Short communication

# Detection of a novel RNA virus with hepatitis E virus-like non-structural genome organization in amphibian, agile frog (*Rana dalmatina*) tadpoles

Gábor Reuter[a,b,*], Ákos Boros[a,b], Zoltán Tóth[c], Beatrix Kapusinszky[d], Eric Delwart[d,e], Péter Pankovics[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 School, University of Pécs Pécs, Hungary
[c] Lendület Evolutionary Ecology Research Group, Plant Protection Institute, Hungarian Academy of Sciences, Budapest, Hungary
[d] Blood Systems Research Institute, San Francisco, CA, USA
[e] University of California, San Francisco, CA, USA

## ABSTRACT

In recent years, relatives (bastrovirus, hepelivirus) of hepeviruses (family *Hepeviridae*) have been reported in a variety of vertebrate hosts. Preliminary studies indicated that inter-viral family recombination events at the junction of the genomes that encodes non-structural (ORF1) and structural protein (ORF2) were implicated in the genesis of hepeviruses. Using viral metagenomics, next generation sequencing and RT-PCR techniques a genetically divergent hepevirus-like RNA virus was identified and characterized from agile frog (*Rana dalmatina*) tadpoles living in aquatic environment in three natural ponds (Mélymocsár, Lake Ilona and Lake Katlan) in the Pilis Mountains, in Hungary. The complete genome of the viral strain agile frog/RD6/2015/HUN (MH330682) is 7188 nt long including a 48-nt 5′ and a 122-nt 3′ non-coding region. Sequence analysis indicated that the agile frog/RD6/2015/HUN genome has potentially three non-overlapping ORFs. ORF1 (4740 nt/1579aa) has a hepevirus-like non-structural genome organization and encodes several hepevirus-like amino acid sequence motifs. The ORF2 is a potential capsid protein. The functions of the ORF3 were not predictable. The study virus was present in 18 (46%) of the 39 faecal specimen pools from agile frog tadpoles. The taxonomic position of this novel virus is presently unknown.

## 1. Introduction

Hepeviruses – including the prototype hepatitis E viruses (HEV) – are a genetically diverse group of viruses in family *Hepeviridae* detected in fish, birds and mammals. Based upon the recent taxonomic classification the family *Hepeviridae* is divided into the genera *Orthohepevirus* (mammalian and avian hepatitis E viruses) and *Piscihepevirus* (cutthroat trout virus) (Smith et al., 2014). Species within the genus *Orthohepevirus* are designated *Orthohepevirus A* (viruses from human, pig, wild boar, deer, mongoose, rabbit and camel), *Orthohepevirus B* (viruses from chickens), *Orthohepevirus C* (viruses from rat, greater bandicoot, shrew, ferret and mink), *Orthohepevirus D* (viruses from bat) (Raj et al., 2012; Drexler et al., 2012; Smith et al., 2013; Johne et al., 2014; Smith et al., 2014; ICTV, 2014) and potentially three currently unassigned species (viruses from moose, kestrel/falcon and little egret) (Lin et al., 2014; Reuter et al., 2016a; Reuter et al., 2016b). Several further unassigned hepeviruses or hepe-like viruses have been discovered recently in vertebrate and invertebrate hosts (Shi et al., 2016; Shi et al., 2018). In general, the members of the genus *Orthohepevirus A* have approximately 6.6–7.3 kb long, positive, single-stranded RNA genome and encodes three open reading frames (ORFs) flanked by a capped 5′ end and a poly (A)-tailed 3′ untranslated regions (UTRs) (Smith et al., 2014). Ortho-hepeviruses and piscihepevirus (cutthroat trout virus) have a similar genome organization to each other (Batts et al., 2011).

The family *Hepeviridae* seems to have originated from inter-viral family recombination. A recent study demonstrated that HEV non-structural protein encoded from ORF1 is related to viruses from families within the "alpha-like" supergroup, whilst the capsid encoded by ORF2 appears to be related to viruses of the *Astroviridae* family that falls within the "Picorna-like" supergroup of viruses (Kelly et al., 2016). These studies shed light on the inter-viral family recombination events at the junction of the genome that encodes non-structural and structural proteins implicated in the genesis of HEV and astroviruses (Dryden et al., 2012; Kelly et al., 2016). Strengthening this hypothesis, a novel group of viruses, tentatively named "bastroviruses" (basal astrovirus), were recently discovered in human and animal faeces and raw sewage that encodes a putative non-structural and structural proteins sharing some homology with members of the *Hepeviridae* and *Astroviridae*, respectively (Munnink et al., 2016; dos Anjos et al., 2017).

