Infection, Genetics and Evolution 41 (2016) 227–232

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Short communication

Sequencing and molecular modeling identifies candidate members of Caliciviridae family in bats

Gábor Kemenesi [a,b], Ákos Gellért [c], Bianka Dallos [a,b], Tamás Görföl [d,e], Sándor Boldogh [f], Péter Estók [g], Szilvia Marton [e], Miklós Oldal [a,b], Vito Martella [h], Krisztián Bányai [e], Ferenc Jakab [a,b,*]

[a] Virological Research Group, János Szentágothai Research Center, University of Pécs, Pécs, Hungary
[b] Institute of Biology, Faculty of Sciences, University of Pécs, Pécs, Hungary
[c] Department of Applied Genomics, Agricultural Institute, Centre for Agricultural Research, Hungarian Academy of Sciences, Martonvásár, Hungary
[d] Department of Zoology, Hungarian Natural History Museum, Budapest, Hungary
[e] Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian Academy of Sciences, Budapest, Hungary
[f] Aggtelek National Park Directorate, Jósvafő, Hungary
[g] Department of Zoology, Eszterházy Károly College, Eger, Hungary
[h] Dipartimento di Medicina Veterinaria, Universitá Aldo Moro di Bari, Valenzano, Italy

ARTICLE INFO

Article history:
Received 15 January 2016
Received in revised form 5 April 2016
Accepted 6 April 2016
Available online 13 April 2016

Keywords:
Calicivirus
Bat
Virus
Protein modeling
Hungary
Daubenton's bat
Serotine bat
Europe

ABSTRACT

Emerging viral diseases represent an ongoing challenge for globalized world and bats constitute an immense, partially explored, reservoir of potentially zoonotic viruses. Caliciviruses are important human and animal pathogens and, as observed for human noroviruses, they may impact on human health on a global scale. By screening fecal samples of bats in Hungary, calicivirus RNA was identified in the samples of *Myotis daubentonii* and *Eptesicus serotinus* bats. In order to characterize more in detail the bat caliciviruses, large portions of the genome sequence of the viruses were determined. Phylogenetic analyses and molecular modeling identified firmly the two viruses as candidate members within the Caliciviridae family, with one calicivirus strain resembling members of the *Sapovirus* genus and the other bat calicivirus being more related to porcine caliciviruses of the proposed genus *Valovirus*. This data serves the effort for detecting reservoir hosts for potential emerging viruses and recognize important evolutionary relationships.

Despite ongoing advances in biomedicine and molecular biology, infectious diseases remain a major threat to human health, economic sustainability and wildlife conservation. In the last decades bats became a possible source for several human infecting agents such as Ebola, SARS and MERS coronaviruses, Nipah virus, and Hendra virus (Wynne and Wang, 2013). With over 1250 species distributed globally, bats represent the most diverse and species rich taxa among mammals (Teeling et al., 2005). The zoonotic potential of numerous bat-harbored viruses has been described also in Europe (Kohl and Kurth, 2014). Additionally, several novel viruses have been detected recently; however, there is no data on their role in association with bat or human diseases (Kemenesi et al., 2014, 2015a,b).

The family Caliciviridae comprises small non-enveloped RNA viruses, which are classified in five genera: *Vesivirus*, *Lagovirus*, *Norovirus*, *Sapovirus* and *Nebovirus*. In addition, unclassified caliciviruses have been identified in monkeys, pigs, avian species and fishes (Farkas et al., 2008; L'Homme et al., 2009; Wolf et al., 2011; Liao et al., 2014; Mikalsen et al., 2014). Calicivirus related sequences have been also reported from different bat species in Hong Kong, Hungary and New Zealand (Tse et al., 2012; Kemenesi et al., 2014; Wang et al., 2015). The common characteristic of these bat-derived calicivirus sequences is the phylogenetic divergence from other known members within the family Caliciviridae. The generation of partial or full-length genome sequence data on these recently discovered viruses is crucial to improve viral taxonomy and classification, to develop reliable diagnostic assays and to better understand the molecular mechanisms driving the evolution and host species adaptation of caliciviruses. Human and animal caliciviruses may evolve by accumulation of punctate mutations and by exchange of genetic material via recombination (Giammanco et al., 2013). Also, interspecies transmission has been documented repeatedly and the origin of some human calicivirus strains may have been triggered by a combination of various evolutionary mechanisms (Smith et al., 1998; Scheuer et al., 2013). In this regard the possible reservoir role of bats has been investigated minimally. Here we report the

