

Genomic analysis of a novel picornavirus from a migratory waterfowl, greater white-fronted goose (*Anser albifrons*)

Ákos Boros^{1,2}, Péter Pankovics^{1,2}, Peter Simmonds³, Tamás Kiss⁴, Tung Gia Phan^{5,6}, Eric Delwart^{5,6}, Gábor Reuter^{1,2}

- 1 Regional Laboratory of Virology, National Reference Laboratory of Gastroenteric Viruses, ÁNTSZ Regional Institute of State Public Health Service, Pécs, Hungary
- 2 Department of Medical Microbiology and Immunology, Medical School, University of Pécs, Szigeti út 12., H-7624 Pécs, Hungary
- 3 Nuffield Department of Medicine, Peter Medawar Building for Pathogen Research, University of Oxford, South Parks Road, Oxford, UK
- 4 Hungarian Ornithological and Nature Conservation Society, Budapest, Hungary
- 5 Blood Systems Research Institute, San Francisco, CA, USA
- 6 University of California, San Francisco, CA, USA

□ Gábor Reuter reuter.gabor@gmail.com

The GenBank[EMBL/DDBJ] accession number for the study sequence: MF358731.

Abstract

The complete genome of goose picornavirus 1 (GPV-1) strain goose/NLSZK2/HUN/2013 (MF358731) was determined by RT-PCR and next-generation sequencing from a cloacal sample of a migratory waterfowl, greater white-fronted goose (*Anser albifrons*) in Hungary. The genome of GPV-1 shows an L-3-3-4 organization pattern with a 5'-terminal origin of replication (ORI) region, a type-IV IRES, and an Hbox/NC-type 2A protein. This virus showed the highest overall sequence identity to the members of the genus *Kobuvirus*, although the phylogenetic position of GPV-1 is different in the analyzed P1, 2C and 3CD phylogenetic trees, which further increases the diversity of known avian picornaviruses.

Long-distance migratory bird species such as greater white-fronted goose (*Anser albifrons*) of the family Anatidae live in large flocks that include birds of different species [1]. These mixed flocks could carry the risk of spreading mostly unknown zoonotic pathogens as well as pathogenic avian viruses [2, 3]. The family *Picornaviridae* consists of genetically diverse viruses of vertebrates, including birds, with a small single-stranded RNA genome of positive polarity [4, 5]. One of the most diverse avian picornavirus phylogenetic clusters is the passerivirus cluster, which includes members of the genera *Passerivirus*, *Gallivirus*, *Oscivirus* and *Sicinivirus* [3]. Picornaviruses of the passerivirus cluster share a same L(leader)-3-3-4 genome organization pattern with a type-II or type-V internal ribosomal entry site (IRES) and predominantly a single, Hbox/NC-type 2A [3]. Here, we report the complete genome sequence of the novel goose picornavirus 1 (GPV-1, MF358731) from a greater white-fronted goose, determined by viral metagenomics and conventional and 5'/3' RACE RT-PCR techniques as described previously [6, 7].

The analyzed cloacal sample was collected from an apparently healthy, one-year-old greater white-fronted goose (*Anser albifrons*) at Lake Zabszék, Szabadszállás, Hungary, in November 2013. The 8051-nt-long complete genome of GPV-1 strain goose/NLSZK2/HUN/2013 was predicted to have a L-3-3-4 genome layout: 5'UTR-L-P1(VP0-VP3-VP1)-P2(2A-C)-P3(3A-D)-3'UTR (Fig. 1a). At the 5'-terminal end, the presence of three hairpins and a possible pseudoknot were predicted using Mfold (Fig. 1b). This structure strongly resembles the origin of replication (ORI) structures of kobu-, sali-, and osciviruses [8]. Based on the results of Mfold analysis and 5'UTR sequence alignments of different IRES elements, the 5'UTR-IRES of GPV-1 is most likely to belong to the hepacivirus/pestivirus-like type-IV (Fig. 1c) [8]. The 227-aa-long leader (L) contains the conserved motifs P₆₅VTLPRNV and L₈₆RMPG. Similar motifs are encoded in the genomes of other phylogenetically related picornaviruses as well (Fig. 1d). The P1 region was predicted to encode only three capsid proteins: a myristoylated VP0, VP3 and VP1 (Fig. 1a). The P2 is predicted to contain a 105-aa long H-box/NC-type 2A peptide (Fig. 1a), and the 2C helicase, the 3C protease and the 3D RNA polymerase are predicted to contain all known functional motifs (Fig. 1a).