The agile frog (*Rana dalmatina*) is a member of the genus *Rana* in the family *Ranidae*. This species is widespread in much of Europe and can be found from France to northern Turkey, but is distributed

---

unevenly, being absent from large areas within its range (Kaya et al., 2009; Sillero et al., 2014). Agile frogs inhabit glades and open sites within light and relatively dry deciduous mixed forests, and breed in a variety of water bodies, ranging from small ephemeral puddles to large permanent ponds and lakes within forests and at their edges (Nöllert and Nöllert, 1992). There is very limited data related to the viruses in amphibians including frog. Virome analysis may help to identified novel, evolutionarily important, viral sequences and potential disease causing agents in this lower vertebrates.

Using random amplification of enriched viral nucleic acids, high-throughput sequencing technology and different RT-PCR techniques the complete genome of a novel RNA virus with hepevirus-like non-structural genome organization was determined and characterized in faeces from the agile frog.

## 2. Materials and methods

On the 2/06/2015, a total of 39 faecal specimen pools were collected from healthy agile frog (*Rana dalmatina*) tadpoles (Gosner stage 25–38; Gosner, 1960) living in three natural ponds (Mélymocsár: 47°42′27″N 19°02′24″E, N = 13, RD1–13; Lake Ilona: 47°42′48″N 19°02′25″E, N = 13, RD21–33 and Lake Katlan: 47°42′42″N 19°02′40″E, N = 13, RD39-51) in the Pilis Mountains, in Hungary. Tadpoles were caught using dip-netting and kept in threes in 0.2 L aeriated, aged tap water for 30–60 min in 0.5 L plastic cups; faeces from each cup was collected using a Pasteur pipette into a 1.5 mL Eppendorf tube and then stored at −80 °C. Therefore, each faecal specimen pool contained faeces from three individual tadpoles (total of 117); this sampling procedure ensured that a sufficient amount of faeces was obtained in each replicate specimen for subsequent analyses. Permission to capture the animals was issued by the national authority of the Middle-Danube-Valley Inspectorate for Environmental Protection, Nature Conservation and Water Management, Hungary (KTF: 2771-3/ 2015).

A specimen pool containing 6 faecal sample pools (from Mélymocsár: RD2, RD6 and RD8; and Lake Katlan: RD40, RD44 and RD45) from 18 individuals was selected for viral metagenomics analysis. Briefly, 200 μL PBS-diluted specimen was 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 (Turbo DNase, Invitrogen; Baseline Zero DNase, Epicentre Biotechnologies; Benzonase Nuclease, Novagen; RNase A, Fermentas) to digest unprotected nucleic acids at 37 °C for 2 h (Phan et al., 2013). Viral-particle protected nucleic acids were extracted using QIAamp spin-column technique (QIAamp Viral RNA Mini Ki, Qiagen) using RNase inhibitor (RiboLock RNase Inhibitor, Fermentas) at the elution step. Sequence independent random RT-PCR amplification (Victoria et al., 2009) with 20 PCR cycles was used and viral cDNA with 250 bases paired-ends library was constructed using the Nextera XT DNA Library Preparation Kit (Illumina). The library was sequenced on the Hiseq Illumina platform according to the manufacturer's instruction. The acquired metagenomic reads were trimmed; assembled *de-novo* (Deng et al., 2015) and analyzed using an in-house pipeline (Phan et al., 2013). Briefly, the singlets and the assembled contigs > 100-bp were compared to the GenBank (Benson et al., 2013) protein database using BLASTx (version 2.2.7) (Altschul et al., 1997) using E-value cut-off of 0.01. Candidate viral hits were then compared to a non-virus non-redundant protein database to remove false positive viral hits. Virus family-level categorization of all viral metagenomic sequences was based on the best BLASTx-scores (E-value ≤ $10^{-10}$). Metagenomic raw sequence reads data are available upon request.

Sequence specific screening primer pairs (frHEV-F1: 5′-GAAAAAC TCATGATGAGAGAATG-3′ corresponding to nt positions 4144–4166 and frHEV-R1: 5′-CACAATACTGTCGTCACCCTT-3′ corresponding to nt positions 4386–4366 of the study strain) were designed based on the viral 3D polymerase coding (ORF1) contig to identify the viral RNA of

the study strain from the specimen pools. In addition, 29 different sets of specific primers were designed based on the sequences of the metagenomic reads/contigs and the amplified PCR-products for the verification of the metagenomic contigs and to obtain the complete viral genome of the study strain (agile frog/RD6/2015/HUN) using a primer-walking strategy (Sverdlov and Azhikina, 2005), 5′/3′RACE (Boros et al., 2011) and TAIL-PCR methods (Liu and Chen, 2007). PCR-products were sequenced directly and then run on an automated sequencer (ABI Prism 310, Applied Biosystems, Stafford, USA). All faecal specimen pools (n = 39) from agile frog were also tested by RT-PCR using frHEV-F2/frHEV-R2 primer pairs (frHEV-F2: 5′-CTCAATCAACTTGGG ATGAAGA-3′ corresponding to nt positions 6118–6139 and frHEV-R2: 5′-CGGTCTCAGCCATAAATCTGAA-3′ corresponding to nt positions 6514–6493 of the study strain). The full genome sequence of the virus agile frog/RD6/2015/HUN has been deposited in GenBank under accession number MH330682.

