Ljungan/Sebokele-like picornavirus in birds of prey, common kestrel (*Falco tinnunculus*) and red-footed falcon (*F. vespertinus*)

Péter Pankovics[a,b], Ákos Boros[a,b], Róbert Mátics[c,d], Beatrix Kapusinszky[e,f], Eric Delwart[e,f], Gábor Reuter[a,b,□]

[a] Regional Laboratory of Virology, National Reference Laboratory of Gastroenteric Viruses, ÁNTSZ Regional Institute of State Public Health Service, Pécs, Hungary

[b] Department of Medical Microbiology and Immunology, Medical Center, University of Pécs, Pécs, Hungary

[c] Department of Pathophysiology, University of Pécs Medical Center, Hungary

[d] Hungarian Nature Research Society, (HuNaReS), Ajka, Hungary

[e] Blood Systems Research Institute, San Francisco, CA, USA

[f] University of California, San Francisco, CA, USA

□ Corresponding author at: Department of Medical Microbiology and Immunology, University of Pécs, Szigeti út 12., H-7624 Pécs, Hungary.

E-mail address: reuter.gabor@gmail.com (G. Reuter).

Keywords: Picornavirus Ljungan virus Sebokele virus Parechovirus Bird

Abstract

Ljungan and Sebokele viruses are thought to be rodent-borne (picorna)viruses in the genus *Parechovirus*. Using random amplification and next generation sequencing method a novel Ljungan/Sebokele-like picornavirus was identified in birds of prey. Viral RNA was detected in total of 1 (9%) of the 11 and 2 (28.6%) of the 7 faecal samples from common kestrels and red-footed falcons in Hungary, respectively. High faecal viral RNA load (4.77×10^6 genomic copies/ml) measured by qPCR. The complete genome of picornavirus strain falcon/ HA18_080/2014/HUN (KY645497) is 7964-nucleotide (nt) long including a 867-nt 5'end and a 101-nt 3'end (excluding the poly(A)-tail). Falcon/HA18_080/2014/HUN has type-II IRES related to hunnivirus IRES, encodes a polyprotein lacking a leader protein, a VP0 maturation cleavage site and it predicted to encode three 2A proteins (2A $^{NPG\downarrow P}$, 2A $^{NPG\downarrow P}$ and 2A $^{H\text{-}Box/NC}$), two of them end with 'ribosome-skipping' sites (DxExNPG$^{\downarrow}$P). Sequence analyses indicated that the ORF1 (6996 nt) polyprotein (2331 amino acid - aa) of falcon/HA18_080/ 2014/HUN shares the highest aa identity, 59% and 57%, to the corresponding polyproteins of Ljungan and Sebokele viruses. This study reports the identification and complete genome characterization of a novel Ljungan/Sebokele-like picornavirus in faeces of birds of prey which suggests that the genetic diversity and the potential host species spectrum of Ljungan/Sebokele-like viruses in genus *Parechovirus* are wider than previously thought.

1. Introduction

Ljungan and Sebokele viruses are two members of the genus *Parechovirus* within the family *Picornaviridae*. The genus *Parechovirus* is comprised of four official species *Parechovirus A* (formerly named Human parechovirus), *Parechovirus B* (formerly named Ljungan virus), *Parechovirus C* (Sebokele virus) and *Parechovirus D* (ferret parechovirus) (www.picornaviridae.com, Adams et al., 2017). Human parechoviruses are important pathogens in humans especially in infants and young children and are associated with a wide spectrum and severity of dis- eases (Harvala et al., 2010). Ljungan and Sebokele viruses are thought to be rodent-borne viruses. Ljungan virus was first isolated from bank voles in Sweden (Niklasson and Le Duc, 1984) and it has been suggested to be the zoonotic etiological agent of myocarditis, type-1 diabetes mellitus and possibly other human diseases (Niklasson et al., 1998; Niklasson et al., 2007). Sebokele virus was originally isolated in 1972 from African wood mice (Digoutte and Germain, 1985; Joffret et al., 2013). Recently, Ljungan virus was also detected in wild birds, in gull species in Japan (Mitake et al., 2016).

In general, picornaviruses are small, non-enveloped viruses with single-stranded, positive-sense, 7.2–10.1 kb-long polyadenylated genomic RNA. Picornaviral genomes have a common organization pattern: viral polyprotein coding single open reading frame (ORF) (except for the genus *Dicipivirus* which has two ORFs separated by an untranslated region) is flanked by the highly structured 5′ and 3′ un- translated regions (UTRs) (Knowles et al., 2012). The picornavirus genomes consist of essential secondary RNA structures functioning as a *cis*-acting RNA sequences e.g. internal ribosomal entry site (IRES) (Martínez-Salas et al., 2015) and *cis*-acting replication element (*cre*) (Tolf et al., 2009). Generally, the viral polyprotein (VP) processed to form three or four capsid monomer proteins (VP0-VP3-VP1 or VP4- VP2-VP3-VP1), and at least seven non-structural proteins: 2A-2B-2C- 3A-3B-3C-3D; however, differences occur in picornaviruses of different genera. The numbers of the cleaved individual proteins can be influ- enced by the presence or absence of a leader (L) protein upstream of the capsid proteins, the cleavage of VP0 into VP4 and VP2 or the presence of multiple 2A (Knowles et al., 2012). Most of the picornaviruses have only one mature 2A, although some picornaviruses in genera *Kunsagi- virus*, *Limnipivirus*, *Parechovirus* (Ljungan and Sebokele virus), *Pasivirus*, *Potamipivirus* and *Aquamavirus* have two; *Avihepatovirus* and *Avisivirus* have three or in case of genus *Megrivirus* and the proposed genus "Aalivirus" up to four and six individual 2A proteins, respectively (http://www.picornaviridae.com), although there is some uncertainty re- garding the actual release of all of these mature 2A peptides during the viral infection cycle (Johansson et al., 2003; Tseng et al., 2007; Boros et al., 2013; Boros et al., 2014a). The presence of aphthovirus 2A–like 'ribosome-skipping' motif DxExNPG$^{\downarrow}$P (where x is any amino acid) is common and present in up to 18 picornavirus genera (http://www. picornaviridae.com). In the members of the genus *Avisivirus* in birds (Boros et al.,

2013) and genus *Limnipivirus* (e.g. bluegill picornavirus) (Barbknecht et al., 2014) have two 2A proteins, in the proposed genus "Aalivirus" in birds have four 2A proteins that end in an NPG↓P motif (Wang et al., 2014).