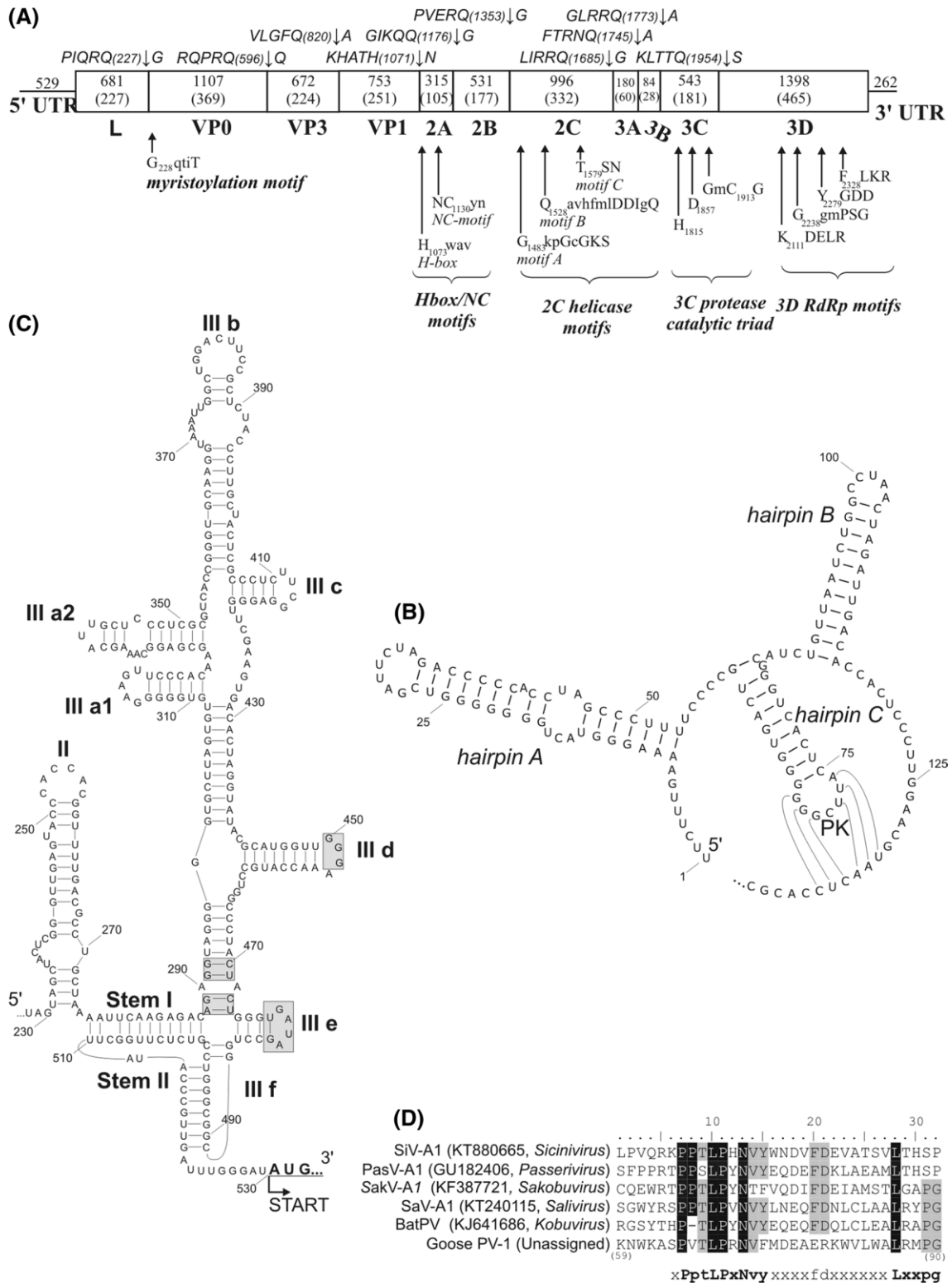


Fig. 1 (A) The main genomic features, the conserved picornaviral motifs, and the predicted cleavage sites of GPV-1 (MF358731). (B and C) The predicted secondary RNA structures of 5'-terminal origin of replication (ORI) region (B) and the 5'UTR-IRES (C) of the 5'UTR. "PK" indicates a pseudoknot. Grey boxes indicate conserved motifs of type-IV IRES elements [8]. (D) Sequence alignment of the leader proteins of GPV-1 and related viruses. Conserved amino acids are indicated with black and grey backgrounds