Secondary protein structure and disorder prediction was made by the Phyre2 web portal for protein modeling, prediction and analysis (Kelley et al., 2015). The pairwise sequence comparisons and identity calculations were conducted by the BioEdit software (Hall, 1999) using the in-built Clustal W algorithm for pairwise alignments.

Sequence alignment was performed using CLUSTAL X (version 2.0.3) (Saitou and Nei, 1987). Phylogenetic and evolutionary analysis was conducted by MEGA6.06 (Tamura et al., 2013), based on 33 RNA helicase/RdRp protein (encoded by ORF1) sequences. The evolutionary history was inferred using Maximum Likelihood (ML) method based on Poisson model with gamma distribution (+ G5) allowing invariable sites (+ I). 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.

## 3. Results and discussion

A specimen pool (containing six faecal samples from eighteen individuals) from agile frog tadpoles was subjected to viral metagenomics analysis. After de novo assembly of the 35,642,306 total number of sequence reads, 261,781 reads were obtained showing similarity to viruses (BLASTx cut-off E score ≤ $10^{-10}$) from this pool. The sequence matches represent the highest number of viral sequence reads (n = 10) and the maximum length of viral sequence contig (n = 10), respectively, are presented in Fig. S1 (Fig. S1). The detected sequences containing > 50 reads were from viruses of *Dicistroviridae* (n = 161,774, including Himetobi P virus n = 42,730), unclassified viruses (n = 52,445), *Circoviridae* (n = 14,793), *Picornaviridae* (n = 7858, including ampivirus A1 n = 1710), *Nanoviridae* (n = 6403), *Microviridae* (n = 6042), *Secoviridae* (n = 5770, including maize necrotic streak virus n = 5541), *Iflaviridae* (n = 2534), *Parvoviridae* (n = 1255), *Marnaviridae* (n = 780), *Tombusviridae* (n = 701), *Nodaviridae* (n = 330), *Potyviridae* (n = 259), *Phycodnaviridae* (n = 134), *Poxviridae* (n = 129), *Podoviridae* (n = 80), *Geminiviridae* (n = 74), *Partitiviridae* (n = 67) and *Hepeviridae* (n = 58) (Fig. S1). The latter, the sequence reads corresponding to family *Hepeviridae* were selected for further analysis. Five (RD2, RD6, RD8, RD40 and RD44) of the 6 specimen pools were RT-PCR positive for hepevirus-like virus using the screening primer (frHEV-F1/frHEV-R1). To characterize the complete viral genome from randomly selected faecal sample pool (RD6) different sets of specific primers were designed on the basis of the metagenomic sequence reads/contigs and were sequenced directly by Sanger sequencing. The complete genome of viral strain agile frog/RD6/2015/HUN (MH330682) is 7188 nt long including a 48-nt 5′UTR and a 122-nt 3′UTR (Fig. 1). Interestingly, in spite of extensive efforts a poly(A)-tail was not found at the 3′ end of the viral genomes (tested in specimens RD6 and RD13) neither by oligo-d(T)-adapter based 3′ RACE nor poly (G)-tagged terminal deoxynucleotidyl transferase (TdT) 3′ RACE reactions (Fig. 1), respectively. According to the results of ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) three main, non-