* Corresponding author at: Virological Research Group, Szentágothai Research Centre, University of Pécs, Ifjúság út 20, H-7624 Pécs, Hungary.
E-mail address: jakabf@gamma.ttk.pte.hu (F. Jakab).

http://dx.doi.org/10.1016/j.meegid.2016.04.004
1567-1348/

Table 1
Amino acid and nucleic acid sequence homology patterns between the examined sequences. Greatest homologies are emphasized in bold.

		M63 (2493 nt)				BS58 (2238 nt)			
		RdRp (951 nt)		Capsid (1542 nt)		RdRp (804 nt)		Capsid (1434 nt)	
		nt (%)	aa (%)	nt (%)	aa (%)	nt (%)	aa (%)	nt (%)	aa (%)
Bat calicivirus	KJ641703	**67**	**67**	**66**	**61**	34	25	28	25
	KJ641700	33	23	29	20	**63**	**60**	**63**	**56**
Bat *Sapovirus*	JN899074	52	43	50	40	35	23	28	23
Bat calicivirus	KJ641701	50	45	54	47	33	25	28	23
Swine *Sapovirus*	KC309421	42	35	44	37	31	24	28	24
Chicken calicivirus	JQ347523	40	31	35	27	33	27	28	24
Norovirus Norwalk	M87661	33	24	30	21	42	32	37	29
Recovirus	EU391643	31	23	31	24	44	39	38	29
Valovirus	FJ355929	31	25	27	19	49	37	47	38
Vesivirus	U13992	33	29	25	16	33	28	22	27
Lagovirus	M67473	37	32	31	22	35	26	28	23
Nebovirus	AY82891	36	31	31	22	32	24	29	23
Human *Sapovirus*	AY646856	49	42	43	40	37	27	35	25

genomic characterization and molecular modeling based analysis for two novel calicivirus strains (BtCalV/BS58/HUN/2013 and BtCalV/M63/HUN/2013) identified from Serotine bat (*Eptesicus serotinus*) and Daubenton's bat (*Myotis daubentonii*).

Bat guano samples were collected during the swarming period in 2013 as part of a countrywide survey program as described previously (Kemenesi et al., 2014). Nucleic acid extractions were performed with DiaExtract Total RNA Isolation Kit (Diagon Ltd., Hungary). 3′ RACE PCRs were fulfilled on both samples with SuperScript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Invitrogen), according to the manufacturer's instructions with using the previously described forward primer by Jiang et al. (1999). 5′ RACE PCR was also performed with 5′/3′ RACE Kit, 2nd Generation (Roche). In order to obtain sequences upstream the RNA dependent RNA polymerase (RdRp) region we performed a two-step RT-PCR protocol. cDNA synthesis was carried out with the Transcriptor First Strand cDNA Synthesis Kit (Roche) using random hexamer oligonucleotide mixture. PCR assays were implemented with Thermo Scientific Phusion High-Fidelity PCR Master Mix following the protocol provided by the manufacturer, using previously published generic calicivirus forward primer (L'Homme et al., 2009) with RdRp specific reverse primer (Jiang et al., 1999).

High throughput sequencing was performed on PCR amplicons to obtain primary sequence data (Mihalov-Kovács et al., 2015). Briefly, an Ion Torrent compatible library was prepared applying the NEBNext® Fast DNA Fragmentation & Library Prep Set for Ion Torrent (New England Biolabs) with the Ion Torrent Xpress barcode adapters (Life Technologies). The emulsion PCR to obtain clonally amplified fragments was carried out using the Ion OneTouch™ 200 Template Kit (Life Technologies) on a OneTouch version 2 equipment (Life Technologies) as recommended by the manufacturer. Templated beads were enriched using an Ion OneTouch™ ES pipetting robot (Life Technologies). The 200 bp sequencing protocol was performed on a 316 chip (Life Technologies) using the Ion Torrent PGM (Life Technologies) semiconductor sequencing equipment. Trimmed sequence reads were used for de novo assembly utilizing the MIRA (version 3.9.17) (Chevreux et al., 1999). Additional bioinformatic analyses, validation with remapping were performed using the CLC Genomics Workbench (version 6.5.1; http://www.clcbio.com) and the DNAStar (version 12; http://www.dnastar.com). GenBank accession numbers for BtCalV/BS58/HUN/2013 and BtCalV/M63/HUN/2013 are KJ652318 and KJ652319, respectively.