Birds are well known reservoirs of numerous viral pathogens such as avian influenza virus, West Nile virus and Japanese encephalitis virus in humans (Reed et al., 2003). However, compared to the ten thousands of known bird species, rather few picornaviruses have been described from avian sources, often from faecal samples, and most of these have been identified in the last decade from domestic birds (Boros et al., 2013; Boros et al., 2014b). Birds have the potential also to serve as powerful biomonitors and can be used to survey for factors that may pose both public and wildlife health concerns (Pollack et al., 2017).

In this study, a novel member of the genus *Parechovirus* was de- tected and characterized from faecal samples of birds of prey, a pi- cornavirus related to Ljungan and Sebokele viruses.

2. Materials and methods

In June 2014, faecal samples were collected from wild birds of prey, from 11 common kestrels (*Falco tinnunculus*) and 7 red-footed falcons (*F. vespertinus*) from Gara (46°03′28.77″N 19°02′38.50″E) and Kardoskút (46°50′43.30″N 20°64′32.92″E) in Hungary, respectively, and stored at −80°C. Samples were collected directly from the birds during regular bird ringing by qualified ornithologists with valid permission (National Inspectorate For Environment, Nature and Water: 14/3858–9/2012). Three falcon samples (as a sample pool) were selected for viral metagenomics analysis. Briefly, 30 v/v% PBS-diluted specimens were passed through a 0.45-μm sterile filter (Millipore) and centrifuged at 6,000 ×*g* for 5 min. Then the filtrate was treated with a mixture of DNases and RNases to digest unprotected nucleic acids (Phan et al., 2013). Viral-particle protected nu- cleic acids were extracted using QIAamp spin-column technique (Qiagen) and subjected to a viral metagenomic analysis using sequence independent random amplification (Victoria et al., 2009). Viral cDNA library was con- structed by Nextera XT DNA Library Preparation Kit (Illumina) and then sequenced on MiSeq Illumina platform according to the manufacturer's in- struction, and as described previously (Phan et al., 2013). The acquired reads were trimmed; *de-novo* assembled and analyzed using an in-house pipeline (Phan et al., 2013). The reads and contigs > 100-bp were com- pared to the GenBank protein database (BlastX). Virus family-level cate- gorization of viral metagenomic reads was based on the best BlastX-scores (*E*-value $\leq 10^{-10}$). For the verification of the metagenomic contigs and for the determination of the complete picornavirus genome direct Sanger dye-terminator sequencing was used. Picornavirus quantification is based on viral cDNA transcribed by reverse primer 5′-AACTATCTGCTCTGGAAAGGT-3′ (corresponding nucleotide (nt) positions 7716–7696 of the study strain) and amplified by primers (R/F: 5′-TTTCCATATCCAGCGGTT-3′/5′-TGTGCACCTATTTGGCTT-3′ corresponding nt positions 7495–7478 and 7391–7408 of the study strain falcon/HA18_080/2014/HUN, respectively) designed

for $3D^{Pol}$ region using a real-time PCR assay (LightCycler FastStart DNA Master SYBR Green I, Roche, Mannheim, Germany). For absolute quantification and the generation of a standard curve, a hundred-fold dilution series of silica-column (Qiagen, Hilden, Germany) purified and spectrophotometrically quantified single PCR amplicon of the picornavirus was used.

Faecal specimens from kestrels and falcons were tested by RT-PCR using specific screening primer-pairs [R: 5′-AACTATCTGCTCTG GAAAGGT-3′ corresponding nt 7716–7696 of the study strain and F: 5′-GTGGTATGCCATCTGGTGCGCCGT-3′ corresponding nt positions 7331–7354 of the study strain falcon/HA18_080/2014/HUN] designed for the conserved RNA-dependent RNA polymerase (RdRp) genome region.

All evolutionary analysis (multiple sequence alignment, best DNA model search, phylogenetic analysis) was conducted in MEGA6 (Tamura et al., 2013). The nt sequences of the study strain and representative picornaviruses were aligned based on codons and pre- tested using the best nt/amino acid(aa) model (ML) search. Dendro- grams were constructed by the Maximum Likelihood (ML) method based on the General Time Reversible model with Gamma distribution (+ G) and invariable sites (+ I).

Possible polyprotein cleavage sites were predicted using the aa se- quence alignment of Ljungan viruses (AF327921, EU854568), ferret parechovirus (KF006989), Sebokele virus (NC_021482) and the study strain and the NetPicoRNA program (Blom et al., 1996). The secondary structures of 5'UTR and *cre* were predicted (but not confirmed by bio- chemical probing) using the Mfold program (Zuker, 2003) and a 2D model was drawn using Corel Draw Graphics Suite ver. 12.

The complete nt and aa sequences of falcon picornavirus, falcon/ HA18_080/2014/HUN and partial 3D/3'UTR sequences of the related viruses have been submitted to GenBank under accession numbers KY645497-KY645499.

3. Results and discussion

A total of 4811 sequence reads (singletons and contigs) showing similarity to viruses were obtained (BLASTx cut-off E score $\leq 10^{-10}$) after de novo assembly of the initial reads from the sample pool. Detected sequences were from viruses of family *Parvoviridae* (N = 2128), *Microviridae* (N = 390), *Picobirnaviridae* (N = 287), *Iridoviridae* (N = 91), *Circoviridae* (N = 49), *Astroviridae* (N = 43), *Picornaviridae* (N = 37) and other (N = 96) or unclassified (N = 1690) virus families. The 37 picornaviral sequence reads were related to VP1 and 3D regions of Ljungan and Sebokele viruses. To characterize the complete genome of the picornavirus different sets of specific primers were designed on the basis of the metagenomic sequence reads and amplicons were sequenced directly by Sanger sequencing. 5′/3′ RACE method was used to obtain the 5′ and 3′ genome ends (Boros et al., 2011). The complete genome of picornavirus strain falcon/HA18_080/ 2014/HUN (KY645497) is 7964 nt long

including a 867-nt 5'end and a 101-nt 3'end (excluding the poly(A)-tail) (Fig. 1). The base composition of the genome was found to be 27% A, 22.7% G, 19.6% C and 30.7% U. The G + C content of the entire genome is 42.3%. Falcon/HA18_080/ 2014/HUN was predicted to have a picornavirus genome organization pattern: VPg-5′UTR$^{IRES\text{-}II}$[P1(VP0-VP3-VP1)-P2(2A^{NPGP}-2A^{NPGP}-2A$^{H\text{-}Box/NC}$-2B-2C)-P3(3A-3B^{VPg}-3C^{Pro}-3D^{Pol})-3′UTR. The coding region is 6996-nt long and it encodes a 2331-aa long polyprotein.