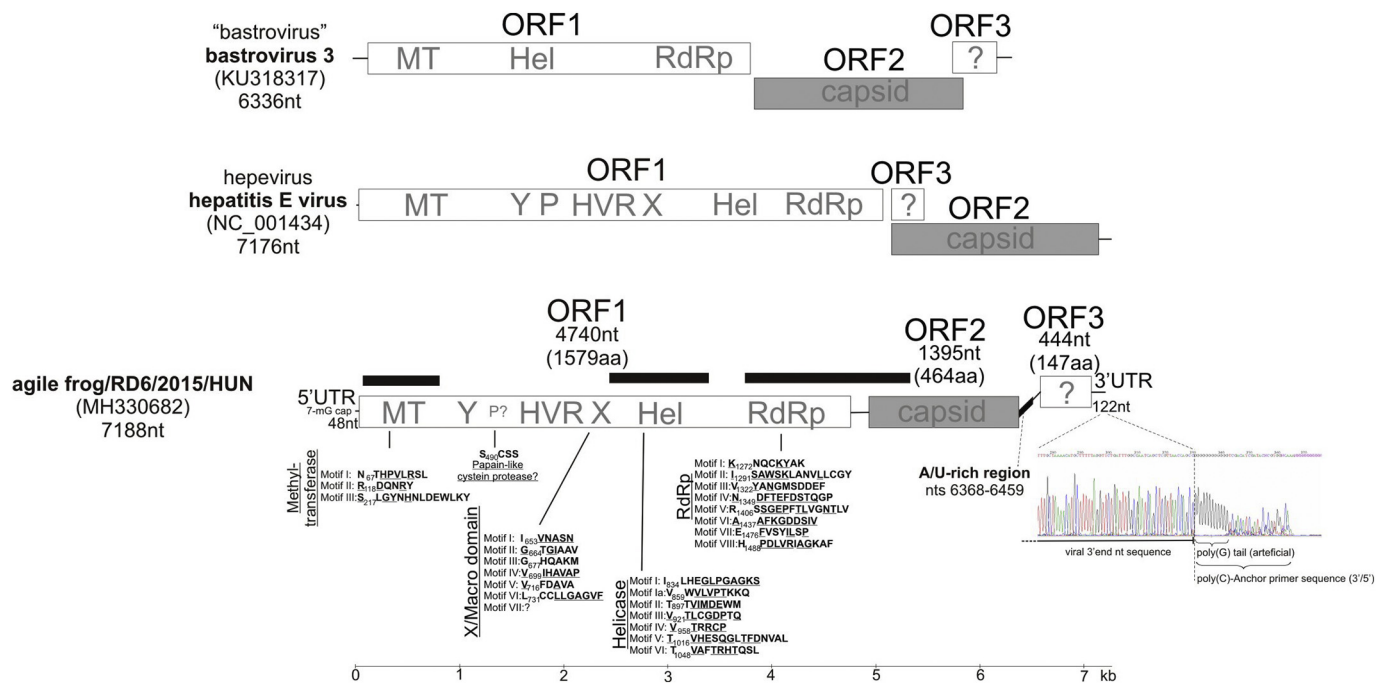
**Fig. 1.** Schematic genome organization and conserved amino acid (aa) motifs in ORF1 of the study strain, agile frog/RD6/2015/HUN (MH330682) from agile frog tadpoles, related to hepeviruses (Koonin et al., 1992; Batts et al., 2011; Lin et al., 2014). Schematic genome organizations of relatives (hepatitis E virus and the unclassified "bastrovirus") of the study strain are shown for comparison. The genome map and each predicted ORFs are drawn to scale. Capsid regions are shown in grey. The positions of the viral metagenomic sequence contigs are marked with black bars. Poly(G)-tagged terminal deoxynucleotidyl transferase (TdT) 3′ RACE reactions (with poly(G) artificial tail) and Sanger sequencing results shows that the potential 3′ end of the viral genome has no poly(A)-tail. A/U-rich region was identified between potential ORF2 and ORF3. The ORF3 protein is translated in a different reading frame to that of ORF1 and ORF2. Abbreviations: MT: methyltransferase; Y: Y domain; P: papain-like cysteine protease; HVR: hypervariable region; X: X/macro domain; Hel: RNA helicase; RdRp: RNA-dependent RNA polymerase; nt: nucleotide; utr: untranslated region.

overlapping ORFs could be predicted. The ORF1 of agile frog/2015/HUN encodes a non-structural polyprotein of 1579 amino acid (aa) (4740 nt) including putative functional domains of hepeviral methyltransferase, domain Y, hypervariable region (HVR), ADP-ribose binding X/Macro domain (A1pp), RNA helicase, RdRp and conserved aa motifs (Meng et al., 2012; Koonin et al., 1992; Neuvonen and Ahola, 2009; Lin et al., 2015) (Fig. 1). The hepevirus aa motifs of, the putative papain-like cysteine protease were not recognizable in agile frog/RD6/2015/HUN. A short serine (S) residues (SSVSS) was found at aa position 558–562 that is present only in the protease region of cutthroat trout HEV (Batts et al., 2011). According to the results of BLASTp analyses the non-structural protein (ORF1) of agile frog/RD6/2015/HUN share the highest identities, 34% (E-value: $2 \times 10^{-95}$, Query cover: 69%), 32% (E-value: $1 \times 10^{-90}$, Query cover: 71%), 32% (E-value: $5 \times 10^{-74}$, Query cover: 63%), 31% (E-value: $2 \times 10^{-65}$, Query cover: 67%) and 29% (E-value: $3 \times 10^{-65}$, Query cover: 54%), to the corresponding non-structural proteins of Dongbei arctic lamprey hepevirus strain QSMS17356 (MG600001), Nanhai ghost shark hepevirus strain NHYJC35031 (MG600008), Bastrovirus/VietNam/Porcine/17489_85 (NC_032423), Bastrovirus/VietNam/Bat/16715_78 (NC_035471) and kestrel hepatitis E virus (KU670940), respectively.

