Infection, Genetics and Evolution 43 (2016) 343–346

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Short communication

Divergent hepatitis E virus in birds of prey, common kestrel (*Falco tinnunculus*) and red-footed falcon (*F. vespertinus*), Hungary

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

[a] Regional Laboratory of Virology, National Reference Laboratory of Gastroenteric Viruses, ÁNTSZ Regional Institute of State Public Health Service, Pécs, Hungary, University of Pécs Medical Center, Pécs, Hungary
[b] Department of Pathophysiology, University of Pécs Medical Center, Hungary
[c] Hungarian Nature Research Society, (HuNaReS), Ajka, Hungary
[d] Blood Systems Research Institute, San Francisco, CA, USA
[e] University of California, San Francisco, CA, USA

ARTICLE INFO

Article history:
Received 22 March 2016
Received in revised form 25 April 2016
Accepted 4 June 2016
Available online 6 June 2016

Keywords:
Hepatitis E virus
Orthohepevirus
Bird
Kestrel
Falcon
Wildlife

ABSTRACT

Hepatitis E virus (HEV), family *Hepeviridae*, has raised considerable public health concerns because of its zoonotic potential; however, the animal to animal transmissions and the natural chain of hepevirus infections in wildlife are less known. Using random amplification and next generation sequencing technology a novel HEV in birds of prey was serendipitously identified in Hungary. HEV RNA was detected in total of 2 (18%) of the 11 and 1 (14%) of the 7 faecal samples from common kestrels and red-footed falcons, respectively. High faecal viral load (2.03×10^8 genomic copies/ml) measured by qPCR. The complete genome of strain kestrel/MR22/2014/HUN (KU670940) HEV is 7033-nt long including a 35-nt 5'end and a 63-nt 3'end (excluding the poly(A)-tail). Sequence analyses indicated that the ORF1 (4920 nt/639 aa), ORF2 (1989 nt/662 aa) and ORF3 (360 nt/119aa) proteins of kestrel/MR22/2014/HUN shared the highest identity (58.1%, 66.8% and 28.5%) to the corresponding proteins of ferret, rat and human genotype 4 *Orthohepeviruses*, respectively. Interestingly, the ORF3 protein is potentially initiated with leucine (L) using an alternate, non-AUG (UUG) start codon. This study reports the identification and complete genome characterization of a novel *Orthohepevirus* species related to mammalian HEVs in birds of prey. It is important to recognize all potential hosts, reservoirs and spreaders in nature and to reconstruct the phylogenetic history of hepeviruses.

1. Introduction

Hepatitis E virus (HEV) is mainly transmitted by the faecal-oral route in human and causes acute liver inflammation and chronic hepatitis in immunosuppressed patients (Kamar et al., 2014; Smith et al., 2014). HEV has raised considerable public health concerns because of its zoonotic potential. Based upon the recent taxonomic scheme the family *Hepeviridae* is divided into the genera *Orthohepevirus* (mammalian and avian hepatitis E viruses) and *Piscihepevirus* (cutthroat trout virus). Species within the genus *Orthohepevirus* are designated *Orthohepevirus A* (viruses from human, pig, wild boar, deer, mongoose, rabbit and camel), *Orthohepevirus B* (viruses from chicken), *Orthohepevirus C* (viruses from rat, greater bandicoot, shrew, ferret and mink) and *Orthohepevirus D* (viruses from bat) (Raj et al., 2012; Smith et al., 2013; Johne et al., 2014; Smith et al., 2014; ICTV). HEV has a ~6.6–7.3 kb long + ssRNA genome and encodes three open reading frames (ORFs) flanked by a capped 5′ end and a poly(A)-tail at the 3′ end (Meng et al., 2012). Using random amplification and high-throughput sequencing technology a novel HEV was identified in birds of prey, in common kestrel and red-footed falcon (family Falconidae).

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). One kestrel sample (MR22) was selected for viral metagenomics analysis. Briefly, 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-

* Corresponding author at: Regional Laboratory of Virology, National Reference Laboratory of Gastroenteric Viruses, ÁNTSZ Regional Institute of State Public Health Service, Szabadság út 7., H-7623 Pécs, Hungary.
E-mail address: reuter.gabor@gmail.com (G. Reuter).