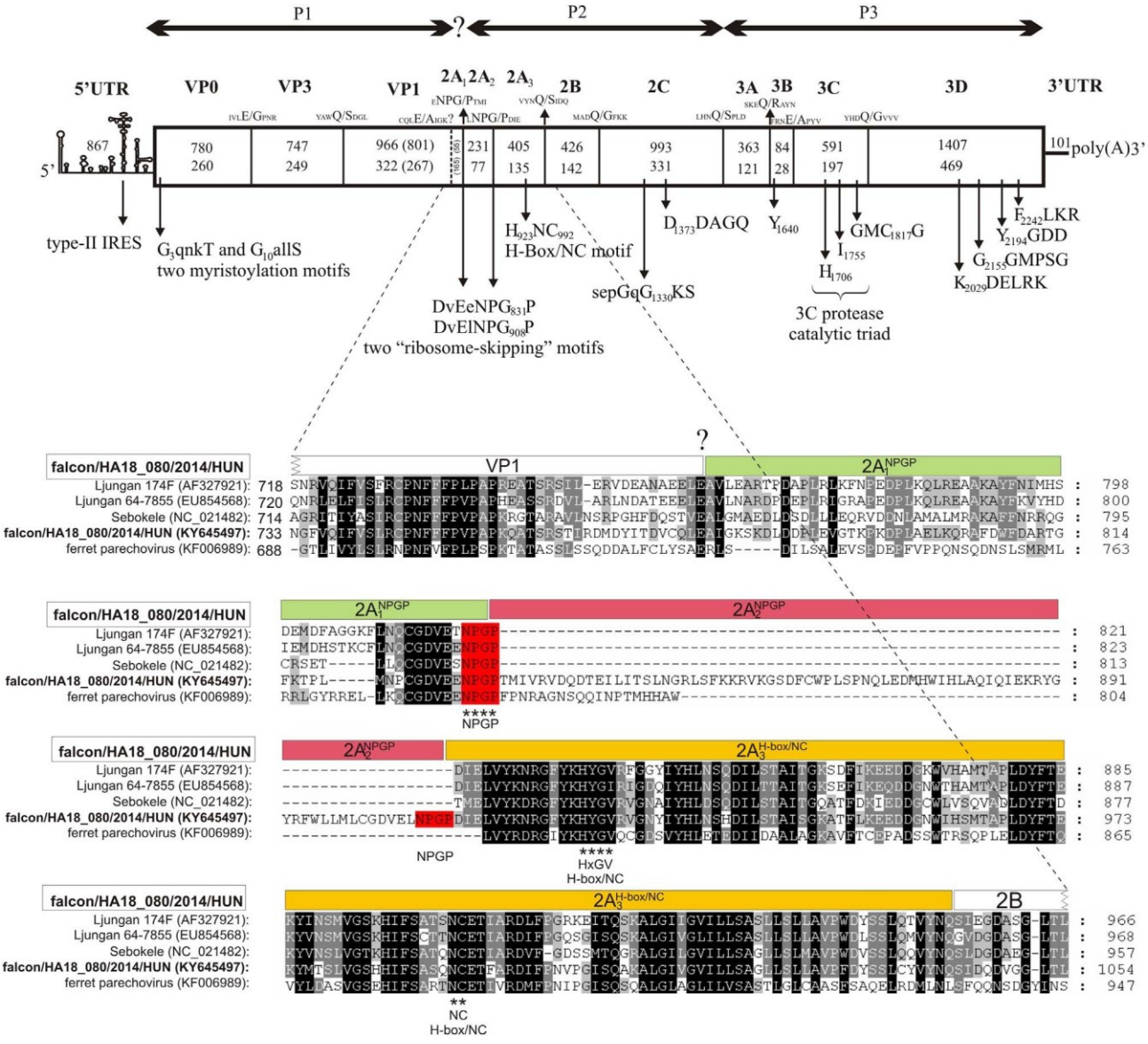

Fig. 1. Schematic genome organization of falcon/HA18_080/2014/HUN (KY645497). P1 (VP0-VP3-VP1) represents viral structural proteins and P2-P3 represent nonstructural proteins. Nucleotide (*upper number*) and amino acid (*lower number*) lengths are indicated in each gene box. Conserved picornaviral amino acid motifs and predicted P4/P4' cleavage sites are indicated at the above of genome map. Question mark indicates the uncertain border of the P1/P2. Amino acid sequence alignment of 2A protein part of two Ljungan viruses, Sebokele virus, ferret parechovirus and falcon/HA18_080/2014/HUN between the C-terminal end of the VP1 and the N-terminal end of the 2B protein is indicated.

3.1. Analysis of untranslated regions (5'UTR and 3'UTR) and cis-acting replication element (cre)

The predicted length of the 5'UTR of falcon/HA18_080/2014/HUN was 867-nt-long. The predicted initiation codon was mapped at nt po- sitions 868–870, which is in an optimal Kozak context (AaaA$_{868}$UGG) (Fig. 2a). A non-in-frame AUG(G) nts was also present 100 nt upstream (at nt positions 768–770) of the initiation codon. A pyrimidine-rich tract was found at nt positions 740–748 (Fig. 2a). Using BlastN (Gen- Bank), up to 79% nt sequence similarity was found between the 3′ end of the falcon/HA18_080/2014/HUN 5′UTR region (from nt 427 to nt 779) and the core domains I-J-K-L of internal ribosomal entry site (IRES) of members of the genera *Hunnivirus* and *Parechovirus*, especially to the rat hunnivirus (KJ950971). Based upon these data and the pre- dicted secondary RNA structure of the 5′UTR, falcon/HA18_080/2014/ HUN had a potential type-II IRES (Fig. 2a). This type of IRES comprised of five major core domains from H to L and conserved nt motifs (GNRA tetraloop, C-rich, A-bulge etc.) which were also recognizable in the study strain (Fig. 2a). The A-bulge and the discontinuous sequence element in domain J, conserved in all cardio- and aphthoviruses and critical for the interaction with the host translation initiator factor eIF4G (Bassili et al., 2004; Saleh et al., 2001), is present in falcon/ HA18_080/2014/HUN (Fig. 2a). Interestingly, and additional potential domain was also found between the domains H and I. The apical 21 nts in domain D of falcon/HA18_080/2014/HUN was completely identical to the corresponding domain of ovine and bovine hunniviruses, mosa- viruses and Ljungan virus 4 (Fig. 2a) (Reuter et al., 2012; Reuter et al., 2014). The 3'UTR is 101 nt long and similar nt sequences were not found in GenBank.