The ORF2 encodes a putative capsid protein of 464 aa; however, no similar protein sequence was found in the GenBank database using different BLAST searches. The presumed capsid of agile frog/RD6/2015/HUN is considerably shorter and shows only low sequence identities to the capsids of hepe-, astro-, bastro and some unassigned HEV-like viruses of fish (Table S1). The highest pairwise sequence identities were measured with the capsids of cutthroat trout (14.4%) and human hepeviruses (14.2%) (Table S1). The 90 aa-long N-terminal region of the presumed capsid is relatively rich in basic amino acids (lysine, arginine and histidine, N = 22) compared to the following 375-aa-long capsid sequence region (basic aa N = 20/375 aa). The basic N

termini of capsid proteins are predominantly present in members of different RNA virus families (Baer et al., 1994; Geigenmüller-Gnirke et al., 1993; Krishna, 2005). Furthermore the secondary structure predictions made by Phyre2 revealed the presence of multiple alpha helices and beta strands in the amino acid sequence of the presumed capsid of the study virus (Fig. S2). The variable number of alpha helices and beta strands are also characteristic features of capsid proteins of other, related RNA viruses (Mori and Matsuura, 2011; Toh et al., 2016).

Between the potential ORF2 and ORF3 coding regions (from nts 6368 to 6459) an A/U-rich (A + U = 80%) non-coding genome region was identified (Fig. 1).

The ORF3, which encodes a potential 147-aa-long protein, is in a different coding frame to ORF1 and ORF2 (Fig. 1). The function of this protein remains unknown, however, interestingly, sequence analyses indicated that this protein of agile frog/RD6/2015/HUN shares 38% aa identity (from aa 71 to aa 128) to the N-terminal part of NS2 (cysteine-protease) protein of equine hepacivirus (KX056117) of the family *Flaviviridae*.

Phylogenetic analysis based on the aa sequences of the RNA helicase/RdRp proteins (encoded by ORF1) showed that agile frog/RD6/2015/HUN is divergent from the known hepatitis E viruses and the unclassified bastroviruses and has an intermediate phylogenetic positions between them (Fig. 2).

Using the screening primer-pairs frHEV-F2/frHEV-R2 18 (46%) of the 39 faecal specimen pools were positive for the study strain by RT-PCR and sequencing: 6 (46%) of 13 in Mélymocsár, 8 (62%) of 13 in Lake Ilona and 4 (31%) of 13 in Lake Katlan.