The P1 of GPV-1 shows highest pairwise aa sequence identity to aichivirus A1 of the genus *Kobuvirus* (43.4%) (Table S1). In the case of the 2C and 3CD peptides, the highest identities are found with sicinivirus A1 (2C: 46.9%) and passerivirus A1 (3CD: 47.3%), although only slightly lower values were found with bat picornavirus 1 of the genus *Kobuvirus* (2C, 46.7%; 3CD, 46.2%; Table S1). GPV-1 shows a distant phylogenetic relationship to kobuviruses in the P1 phylogenetic trees and to siciniviruses in the 2C phylogenetic trees (Fig. 2a, and b), while in the 3CD region, GPV-1 is clustered together with the osciviruses (Fig. 2c).

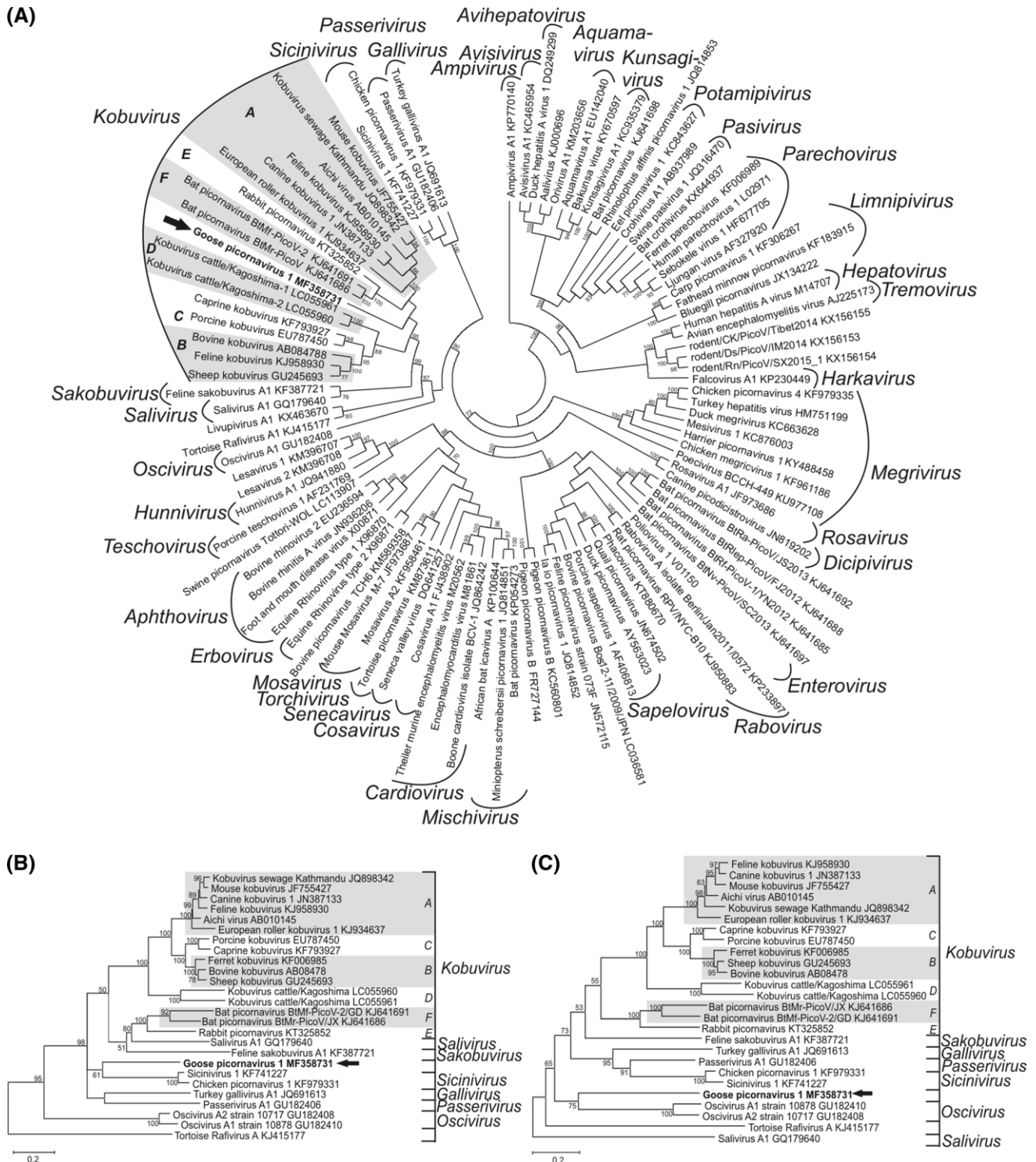


Fig. 2 Phylogenetic analysis of GPV-1 (indicated in bold and with an arrow), representative members of the family *Picornaviridae* (P1), and the closest relatives of the strain from this study (2C and 3CD trees) based on amino acid sequences of the picornavirus proteins P1 (A), 2C (B) and 3CD (C). The trees were generated using the maximum-likelihood method based on the Le_Gascuel_2008 model of MEGA Ver.7.0 [9]. Bars indicate amino acid substitutions per site. A-F designations for the genus *Kobuvirus* represent the members of the species *Aichivirus A-F*

In this study, the complete genome sequence of a novel goose picornavirus (MF358731) detected from a cloacal sample of a greater white-fronted goose (*Anser albifrons*) was determined and characterized. GPV-1 shows an L-3-3-4 genome layout with a characteristic 5'-terminal ORI structure, a type-IV IRES and an Hbox/NC-type 2A protein. The conserved aa motifs found in the L proteins of GPV-1 and various mammalian and avian picornaviruses suggest a similar but currently unknown function and a common origin of these proteins. Based on the results of sequence comparisons, GPV-1 could be assigned to the genus *Kobuvirus*, although the phylogenetic position of the virus in the 2C and 3CD phylogenetic trees do not support this classification.