The *cis*-acting replication element (*cre*) is a short stem-loop structure including an internal or a terminal loop with three unpaired adenine (A) nucleotides. This structure is critical for picornaviral replication by its implication in the uridylylation of the viral VPg peptide (Paul et al., 2000) and has been located in different genome regions in picorna- viruses. Sequence analysis revealed that the potential *cre* is located between nt positions 5777 and 5826 in falcon/HA18_080/2014/HUN (Fig. 2b). This *cre* structure and the genome position in 3B^{VPg} region are similar to those proposed for Ljungan viruses (Al-Sunaidi et al., 2007; Tolf et al., 2009).

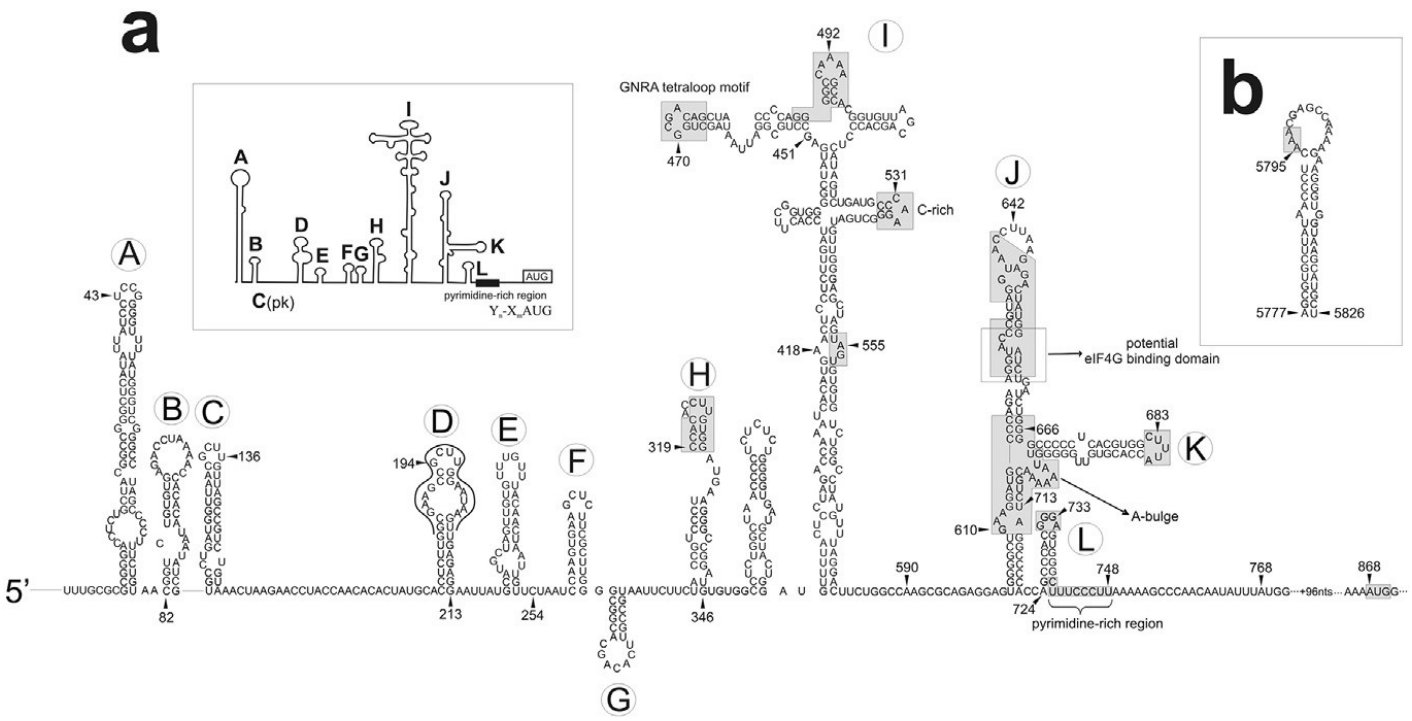

Fig. 2. a) Predicted RNA secondary structure of falcon pi- cornavirus (KY645497) 5'UTR including the internal ribosomal entry site (IRES) using the Mfold program. The complete struc- ture of the 5′UTR including the domains from A to L and the type-II IRES has been annotated as previously proposed for hunniviruses (Reuter et al., 2012) and human parechoviruses (Ghazi et al., 1998) (schematic figure in frame). The central five IRES domains are labelled from H to L to maintain the continuity of the current nomenclature. The positions of conserved type-II IRES motifs, the highly identical nucleotides to hunniviruses and human parechoviruses in IRES domains H, I, J and L, the pyrimidine-rich region at the 3′ end and the predicted polyprotein AUG start codon are indicated by shaded boxes. In domain D, continuous black line shows the position of the identical 21 nucleotides between falcon picornavirus, hunniviruses (HM153767 and JQ941880) and Ljungan virus 4 (EU854568). b) Secondary RNA structure predicted for a putative falcon picornavirus *cre* in the VPg-encoding gene (3B). The AAA sequence, which is conserved in picornavirus *cre*, is indicated by shaded box.

3.2. Analysis of the polyprotein

The possible cleavage sites of the polyprotein of falcon/HA18_080/ 2014/HUN were mapped based on (i) the aa alignment with the selected strains of the closest relatives Ljungan viruses (AF327920- AF327922, LC133331, AF538689, EU854568), Sebokele virus (NC_021482), ferret picornavirus (KF006989) and (ii) NetPicoRNA predictions. Except the border of the VP1/2A (junction of P1/P2) the polyprotein cleavage sites of falcon/HA18_080/2014/HUN were strongly supported (Fig. 1). The prediction of cleavage site of VP1/2A presented a special challenge. During the polyprotein analysis two au- tocleavage motifs ('ribosome-skipping' sites) DxExNP$^{\downarrow}$G$_{831 \text{and } 908}$/P were located at the presumed VP1/2A site (Fig. 1). Based on the pre- viously published cleavage site prediction (Johansson et al., 2003; Zhu et al., 2015) for Ljungan viruses (which viruses have only one NPG$^{\downarrow}$P motif), our hypothetically mapped VP1/2A cleavage site is located at the first DxExNP$^{\downarrow}$G$_{831}$ motif. In this case, two 2A are present in falcon/ HA18_080/2014/HUN. However, the presence of another potential cleavage site (E$_{776}$/A, which motif is also presented at presumed clea- vage site

of 3B/3C of falcon/HA18_080/2014/HUN and located at the same aa position in the alignment as found in Ljungan and Sebokele viruses) upstream from the first NPG↓P motif could not be ruled out (Fig. 1). In this case the number of presumed 2As would increase to three in falcon/HA18_080/2014/HUN (Fig. 1).