In recent years, many novel, taxonomically unassigned RNA viruses (hepelivirus, bastrovirus) have been discovered (Ng et al., 2012; Munnink et al., 2016) which are (phylo)genetically partially related to the classical hepeviruses including hepatitis E virus. Recent analyses also suggest that members of the family *Hepeviridae* potentially
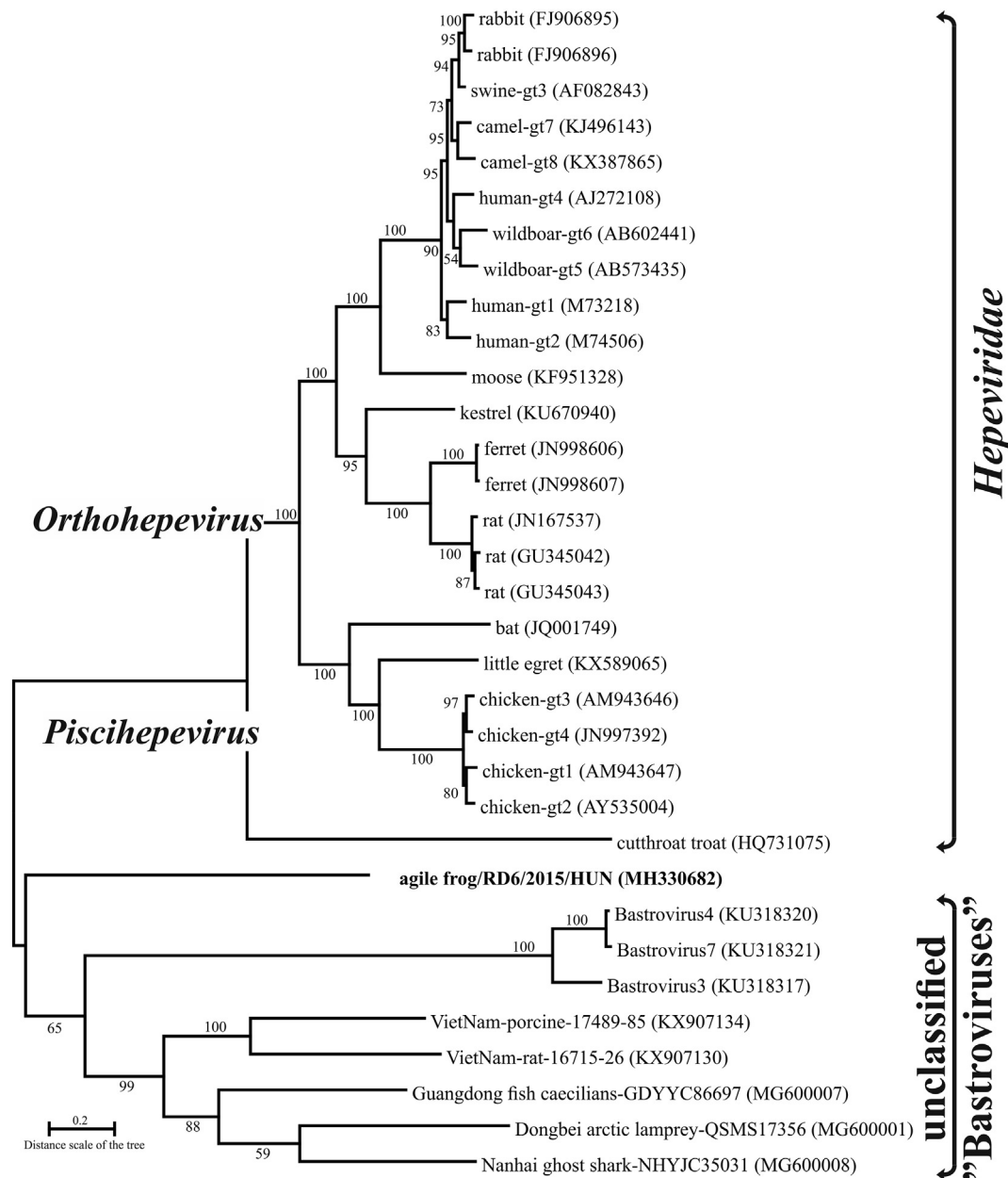
**Fig. 2.** Phylogenetic analysis of corresponding amino acid sequences of representative hepatitis E viruses, unclassified "Bastroviruses" and the study strain agile frog/RD6/2015/HUN (MH330682, shown in bold). The evolutionary analysis was conducted by MEGA6.06 (Tamura et al., 2013), based on 33 RNA helicase/RdRp protein (encoded by ORF1) sequences. The evolutionary history was inferred using the Maximum Likelihood (ML) method based on a Poisson model with gamma distribution (+ G5) allowing invariable sites (+ I). Bootstrap values were determined for 1000 replicates using the whole data subset and the tree is drawn to scale, with branch length measured in the number of substitutions per site. The potential host of the viruses are shown. The two genera, *Orthohepevirus* and *Piscihepevirus*, of the family *Hepeviridae* (Batts et al., 2011; Smith et al., 2014; Purdy et al., 2017) and the unclassified "Bastroviruses" (Munnink et al., 2016) are also indicated. gt: genotype.

represent genetically recombinant viruses from different virus families (Kelly et al., 2016). These data suggest that there are, most likely, some unknown relatives and potential parental virus strains in nature which have played an important role in hepevirus evolution. It also seems that the genetic and host species diversity and the environmental distribution (e.g. aquatic environment) of hepevirus and hepevirus-like virus are wider than previously thought (Batts et al., 2011; Shi et al., 2016; Reuter et al., 2016b; Shi et al., 2018). Until now, hepeviruses have been detected in three classes of vertebrates, in fish, birds, and mammals including humans. Recent research also indicates that hepevirus-like viruses are also present among invertebrates (Shi et al., 2016). Our study reports the identification and complete genome characterization of a novel RNA virus in an amphibian, the agile frog, where the putative

viral ORF1 protein shares limited aa identity with the *Hepeviridae* and bastroviruses, while the putative non-structural ORF2 capsid protein has no significant similarities to these or any other virus families. The aa sequence of the ORF1 non-structural protein of the study strain is predicted to contain functional domains that are found in members of the *Hepeviridae*. However, fundamental differences in sequence (e.g. low aa sequence identity to members of any virus families, potentially missing poly(A)-tail) and genome organization (e.g. positions of ORFs in the genome) are also present which complicates the classification and taxonomy of the study virus.