Basic sequence alignments and sequence manipulations were fulfilled with ClustalX v2.1 and GeneDoc v2.7 softwares. Multiple evolutionary models were tested on our sequence alignments with PhyML v3.0 software in order to designate the best phylogenetic model for our data set. The phylogenetic trees were constructed with MEGA v5.0 software using the Maximum-Likelihood method, based on the General Time Reversible model with Gamma Distribution (GTR + G). Number of bootstraps for simulations was 1000.

The 3D structure models of mature BtCalV/M63/HUN/2013 (aa. 317 to 831) and BtCalV/BS58/HUN/2013 (aa. 268 to 746) bat calicivirus VP1 capsid protein (CP) structures were generated with I-TASSER (Zhang, 2008; Roy et al., 2010). The models were built using the recently identified bat VP1 CP amino acid sequences. The following templates were used to thread the VP1 CP structure of the bat calicivirus BtCalV/M63/HUN/2013: PDB ID codes 3M8L (Feline calicivirus (FCV) capsid protein) and 2GH8 (X-ray structure of a native calicivirus). The structure of Norwalk virus capsid (PDB ID code 1IHM) was used for modeling of the VP1 capsid protein of the bat calicivirus BtCalV/BS58/HUN/2013. The raw protein model structures were refined with the MacroModel energy minimization module of the Schrödinger Suite (Schrödinger, 2015) to eliminate the steric conflicts between the side-chain atoms. Pairwise protein sequence alignments were calculated with the NeedleP tool of the SRS bioinformatics software package. The T = 3 virion models were created with the Oligomer Generator application of VIPERdb (available at http://viperdb.scripps.edu/oligomer_multi.php). Prior to virion model generations the asymmetric units were constructed with Schrödinger Suite using the coordinates of subunit A, B and C of FCV and Norwalk virus in vdb convention format. Molecular graphics and sequence alignment visualization were prepared using VMD v1.9.1 (Humphrey et al., 1996) and the Multiple Sequence Viewer of the Schrödinger Suite, respectively.

Since 5′ RACE PCR protocols were not successful for the two samples, we obtained a 3278 and 3130 nt long partial genome sequences, respectively, for the strains BtCalV/M63/HUN/2013 and BtCalV/BS58/HUN/2013. The sequenced genomic fragment spanned the partial RNA-dependent RNA polymerase (RdRp) gene (M63, 879 nt; BS58, 841 nt) and the whole coat protein coding region (1545 nt and 1436 nt) along with the complete VP2 protein coding region (656 nt and 668 nt). The calicivirus conserved aa motifs GLPSG and YGDD (Smiley et al., 2002) were identified in the RdRp region. A third calicivirus RdRp motif

Fig. 1. Structure comparison of mature calicivirus (ORF2, VP1 segment) CP proteins and virions. Structure based amino acid sequence alignments of the newly reported bat calicivirus capsid proteins and the template FCV and Norwalk calicivirus VP1 proteins (a) and (b). The background of the sequence alignments reflects the homology levels of the two–two related capsid protein sequences: identical amino acids are red, similar aas. are light orange while different aas. are light pink. The main structural differences are indicated by shades of magenta and green color codes on the sequence alignment and on the superimposed CP structures. The template calicivirus VP1 protein structures are illustrated in cyan cartoon representation, while the new bat calicivirus VP1 model structure are pink (c) and (d). Molecular surface representation of the superimposed template and the newly described bat calicivirus virions (e) and (f). The molecular surface is colored by radial extension of the amino acids from the virion center. Dark blue represents the most protruding CP parts. The structural differences were colored in the same way as for the capsid monomers.