The complete P1 (2493 nt/831aa or in the case of three 2A 2328 nt/ 776aa), P2 (2055 nt/685aa or in the case of three 2A 2220 nt/740aa) and P3 (2448 nt/815aa) regions show 62%, 57% and 61% aa identity to the corresponding proteins of the closest relatives, Ljungan virus M1146 (AF538689), respectively.

The analysis of the P1 region alignments did not support the pre- sence of L protein or the cleavage of VP0 into VP4 and VP2, therefore the falcon/HA18_080/2014/HUN virions are probably built up from only three capsid monomers (VP0-VP3-VP1) (Fig. 1). Two presumed myristoylation motifs (GxxxT/S, x = variable; G_3qnkT and G_{10}allS) were found at the N-terminal of VP0 (Fig. 1). The integrin-binding RGD motif typical in some types of human parechoviruses was not found in the VP1. The VP0 (260aa), VP3 (249aa) and VP1 (322aa or in the case of three 2A 267aa) of falcon/HA18_080/2014/HUN have 68%, 66% and 50% aa identity to the corresponding proteins of Ljungan virus (EU854568), as the closest match in GenBank, respectively.

The predicted $2A_1$ is 55aa long if this region is not an extension of the C-terminal end of VP1 protein. This protein has 40% aa identity to the corresponding protein of Sebokele virus (NC_021482). The $2A_2$ is 77aa-long and it has 35% aa identity (between aa positions 22 and 72 of $2A_2$) to cytochrome P450 CYP1C1 (family1, subfamily C polypeptide 1 – DQ007044) of zebrafish, *Danio rerio*. The 135aa-long $2A_3$ of falcon/HA18_080/2014/HUN possesses the highest (77%) aa sequence iden- tity to the corresponding 2A of Ljungan virus (EU854568), contains H- box/NC regions (Fig. 1), therefore it appears to belong to the par- echovirus-like 2A proteins (Hughes and Stanway, 2000). The 142-aa- long 2B protein has 71% aa identity to the corresponding protein of Ljungan virus (AF327921). The 2C protein of falcon/HA18_080/2014/HUN possesses the highly conserved GxxGXGKS (X = uncharged, x = variable) motif for NTP binding sites with modification (sepGqG$_{1330}$KS) and the DDLxQ motif (x = variable; with modification (D$_{1373}$DagQ) for putative helicase activity (Fig. 1). The 2C and 3A proteins have 64% and 44% aa identity to 2C and 3A proteins of Se- bokele virus (NC_021482). The 28-aa-long 3B protein (small viral protein genome-linked or VPg) has Y$_{1640}$ residue at the third aa position and has 63% aa sequence identity to the VPg of Ljungan virus 145SL (FJ384560). The conservative catalytic triad (H, I, C) of the 3C viral cysteine-active-centre protease was also seen in the falcon/HA18_080/ 2014/HUN and the active site cysteine in the motif GxCG (x = variable, GMC$_{1817}$G) was also present (Fig. 1). The highly conserved motifs (e.g. KDELR, GxxPSG, YGDD and FLKR, x = variable) of the $3D^{Pol}$ (RNA- dependent-RNA-polymerase) were also presented in the study sequence (Fig. 1). The 3C and 3D proteins of the falcon/HA18_080/2014/HUN have 64%–64% aa identity to the corresponding proteins of Ljungan virus (AF538689).

The phylogenetic trees based on the P1, 2C and 3CD nt sequence of falcon/HA18_080/2014/HUN and

representative members of picorna- viruses could indicate that the study strain is a member of the parechovirus lineage containing ferret picornavirus, Ljungan virus, Sebokele virus and human parechoviruses (Fig. 3A-C).

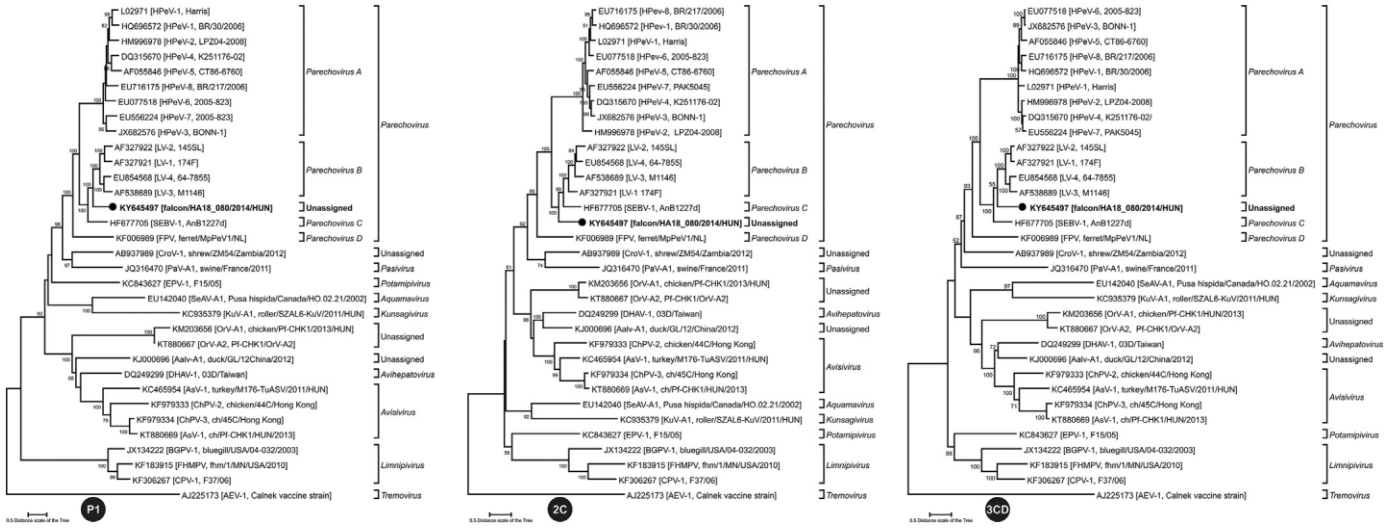

Fig. 3. Phylogenetic analysis of a novel Ljungan/Sebokele-like picornavirus strain falcon/HA18_080/2014/HUN (KY645497 in bold) and representative picornaviruses based on the complete P1, 2C and 3CD nucleotide sequences. The strain Calnek vaccine strain (AJ225173, AEV-1, genus *Tremovirus*) was used as an outgroup in each phylogenetic tree. The multiple alignment of nucleotide sequences, finding the best DNA model search and the molecular phylogenetic analysis were conducted in MEGA6. The evolutionary history was inferred using the Maximum Likelihood method based on the General Time Reversible model with discrete Gamma distribution (+ G), allowing evolutionarily invariable sites (+ I). Using all sites in the multiple alignments the bootstrap values were determined with 1000 replicates. The tree was drawn to scale with branch lengths measured in the number of substitutions per site.