Using viral metagenomics and next generation sequencing technologies viruses of prokaryotes, algae, plants, insects and invertebrates/vertebrates classified into > 18 known virus families were identified in

wild agile frogs some with high numbers of sequence reads (e.g. dicistrovirus). Some of these virus sequences were identified previously in exotic natural aquatic sampling areas [(e.g. Antarctic picorna-like viruses (López-Bueno et al., 2015))] and wastewater [e.g. Marine RNA viruses (Greninger and DeRisi, 2015)]. Further studies are needed to investigate the biology and host species spectrum of agile frog/RD6/2015/HUN-like viruses.

Discovery and complete genomic characterization of relatives of hepevirus may help to more fully understand the diversity, cross-species genetic jumps and the potential genetic evolution of hepevirus and hepevirus-like viruses.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2018.07.029.

## Acknowledgements

## References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Baer, M.L., Houser, F., Loesch-Fries, L.S., Gehrke, L., 1994. Specific RNA binding by amino-terminal peptides of alfalfa mosaic virus coat protein. EMBO J. 13, 727–735.

Batts, W., Yun, S., Hedrick, R., Winton, J., 2011. A novel member of the family *Hepeviridae* from cutthroat trout (*Oncorhynchus clarkii*). Virus Res. 158, 116–123.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2013. GenBank. Nucleic Acids Res. 41, D36–D42.

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

Deng, X., Naccache, S.N., Ng, T., Federman, S., Li, L., Chiu, C.Y., Delwart, E., 2015. An ensemble strategy that significantly improves de novo assembly of microbial genomes from metagenomic next-generation sequencing data. Nucleic Acids Res. 43, e46.

dos Anjos, K., Nagat, T., de Melo, F.L., 2017. Complete genome sequence of a novel bastrovirus isolated from raw sewage. Genome Announc. 5 (40), e01010–e01017.

Drexler, J.F., Seelen, A., Corman, V.M., Tateno, A.F., Cottontail, V., Zerbinati, R.M., Gloza-Rausch, F., Klose, S.M., Adu-Sarkodie, Y., Oppong, S.K., Kalko, E.K.V., Osterman, A., Rasche, A., Adam, A., Müller, M.A., Ulrich, R.G., Leroy, E.M., Lukashev, A.N., Drosten, C., 2012. Bats worldwide carry hepatitis E virus-related viruses that form a putative novel genus within the family *Hepeviridae*. J. Virol. 86, 9134–9147.

Dryden, K.A., Tihova, M., Nowotny, N., Matsui, S.M., Mendez, E., Yeager, M., 2012. Immature and mature human astrovirus: structure, conformational changes, and similarities to hepatitis E virus. J. Mol. Biol. 422 (5), 650–658.

Geigenmüller-Gnirke, U., Nitschko, H., Schlesinger, S., 1993. Deletion analysis of the capsid protein of Sindbis virus: identification of the RNA binding region. J. Virol. 67, 1620–1626.

Gosner, K.L., 1960. A simplified table for staging anuran embryos and larvae with notes on identification. Herpetologica 16 (3), 183–190.

Greninger, A.L., DeRisi, J.L., 2015. Draft genome sequences of Marine RNA viruses SF-1, SF-2, and SF-3 recovered from San Francisco wastewater. Genome Announc. 3 (3) e00653-15.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

ICTV, 2014. Master Species List 2014 v4. http://talk.ictvonline.org/files/ictv_documents/m/msl/5208.aspx.

Johne, R., Dremsek, P., Reetz, J., Heckel, G., Hess, M., Ulrich, R.G., 2014. *Hepeviridae*: an expanding family of vertebrate viruses. Infect. Genet. Evol. 27, 212–229.

Kaya, U., Kuzmin, S., Sparreboom, M., Ugurtas, I.H., Tarkhnishvili, D., Anderson, S., Andreone, F., Corti, C., Nyström, P., Schmidt, B., Anthony, B., Ogrodowczyk, A., Ogielska, M., Bosch, J., Tejedo, M., 2009. Rana dalmatina. In: The IUCN Red List of Threatened Species 2009:e.T58584A11790570, https://doi.org/10.2305/IUCN.UK.2009.RLTS. T58584A11790570.en. Downloaded on 14 May 2018.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10 (6), 845.

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.

Koonin, E.V., Gorbalenya, A.E., Purdy, M.A., Rozanov, M.N., Reyes, G.R., Bradley, D.W., 1992. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. Proc. Natl. Acad. Sci. U. S. A. 89, 8259–8263.