Acknowledgements This work was supported by a grant from the Hungarian Scientific Research Fund (NKFIH/OTKA K111615). Á.B. and P.P. are supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Compliance with ethical standards

Funding This study was funded by Hungarian Scientific Research Fund (OTKA/NKFIH K111615).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

References

1. Carboneras C (1992) Family Anatidae. In: Hoyo J, Elliott A, Sargatal J (eds) Handbook of birds of the world. volume 1: ostrich to ducks. Lynx Editions, Barcelona, pp 536–629
2. Woo PC, Lau SK, Yuen KY (2006) Infectious diseases emerging from Chinese wet-markets:

zoonotic origins of severe respiratory viral infections. *Curr Opin Infect Dis* 19:401–407

3. Boros Á, Pankovics P, Reuter G (2014) Avian picornaviruses: molecular evolution, genome diversity and unusual genome features of a rapidly expanding group of viruses in birds. *Infect Genet Evol* 28:151–166
4. Zell R (2017) *Picornaviridae* *Arch Virol*, pp 1–19. <https://doi.org/10.1007/s00705-017-3614-8>
5. Palmenberg A, Neubauer D, Skern T (2010) Chapter 1: Genome organization and encoded proteins. In: Ehrenfeld E, Domingo E, Roos RP (eds) *The picornaviruses*. ASM Press, Washington, DC, pp 3–17
6. Phan TG, Vo NP, Boros A, Pankovics P, Reuter G, Li OTW, Wang C, Deng X, Poon LLM, Delwart E (2013) The viruses of wild pigeon droppings. *PLoS One* 8(9):e72787
7. Boros Á, Pankovics P, Knowles NJ, Reuter G (2012) Natural inter-species recombinant bovine/porcine enterovirus in sheep. *J Gen Virol* 93:1941–1951
8. Asnani M, Kumar P, Hellen CUT (2015) Widespread distribution and structural diversity of Type IV IRESs in members of *Picornaviridae*. *Virology* 478:61–74
9. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874

Table S1: Pairwise amino acid sequence identities (%) between the P1, 2C and 3CD proteins of GPV-1 strain goose/NLSZK2/HUN/2013 (MF358731) compared to the representative members of the 35 officially recognized and 7 candidate picornavirus genera. Boldface numbers indicate the highest levels of amino acid identities.