Falcon/HA18_080/2014/HUN-like picornaviral RNA was detected in total of 1 (9%) of the 11 and 2 (28.6%) of the 7 faecal samples from common kestrels and red-footed falcons, respectively (the distance between the two sampling areas is 135 km away). All amplicons were confirmed by direct nucleotide sequencing and submitted to GenBank under accession numbers KY645497-KY645499. These 3 picornaviral sequences from birds of prey showed 94–98% nt sequence identity to each other in the amplified partial $3D^{Pol}$/3'UTR regions. The result of the qPCR quantification reveals the presence of 4.77×10^6 (standard deviation $\pm 5.42 \times 10^5$) genomic copies/ml in original faeces in falcon sample HA18_080.

This study reports the identification and complete genome char- acterization of a novel parechovirus related to Ljungan and Sebokele viruses from a wild bird. Based upon the sequence- and phylogenetic analyses falcon/HA18_080/2014/HUN had the highest aa identity to the Ljungan viruses in species *Parechovirus B*. However, while the four known Ljungan virus genotypes (genotypes 1–4) had only 11–21% aa difference between each other, the falcon/HA18_080/2014/HUN ORF1 polyprotein had a 41%

difference to the Ljungan viruses (and 43% to Sebokele virus). The aa difference between Ljungan viruses (species *Parechovirus B*) and Sebokele virus (species *Parechovirus C*) is 38–39% on the ORF1 polyprotein. These data suggests that falcon/HA18_080/ 2014/HUN may represent a novel parechovirus species in genus *Parechovirus*. The known geographic distribution of Ljungan virus is limited. It has been detected in Northern Europe (Sweden, Denmark, Finland), Italy, UK, the United States (Pounder et al., 2015), Germany, Thailand (Kallies, 2010) and in Japan (Mitake et al., 2016). Until now, Ljungan virus genome sequence has been found in bank voles (*Myodes glareolus*) (Niklasson et al., 1998; Johansson et al., 2002; Pounder et al., 2015), montane vole (*Microtus montanus*) (Johansson et al., 2003), red-backed vole (*Myodes gapperi*) (Tolf et al., 2009), yellow-necked mice (*Apodemus flavicollis*) (Hauffe et al., 2010), Eurasian red squirrels (*Sciurus vulgaris*) (Romeo et al., 2014), laboratory rats (*Rattus norvegicus*) (Kallies, 2010) and other rodents (Kallies, 2010). However, very recently, Ljungan virus has also been detected in wild birds, in gull (*Larus* sp.) as a host species from Japan, indicating the possibility of Ljungan virus infection and replication in birds (Mitake et al., 2016). This opens a new possi- bility of further natural host species of these viruses. It represents an interesting scientific direction, and may underline the possible importance of the Ljungan viruses, that several scientific studies have reported that the Ljungan virus is associated with serious diseases in- cluding diabetes mellitus, myocarditis, encephalitis, intrauterine fetal death in rodents, and based on this results maybe in humans (Niklasson, 2011; Zheng et al., 2015). Recent results of a seroepidemiological study also suggested a human-to-human transmission of Ljungan and - much more importantly - Ljungan-like viruses in Finland (Jääskeläinen et al., 2015). In this context, the knowledge of the distribution of these viruses among wild and domestic animals is crucial to assess their potential transmission and importance as human pathogens. In addition, the novel Ljungan-like genome sequence from different location and host help to better understand the evolutionary and epidemiological prop- erties of these potential zoonotic viruses. Finally, the wider genetic diversity of Ljungan-like viruses suggests that the sensitivity of the presently used molecular methods including PCRs (Mantke et al., 2007) is sub-optimal resulting an underreporting the incidence of these viruses.

In spite of the potential importance of Ljungan viruses there are only a total of 14 published Ljungan virus strains available in the GenBank and only two studies reported results on viral nucleic acid quantifica- tion. In experimentally infected laboratory mice 10^6–10^{10} RNA copies per gram of Ljungan virus were detected in different tissue samples (in increasing order: kidney, liver, lung, pancreas, heart and brain) when clinical signs of encephalitis appeared (Mantke et al., 2007). In another field study, Ljungan virus was detected in the liver in bank vole and yellow-necked mice with copy numbers per gram tissue ranging from 10 to 5×10^4 (Hauffe et al., 2010). There is no report on quantification of Ljungan virus in faeces; however, we found a relatively high faecal viral load (4.77×10^6 copies per ml viral RNA concentration) of falcon/HA18_080/2014/HUN in the analyzed falcon sample – com- pared to other known stool-associated

viruses in faeces. This represents a potential indication of possible, although not confirmed, viral re- plication in the avian host.

The avian host origin is also supported by the genome character- istics of falcon/HA18_080/2014/HUN. While the type-IV IRES is pre- dominant, the type-II IRES – as a potential IRES in falcon/HA18_080/2014/HUN - is also common among avian picornaviruses (Boros et al., 2014b). The other characteristic feature of avian picornaviruses is the multiple 2A proteins encoded by the genome (Boros et al., 2014b). In addition, the members of the picornaviruses with multiple 2A proteins with 'ribosome-skipping' sites (DxExNPG$^{\downarrow}$P) are grouped in one particular phylogenetic lineage (Boros et al., 2014b). In this lineage – which also contains the Ljungan and Sebokele viruses - there are two pi- cornavirus genera which encode two NPG$^{\downarrow}$Ps: bluegill picornavirus 1, carp picornavirus 1 and fathead minnow picornavirus 1 in genus *Lim- nipivirus* (Barbknecht et al., 2014) and the bird origin turkey and chicken avisiviruses in genus *Avisivirus* (Boros et al., 2013). In this phylogenetic lineage, the duck picornavirus of the proposed bird-origin genus "Aalivirus" is predicted to possess six 2A proteins, four of which have NPG$^{\downarrow}$P motifs (Wang et al., 2014). The affiliation of the first NPG$^{\downarrow}$P in picornavirus genomes with multiple NPG$^{\downarrow}$P motifs, and in consequence of the border of the VP1/2A, is an open question (Boros et al., 2013; Barbknecht et al., 2014; Wang et al., 2014). The presence of an NPG$^{\downarrow}$P motif at the C-terminal end of a capsid protein VP1 was experimentally demonstrated in Ljungan virus (Johansson et al., 2003). Without experimental results we cannot rule out the possibility of the presence of an unusual VP1/2A cleavage site upstream from the first NPG$^{\downarrow}$P motif and therefore the release of up to three different 2A polypeptides in falcon/HA18_080/2014/HUN. However, it is also pos- sible that the first 2A motif is actually a part of VP1 that does not trimmed off by 3C^{Pro}. Either way, the falcon/HA18_080/2014/HUN with two NPG$^{\downarrow}$P polyprotein motifs and more than one 2A proteins represents an interesting genetic and evolutionary variant among the known parechoviruses.

