., Delwart, E., Reuter, G., 2015b. Detection of a mammalian-like astrovirus in bird, European roller *(Coracias garrulus).* Infect. Genet. Evol. 34, 114-121. https://doi.org/10.1016/j.meegid.2015.06.020.

Pankovics, P., Boros, Á., Tóth, Z., Phan, T.G., Delwart, E., Reuter, G., 2017. Genetic characterization of a second novel picornavirus from amphibian host, smooth newt *(Lissotriton vulgaris).* Arch. Virol. 162, 1043-1050. https://doi.org/10.1007/s00705-016-3198-8.

Pantin-Jackwood, M., Todd, D., Koci, M.D., 2013. Chapter 9. Avian astroviruses. In: Schultz-Cherry, S. (Ed.), Astrovirus Research, Essential Ideas, Everyday Impacts, Future Directions. Springer Science +

Business Media, New York, pp. 151-180. ISBN: 978-1-4614-4734-4. https://doi.org/10.1007/978-1-4614-4735-1.

Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., Finn, R.D., 2018. HMMER web server: 2018 update. Nucleic Acids Res. 46, 200-204. https://doi.org/10.1093/nar/ gky448.

Prilusky, J., Felder, C.E., Zeev-Ben-Mordehai, T., Rydberg, E.H., Man, O., Beckmann, J.S., Silman, I., Sussman, J.L., 2005. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics 21, 3435-3448. https:// doi.org/10.1093/bioinformatics/bti537.

Purdy, M.A., Tam, A.W., Huang, C.C., Yarbough, P.O., Reyes, G.R., 1993. Hepatitis E virus: a non-enveloped member of the "alpha-like" RNA virus supergroup? Semin. Virol. 4, 319-326. https://doi.org/10.1006/smvy.1993.1029.

Reuter, G., Boros, Á., Kiss, T., Delwart, E., Pankovics, P., 2014. Complete genome characterization of mosavirus (family *Picornaviridae)* identified in droppings of a European roller *(Coracias garrulus)* in Hungary. Arch. Virol. 159, 2723-2729. https:// doi.org/10.1007/s00705-014-2113-4.

Reuter, G., Boros, Á., Tóth, Z., Phan, T.G., Delwart, E., Pankovics, P., 2015. A highly divergent picornavirus in an amphibian, the smooth newt *(Lissotriton vulgaris).* J. Gen. Virol. 96, 2607-2613.

Reuter, G., Boros, Á., Pál, J., Kapusinszky, B., Delwart, E., Pankovics, P., 2016. Detection and genome analysis of a novel (Dima)rhabdovirus (Riverside virus) from *Ochlerottatus sp.* mosquitoes in Central Europe. Infect. Genet. Evol. 39, 336-341.

Reuter, G., Boros, Á., Tóth, Z., Kapusinszky, B., Delwart, E., Pankovics, P., 2018. Detection of a novel RNA virus with hepatitis E virus-like non-structural genome organization in amphibian, agile frog (Rana dalmatina) tadpoles. Infect. Genet. Evol. 65, 112-116. https://doi.org/10.1016/j.meegid.2018.07.029.

Shi, M., Lin, X.D., Chen, X., Tian, J.H., Chen, L.J., Li, K., Wang, W., Eden, J.S., Shen, J.J., Liu, L., Holmes, E.C., Zhang, Y.Z., 2018. The evolutionary history of vertebrate RNA viruses. Nature 556, 197-202. https://doi.org/10.1038/s41586-018-0012-7.

Smith, D.B., Simmonds, P., Jameel, S., Emerson, S.U., Harrison, T.J., Meng, X.J., Okamoto, H., Van der Poel, W.H., Purdy, M.A., 2014. Consensus proposals for classification of the family *Hepeviridae.* J Gen. Virol. 95, 2223-2232. https://doi.org/10. 1099/vir.0.000115.

Speroni, S., Rohayem, J., Nenci, S., Bonivento, D., Robel, I., Barthel, J., Luzhkov, V., Coutard, B., Canard, B., Mattev, A., 2009. Structural and biochemical analysis of human pathogenic astrovirus serine protease at 2.0 A resolution. J. Mol. Biol. 387, 1137-1152. https://doi.org/10.1016/jjmb.2009.02.044.

Sverdlov, E., Azhikina, T., 2005. Primer Walking. In: eLS. John Wiley & Sons Ltd, Chichester. http://www.els.net. https://doi.org/10.1038/npg.els.0005382.

Tengs, T., Kristoffersen, A.B., Bachvaroff, T.R., Jonassen, C.M., 2013. A mobile genetic element with unknown function found in distantly related viruses. Virol. J. 10, 132. https://doi.org/10.1186/1743-422X-10-132.

Web references

Wildscreen Arkive https://www.arkive.org/european-roller/coracias-garrulus/ (accessed 25.01.2019).

Reche, P., 2019. Sequence Identity and Similarity (SIAS). http://imed.med.ucm.es/Tools/ sias.html (accessed 25.01.2019).