Genus	Type/virus name [<i>Species</i>]	Accession Number	P1	2C	3CD
"Aalivirus"	Duck picornavirus GL12	KJ000696	13.2	23.7	22.0
<i>Ampivirus</i>	Ampivirus A1	KP770140	14.4	17.8	20.7
<i>Aphthovirus</i>	Foot-and-mouth disease virus 1	AF308157	20.4	26.5	26.6
<i>Aquamavirus</i>	Aquamavirus A1	EU142040	13.6	22.2	21.2
<i>Avihepatovirus</i>	Duck hepatitis A virus 1	DQ226541	15.0	18.4	21.2
<i>Avisivirus</i>	Avisivirus A1	KC465954	15.4	22.5	25.9
<i>Cardiovirus</i>	Encephalomyocarditis virus 1	M81861	20.1	23.7	27.7
<i>Cosavirus</i>	Cosavirus A1	FJ438902	18.2	27.7	28.3
"Crohivirus"	Crohivirus 1	AB937989	13.6	19.5	21.0
<i>Dicipivirus</i>	Cadicivirus A1	JN819202	19.0	27.4	32.2
<i>Enterovirus</i>	Poliovirus 1	V01149	19.6	21.5	28.2
<i>Erbovirus</i>	Equine rhinitis B virus 1	X96871	21.6	25.7	29.0
<i>Gallivirus</i>	Gallivirus A1	JQ691613	21.9	41.0	44.3
<i>Harkavirus</i>	Falcovirus A1	KP230449	14.5	20.8	19.4
<i>Hepatovirus</i>	Hepatitis A virus 1	M14707	13.5	25.2	22.6
<i>Hunnivirus</i>	Hunnivirus A1	JQ941880	21.5	26.2	27.9
<i>Kobuvirus</i>	Aichivirus A1 [<i>Aichivirus A</i>]	AB010145	43.4	44.0	44.9
<i>Kobuvirus</i>	Bovine kobuvirus 1 [<i>Aichivirus B</i>]	AB084788	40.1	44.1	45.5
<i>Kobuvirus</i>	Porcine kobuvirus 1 [<i>Aichivirus C</i>]	EU787450	40.3	44.4	43.3
<i>Kobuvirus</i>	Kagovirus 1 [<i>Aichivirus D</i>]	LC055961	39.9	42.5	42.8
<i>Kobuvirus</i>	Rabbit picornavirus 1 [<i>Aichivirus E</i>]	KT325852	39.6	46.2	45.3
<i>Kobuvirus</i>	Bat picornavirus BtMr-PicoV [<i>Aichivirus F</i>]	KJ641686	41.4	46.7	46.2
<i>Kunsagivirus</i>	Kunsagivirus A1	KC935379	16.4	23.1	21.6
"Lesavirus"	Lesavirus 1	KM396707	22.1	25.0	26.9
<i>Limnipivirus</i>	Limnipivirus B1	KF306267	12.4	18.4	20.3
"Livupivirus"	Livupivirus 1	KX463670	34.5	37.4	41.3
<i>Megrivirus</i>	Melegrivirus A1 (THV-1)	HM751199	16.7	30.4	35.0
<i>Mischivirus</i>	Mischivirus A1	JQ814851	21.3	29.2	27.2
<i>Mosavirus</i>	Mosavirus A1	JF973687	22.3	25.1	29.3
"Orivirus"	Orivirus A1	KM203656	14.6	21.3	18.8
<i>Oscivirus</i>	Oscivirus A1	GU182408	21.7	41.2	45.2
<i>Parechovirus</i>	Human parechovirus 1	AJ005695	14.0	23.4	21.2
<i>Pasivirus</i>	Pasivirus A1	JQ316470	13.5	24.2	20.0
<i>Passerivirus</i>	Passerivirus A1	GU182406	25.8	43.0	47.3

"Poecivirus"	Poecivirus 1	KU977108	15.5	30.5	31.1
<i>Potamipivirus</i>	Eel picornavirus 1	KC843627	11.2	20.9	22.5
<i>Rabovirus</i>	Rabovirus A1	KP233897	18.3	21.1	28.3
"Rafivirus"	Tortoise rafivirus A1	KJ415177	26.9	32.6	37.5
<i>Rosavirus</i>	Rosavirus A1	JF973686	17.2	31.1	31.4
<i>Sakobuvirus</i>	Sakobuvirus A1	KF387721	37.4	42.6	45.3
<i>Salivirus</i>	Salivirus A1	GQ179640	36.8	43.8	37.1
<i>Sapelovirus</i>	Porcine sapelovirus 1	AF406813	20.0	19.9	28.1
<i>Senecavirus</i>	Seneca Valley virus 1	DQ641257	19.7	25.4	26.9
<i>Sicinivirus</i>	Sicinivirus A1	KF741227	23.9	46.9	43.7
<i>Teschovirus</i>	Porcine teschovirus 1	AJ011380	19.9	24.8	27.5
<i>Torchivirus</i>	Tortoise picornavirus 1	KM873611	21.4	27.6	29.9
<i>Tremovirus</i>	Avian encephalomyelitis virus 1	AJ225173	14,6	22,9	23.3