In the light of the possible association of Ljungan virus replication in bird (Mitake et al., 2016), one might hypothesize that the Ljungan and Ljungan-like viruses were adapted to infect birds as a result of the long lasting (predator/food prey) relationship between rodent hosts and carnivorous or omnivorous wild birds. Further systematic studies are needed to investigate the host species spectrum and switch, incidence and genetic diversity of Ljungan-like viruses in different bird species.

Acknowledgements

This work was supported by grants from the Hungarian Scientific Research Fund (OTKA/NKFIH K111615), by the Hungarian Nature Research Society and by NHLBI R01-HL105770. P.P and Á.B. were supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

Adams, M.J., Lefkowitz, E.J., King, A.M.Q., Harrach, B., Harrison, R.L., Knowles, N.J., Kropinski, A.M., Krupuvic, M., Kuhn, J.H., Mushegian, A.R., Nibert, M., Sabanadzovic, S., Sanfaçon, H., Simmonds, P., Varsani, A., Zerbini, F.M., Gorbelenya, A.E., Davison, A.J., 2017. Changes to taxonomy and the international code of virus classification and nomenclature ratified by the international committee on taxonomy of viruses (2017). Arch. Virol. 162, 2505–2538.

Al-Sunaidi, M., Williams, C.H., Hughes, P.J., Schnurr, D.P., Stanway, G., 2007. Analysis of a new human parechovirus allows the definition of parechovirus types and the identification of RNA structural domains. J. Virol. 81, 1013–1021.

Barbknecht, M., Sepsenwol, S., Leis, E., Tuttle-Lau, M., Gaikowski, M., Knowles, N.J., Lasee, B., Hoffman, M.A., 2014. Characterization of a new picornavirus isolated from the freshwater fish *Lepomis macrochirus*. J. Gen. Virol. 95, 601–613.

Bassili, G., Tzima, E., Song, Y., Saleh, L., Ochs, K., Niepmann, M., 2004. Sequence and secondary structure requirements in a highly conserved element for foot-and-mouth disease virus internal ribosome entry site activity and eIF4G binding. J. Gen. Virol. 85, 2555–2565.

Blom, N., Hansen, J., Blaas, D., Brunak, S., 1996. Cleavage site analysis in picornaviral polyproteins: discovering cellular targets by neural networks. Protein Sci. 5, 2203–2216.

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, Á., Nemes, C., Pankovics, P., Kapusinszky, B., Delwart, E., Reuter, G., 2013. Genetic characterization of a novel picornavirus in turkeys (*Meleagris gallopavo*) distinct from turkey galliviruses and megriviruses and distantly related to the members of the genus *Avihepatovirus*. 94. pp. 1496–1509.

Boros, Á., Pankovics, P., Knowles, N.J., Nemes, C., Delwart, E., Reuter, G., 2014a. Comparative complete genome analysis of chicken and turkey megriviruses (family *Picornaviridae*): long 3′ untranslated regions with a potential second open reading frame and evidence for possible recombination. J. Virol. 88, 6434–6443.

Boros, Á., Pankovics, P., Reuter, G., 2014b. Avian picornaviruses: molecular evolution, genome divesrity and unusual features of a rapidly expanding group of viruses in birds. Infect. Genet. Evol. 28, 151–166.

Digoutte, J.P., Germain, M., 1985. Sebokele (An B 1227d), nouvel Arbovirus non groupé.

In: Karabatsos, N. (Ed.), International Catalogue of Arboviruses Including Certain Other Virus of Vertebrates, 3rd edn. American Society of Tropical Medicine and Hygiene, San Antonio, Texas, pp. 917–918.

Ghazi, F., Hughes, P.J., Hyypiä, T., Stanway, G., 1998. Molecular analysis of human parechovirus type 2 (formerly echovirus 23). J. Gen. Virol. 79, 2641–2650.

Harvala, H., Wolthers, K.C., Simmonds, P., 2010. Parechoviruses in children: under- standing a new infection. Curr. Opin. Infect. Dis. 23, 224–230.

Hauffe, H.C., Niklasson, B., Olsson, T., Bianchi, A., Rizzoli, A., Klitz, W., 2010. Ljungan virus detected in bank voles (*Myodes glareolus*) and yellow-necked mice (*Apodemus flavicollis*) from northern Italy. J. Wildl. Dis. 46, 262–266.

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, 201–207.

Jääskeläinen, A.J., Voutilainen, L., Lehmustro, R., Henttonen, H., 2015. Serological survey in the Finnish human population implies human-to-human transmission of Ljungan virus or antigenically related viruses. Epidemiol. Infect. 144, 1278–1285.

Joffret, M.L., Bouchier, C., Grandadam, M., Zeller, H., Maufrais, C., Bourhy, H., Despres, P., Delpeyroux, F., Dacheux, L., 2013. Genome characterization of Sebokele virus 1 (SEBV1) reveals a new candidate species among the genus *Parechovirus*. J. Gen. Virol. 94, 1547–1553.

Johansson, S., Niklasson, B., Maizel, J., Gorbalenya, A.E., Lindberg, A.M., 2002. Molecular analysis of three Ljungan virus isolates reveals a new, close-to-root lineage of the *Picornaviridae* with a cluster of two unrelated 2A proteins. J. Virol. 76, 8920–8930.