http://dx.doi.org/10.1016/j.meegid.2016.06.013
1567-1348/

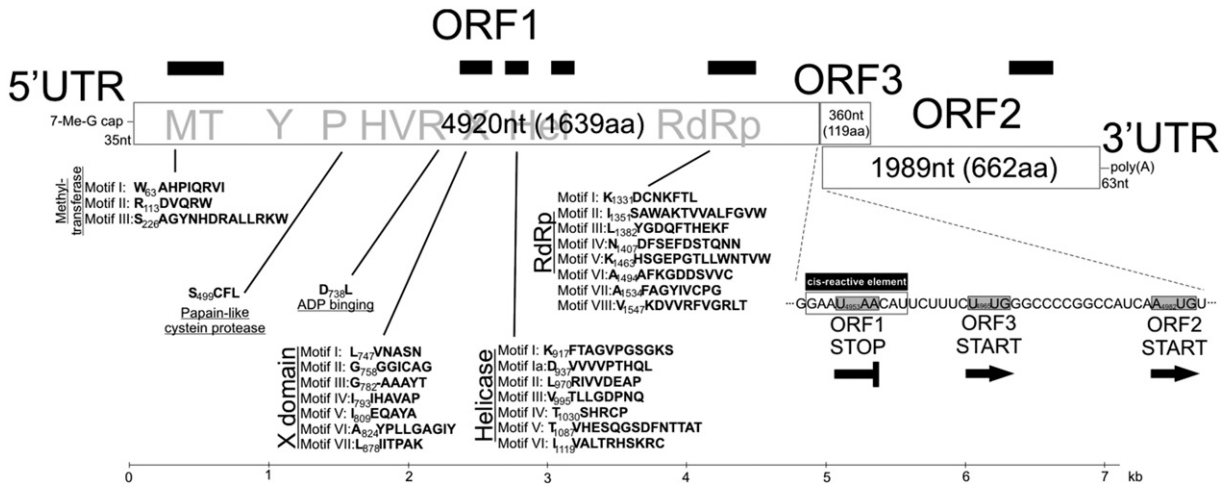

Fig. 1. Schematic genome organization and conserved amino acid (aa) motifs (Koonin et al., 1992; Lin et al., 2014) of hepatitis E virus (kestrel/MR22/2014/HUN, KU670940) from kestrel. The genome map and each ORF were drawn to scale. The positions of the viral metagenomic sequence contigs were marked with black bars. The nucleotide (nt) sequence of the junction region between the stop codon of ORF1 and the start codon of ORF2 is also shown. The ORF3, which encodes a 119-aa-long protein is potentially initiated with leucine (L) using an alternate (Ivanov et al., 2011), non-AUG start codon, $U_{4965}UG$. ORF3 protein is translated in a similar reading frame as ORF1. The highly conserved cis-reactive element (Graff et al., 2005) indicated by an empty box. 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.

particle protected nucleic 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 constructed by Nextera XT DNA Library Preparation Kit (Illumina) and then the library was sequenced by the Miseq Illumina platform according to the manufacturer's instruction, 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 greater than 100-bp were compared to the GenBank protein database (BLASTx). Virus family-level categorization of viral metagenomic reads was based on the best BLASTx-scores (E-value ≤ 10^{-10}). To characterize the complete genome of kestrel HEV (kestrel/MR22/2014/HUN) 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). HEV quantification is based on HEV cDNA transcribed by reverse primer 5′-AAACCGACTCGGCGCCAGAGCTCACAA-3′ (corresponding nt positions 6885-6859 of the study strain) and amplified by primers (R/F: 5′-TGGTAAAGGCGGTGACAGCAT-3′/5′-TGGGTGTCGCACGGGGTTGAT-3′ corresponding nt positions 6833-6813 and 6701-6721 of the study strain kestrel/MR22/2014/HUN, respectively) designed for the ORF2 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 HEV was used. The qPCR analyses contained three technical repeats.

Faecal specimens from kestrels and falcons were tested by RT-PCR using generic screening primer-pairs [kHEVscreen-R: 5′-CTGGCCCCCAATTCTTYTCCG-3′ corresponding nt positions 4708-4728 and kHEVscreen-F: 5′-CATGGCAAAGTYGGGCAGGG-3′ corresponding nt positions 4065-4084 of the study strain kestrel/MR22/2014/HUN] designed for the conserved RNA-dependent RNA polymerase (RdRp) genome region (ORF1).

Table 1
Amino acid (aa) sequence identities of ORF1 (non-structural polyprotein), ORF2 (capsid protein) and ORF3 proteins of the prototype HEV strains compared with HEV (kestrel/MR22/2014/HUN, KU670940) from common kestrel using SIAS web server (http://imed.med.ucm.es/Tools/sias.html). Bold numbers indicate the highest aa identity.

HEV strains (GenBank accession no.)	% aa identity of HEV (KU670940) from kestrel		
	ORF1	ORF2	ORF3
Genotype 1 HEV (M73218)	51.9	62.2	24.8
Genotype 2 HEV (M74506)	51.3	62.1	24.1
Genotype 3 HEV (AF082843)	51.4	61.8	25.5
Genotype 4 HEV (AJ272108)	51.1	59	**28.5**
Rabbit HEV (FJ906895)	50.2	60.4	24.8
Moose HEV (KF951328)	55	59.5	22.6
Ferret HEV (JN998606)	**58.1**	64.6	25.5
Rat HEV (GU345042)	57.2	**66.8**	25.5
Bat HEV (JQ001749)	47.5	50.3	12.4
Avian HEV (AM943646)	47.4	47.3	22.6
Cutthroat trout HEV (HQ731075)	27.9	19.2	13.3

3. Results and discussion

A total of 1310 sequence reads (singletons and contigs) showing similarity to viruses were obtained (BLASTx cut-off E score ≤ 10^{-10}) after *de novo* assembly from 77,038 initial reads from sample MR22. Detected sequences were from viruses of family *Parvoviridae* (N = 674), *Picobirnaviridae* (N = 353), *Microviridae* (N = 97), *Astroviridae* (N = 57), *Picornaviridae* (N = 22), *Hepeviridae* (N = 17), *Adenoviridae* (N = 11), *Dicistroviridae* (N = 9), *Circoviridae* (N = 9), and other (N = 7) or unclassified (N = 54) virus families. The 17 HEV sequence reads that were related to ferret and rat HEVs could be assembled into 6 contigs covering ≈20% of the kestrel HEV genome (Fig. 1). The complete genome of strain kestrel/MR22/2014/HUN (KU670940) is 7033 nt long including a 35-nt 5'end and a 63-nt 3'end (excluding the poly(A)-tail) (Fig. 1). The ORF1 of kestrel/MR22/2014/HUN encodes a non-structural polyprotein of 1639 aa including putative functional domains of methyltransferase, Y-domain, papain-like cysteine protease, hypervariable region (HVR), X/macro-domain, RNA helicase, RdRp and conserved aa motifs (Koonin et al., 1992; Lin et al., 2014) (Fig. 1). The ORF2 encodes the capsid protein of 662 aa. Interestingly, the ORF3, which encodes a 119-aa-long protein is partially overlapping with the ORF2, potentially initiated with leucine (L) using an alternate (Ivanov et al., 2011), non-AUG start codon