Krishna, N.K., 2005. Identification of structural domains involved in astrovirus capsid biology. Viral Immunol. 18 (1), 17–26.

Lin, J., Norder, H., Uhlhorn, H., Belák, S., Widén, F., 2014. Novel hepatitis E like virus found in Swedish moose. J. Gen. Virol. 95, 557–570.

Lin, J., Karlsson, M., Olofson, A.S., Belák, S., Malmsten, J., Dalin, A.M., Widén, F., Norder, H., 2015. High prevalence of hepatitis E virus in Swedish moose – a phylogenetic characterization and comparison of the virus from different regions. PLoS One 10, e0122102.

Liu, Y.G., Chen, Y., 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. BioTechniques 43, 649–650.

López-Bueno, A., Rastrojo, A., Peiro-Pastor, R., Arenas, M., Alcami, A., 2015. Ecological connectivity shapes quasispecies structure of RNA viruses in an Antarctic lake. Mol. Ecol. 24 (19), 1825–4812.

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*. In: Virus Taxonomy, Classification and Nomenclature of Viruses. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, pp. 1021–1028.

Mori, Y., Matsuura, Y., 2011. Structure of hepatitis E viral particle. Virus Res. 161 (1), 59–64.

Munnink, B.B.O., Cotton, M., Canuti, M., Deijs, M., Jebbink, M.F., van Hemert, F.J., Phan, M.V.T., Bakker, M., Farsani, S.M.J., Kellam, P., van der Hoek, L., 2016. A novel astrovirus-like RNA virus detected in human stool. Virus Evol. 2 (1), vew005.

Neuvonen, M., Ahola, T., 2009. Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. J. Mol. Biol. 385, 212–225.

Ng, T.F.F., Marine, R., Wang, C., Simmonds, P., Kapusinszky, B., Bodhidatta, L., Oderinde, B.S., Wommack, K.E., Delwart, E., 2012. High variety of known and new RNA and DNA viruses of diverse origins in untreated sewage. J. Virol. 86, 12161–12175.

Nöllert, A., Nöllert, C., 1992. Die Amphibien Europas. Bestimmung-Gefährdung-Schutz, Stuttgart, Franckh-Kosmos Verlag.

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.

Purdy, M.A., Harrison, T.J., Jameel, S., Meng, X.-J., Okamoto, H., Van der Poel, W.H.M., Smith, D.B., ICTV Report Consortium, 2017. ICTV virus taxonomy profile: *Hepeviridae*. J. Gen. Virol. 98, 2645–2646.

Raj, V.S., Smits, S.L., Pas, S.D., Provacia, L.B.V., Moorman-Roest, H., Osterhaus, A.D.M.E., Haagmans, B.L., 2012. Novel hepatitis E virus in ferrets, the Netherlands. Emerg. Infect. Dis. 18 (8), 1369–1370.

Reuter, G., Boros, Á., Mátics, R., Kapusinszky, B., Delwart, E., Pankovics, P., 2016a. Divergent hepatitis E virus in birds of prey, common kestrel (*Falco tinnunculus*) and red-footed falcon (*Falco vespertinus*), Hungary. Infect. Genet. Evol. 43, 343–346.

Reuter, G., Boros, Á., Mátics, R., Kapusinszky, B., Delwart, E., Pankovics, P., 2016b. A novel avian-like hepatitis E virus in wild aquatic bird, little egret (*Egretta garzetta*), in Hungary. Infect. Genet. Evol. 46, 74–77.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., Qin, X.C., Li, J., Cao, J.P., Eden, J.S., Buchmann, J., Wang, W., Xu, J., Holmes, E.C., Zhang, Y.Z., 2016. Redefining the invertebrate RNA virosphere. Nature 540, 539–543.

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.

Sillero, N., Campos, J., Bonardi, A., Corti, C., Creemers, R., Crochet, P.-A., Crnobrnja Isailovic, J., Denoël, M., Ficetola, G.F., Gonçalves, J., Kuzmin, S., Lymberakis, P., de Pous, P., Rodríguez, A., Sindaco, R., Speybroeck, J., Toxopeus, B., Vieites, D.R., Vences, M., 2014. Updated distribution and biogeography of amphibians and reptiles of Europe. Amphibia-Reptilia 35, 1–31.

Smith, D.B., Purdy, M.A., Simmonds, P., 2013. Genetic variability and the classification of hepatitis E virus. J. Virol. 87 (8), 4161–4169.

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.

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.

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.

Toh, Y., Harper, J., Dryden, K.A., Yeager, M., Arias, C.F., Méndez, E., Tao, Y.J., 2016. Crystal structure of the human astrovirus capsid protein. J. Virol. 90 (20), 9008–9017.

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.