Johansson, S., Niklasson, B., Tesh, R.B., Shafren, D.R., Travassos da Rosa, A.A., Lindberg, A.M., 2003. Molecular characterization of M1146, an American isolate of Ljungan virus (LV) reveals the presence of a new LV genotype. J. Gen. Virol. 84, 837–844.

Kallies, R., 2010. Ljungan Virus – Prevalence in Rodents and Virus Pathogenesis. (Dissertation) Freie Universität Berlin, Robert Koch-Insitute, Berlin(http://www.diss. fu-berlin.de/diss/servlets/MCRFileNodeServlet/FUDISS_derivate_000000012396/ Diss_RKallies_online.pdf).

Knowles, N.J., Hovi, T., Hyypiä, T., King, A.M.Q., Lindberg, A.M., Pallansch, M.A., Palmenberg, A.C., Simmonds, P., Skern, T., Stanway, G., Yamashita, T., Zell, R., 2012. Picornaviridae. 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. 855–880.

Mantke, O.D., Kallies, R., Niklasson, B., Nitsche, A., 2007. A new quantitative real-time reverse transciptase PCR assay and melting curve analysis for detection and geno- typing of Ljungan virus strains. J. Virol. Methods 141, 71–77.

Martínez-Salas, E., Francisco-Velilla, R., Fernandez-Chamorro, J., Lozano, G., Diaz- Toledano, R., 2015. Picornavirus IRES elements: RNA structure and host protein interactions. Virus Res. 206, 62–73.

Mitake, H., Fujii, Y., Nagai, M., Ito, N., Okadera, K., Okada, K., Nakagawa, K., Kishimoto, M., Mizutani, T., Okazaki, K., Sakuda, Y., Takada, A., Sugiyama, M., 2016. Isolation of a sp. nov. Ljungan virus from wild birds in Japan. J. Gen. Virol. 97, 1818–1822.

Niklasson, B., 2011. Current views on ljungan virus and its relationship to human diabetes. J. Med. Virol. 83, 1673.

Niklasson, B., Le Duc, J., 1984. Isolation of the nephropathia epidemica agent in Sweden. Lancet 323, 1012–1013.

Niklasson, B., Hörnfeldt, B., Lundman, B., 1998. Could myocarditis, insulin-dependent diabetes mellitus, and Guillain-Barré syndrome be caused by one or more infectious agents carried by rodents? Emerg. Infect. Dis. 4, 187–193.

Niklasson, B., Samsioe, A., Papadogiannakis, N., Kawecki, A., Hörnfeldt, B., Saade, G.R., Klitz, W., 2007. Association of zoonotic Ljungan virus with intrauterine fetal deaths. Birth Defects Res. A Clin. Mol. Teratol. 79, 488–493.

Paul, A.V., Rieder, E., Kim, D.W., van Boom, J.H., Wimmer, E., 2000. Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the *in vitro* uridylylation of VPg. J. Virol. 74, 10359–10370.

Phan, T.G., Vo, N.P., Boros, Á., Pankovics, P., Reuter, G., Li, O.T., Wang, C., Deng, X., Poon, L.L., Delwart, E., 2013. The viruses of wild pigeon droppings. PLoS ONE 8 (9), e72787.

Pollack, L., Ondrasek, N., Calisi, R., 2017. Urban health and ecology: the promise of an avian biomonitoring tool. Curr. Zool. 2017http://dx.doi.org/10.1093/cz/zox01. (in press).

Pounder, K.C., Watts, P.C., Niklasson, B., Kallio, E.R., Martson, D.A., Fooks, A.R., Begon, M., McElhinney, L.M., 2015. Genome characterization of two Ljungan virus isolates from wild bank voles (*Myodes glareolus*) in Sweden. Infect. Genet. Evol. 36, 156–164. Reed, K.D., Meece, J.K., Henkel, J.S., Shukla, S.K., 2003. Birds, migration and emerging zoonosis: west nile virus, lyme disease, influenza a and enteropathogens. Clin. Med. Res. 1, 5–12.

Reuter, G., Pankovics, P., Knowles, N.J., Boros, Á., 2012. Two closely related novel picornaviruses in cattle and sheep in Hungary from 2008 to 2009, proposed as member of a new genus in the family *Picornaviridae*. J. Virol. 86, 13295–13302.

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 garrulous*) in Hungary. Arch. Virol. 159, 2723–2729.

Romeo, C., Ferrari, N., Rossi, C., Everest, D.J., Grierson, S.S., Lanfranchi, P., Martinoli, A., Sino, N., Wauters, L.A., Hauffe, H.C., 2014. Ljungan virus and an adenovirus in Italian squirrel population. J. Wildl. Dis. 50, 409–411.

Saleh, L., Rust, R.C., Füllkrug, R., Beck, E., Bassili, G., Ochs, K., Niepmann, M., 2001. Functional interaction of translation initiation factor eIF4G with the foot-and-mouth disease virus internal ribosome entry site. J. Gen. Virol. 82, 757–763.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tolf, C., Gullberg, M., Johansson, E.S., Tesh, R.B., Andersson, B., Lindberg, A.M., 2009. Molecular characterization of a novel Ljungan virus (*Parechovirus*; *Picornaviridae*) reveals a fourth genotype and indicates ancestral recombination. J. Gen. Virol. 90, 843–853.

Tseng, C.H., Knowles, N.J., Tsai, H.J., 2007. Molecular analysis of duck hepatitis virus type 1 indicates that it should be assigned to a new genus. Virus Res. 123, 190–203.

Victoria, J.G., Kapoor, A., Li, L., Blinkova, O., Slikas, B., Wang, C., Naeem, A., Zaidi, S., Delwart, E., 2009. Metagenomic analysis of viruses in stool samples from children with acute flaccid paralysis. J. Virol. 83, 4642–4651.

Wang, X., Liu, N., Wang, F., Ning, K., Li, Y., Zhang, D., 2014. Genetic characterization of a novel duck-origin picornavirus with six 2A proteins. J. Gen. Virol. 95, 1289–1296.

Zheng, L., Wang, F., Huang, J., Xin, H., 2015. Evaluation of the association of zoonotic Ljungan virus with perinatal deaths and fetal malformation. Birth Defects Res. C Embryo Today 105, 81–85.

Zhu, L., Wang, X., Ren, J., Porta, C., Wenham, H., Ekström, J.O., Panjwani, A., Knowles, N.J., Kotecha, A., Siebert, C.A., Lindberg, A.M., Fry, E.E., Rao, Z., Tuthill, T.J., Stuart, D.I., 2015. Structure of Ljungan virus provides insight into genome packaging of this picornavirus. Nat. Commun. 6, 8316.

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