Virologica Sinica (2020) 35:528–537
https://doi.org/10.1007/s12250-020-00210-9

www.virosin.org
www.springer.com/12250

RESEARCH ARTICLE

Characterization of the First Genome of *Porcine mastadenovirus B* (HNU1 Strain) and Implications on Its Lymphoid and Special Origin

Shu-Jing Liu[1] · Qiong Wang[1] · Ting-Ting Li[2] · Si-Hua Zhang[3] · Jin-Yan Li[1] · Li-Jun Wu[4] · Ye Qiu[1] · Xing-Yi Ge[1]

Received: 5 December 2019 / Accepted: 4 February 2020 / Published online: 31 March 2020

Abstract

Porcine adenoviruses (PAdVs) are classified into three species, PAdV-A, PAdV-B, and PAdV-C. The genomes of PAdV-A and PAdV-C have been well characterized. However, the genome of PAdV-B has never been completely sequenced, and the epidemiology of PAdV-B remains unclear. In our study, we have identified a novel strain of PAdV-B, named PAdV-B-HNU1, in porcine samples collected in China by viral metagenomic assay and general PCR. The genome of PAdV-B-HNU1 is 31,743 bp in length and highly similar to that of California sea lion adenovirus 1 (C. sea lion AdV-1), which contains typical mastadenoviral structures and some unique regions at the carboxy-terminal end. Especially, PAdV-B-HNU1 harbors a dUTPase coding region not clustering with other mastadenoviruses except for C. sea lion AdV-1 and a fiber coding region homologous with galectin 4 and 9 of animals. However, the variance of GC contents between PAdV-B-HNU1 (55%) and C. sea lion AdV-1 (36%) indicates their differential evolutionary paths. Further epidemiologic study revealed a high positive rate (51.7%) of PAdV-B-HNU1 in porcine lymph samples, but low positive rates of 10.2% and 16.1% in oral swabs and rectal swabs, respectively. In conclusion, this study characterized a novel representative genome of a lymphotropic PAdV-B with unique evolutionary origin, which contributes to the taxonomical and pathogenic studies of PAdVs.

Keywords Porcine adenovirus (PAdV) · PAdV-B · PAdV-B-HNU1 · Lymphoid

Introduction

Adenoviruses (AdVs), belonging to the family *Adenoviridae*, are non-enveloped, double-stranded DNA viruses with a genome of 26–45 kb in size. Their genomes are flanked by an inverted terminal repeat (ITR) on each end. At present, five genera in the family *Adenoviridae* have been identified, including *Mastadenovirus, Aviadenovirus, Atadenovirus, Ichtadenovirus*, and *Siadenovirus* (Harrach *et al*. 2011). Mastadenoviruses infect mammals and aviadenoviruses infect birds (Benkö and Harrach 2003). Atadenoviruses were named because of their unusually high AT content, which have been found in mammals, birds and reptiles (Hess *et al*. 1997). Siadenoviruses infect amphibians and birds (Davison *et al*. 2000; Kovács *et al*. 2010). The *Ichtadenovirus* genus only contains one species which was isolated from the white sturegeon (Kovács *et al*. 2003).

AdVs are important, widespread, and occasionally fatal pathogens in humans, as well as wild and domestic

Shu-Jing Liu and Qiong Wang contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12250-020-00210-9) contains supplementary material, which is available to authorized users.

✉ Ye Qiu
qiuye@hnu.edu.cn

✉ Xing-Yi Ge
xyge@hnu.edu.cn

[1] Hunan Provincial Key Laboratory of Medical Virology, Institute of Pathogen Biology and Immunology, College of Biology, Hunan University, Changsha 410082, Hunan, China

[2] Hubei Animal Disease Prevention and Control Center, Wuhan 430070, China

[3] Wuhan Animal Disease Prevention and Control Center, Wuhan 430016, China

[4] Wuhan Academy of Agricultural Sciences, Wuhan 430065, China

animals. So far 103 types of human adenoviruses (HAdVs) have been recognized by The HAdV Working Group (The HAdV Working Group 2019). HAdVs are known to cause cryptic enteric infection, conjunctivitis, acute respiratory disease, hepatitis, and pneumonia (Hiroaki *et al*. 2008; Matsushima *et al*. 2013; Ghebremedhin 2014). AdVs are involved both in asymptomatic infections and clinical diseases, whereby virulence, tropism and pathogenesis are typically associated with specific types, and that disease manifestation is additionally host immune status- and age-related. AdVs have been described in every vertebrate class and in general exhibit a strong species specificity. Interestingly, several reports indicated cross-species transmission of viruses in *Mastadenovirus*. For example, canine adenoviruses may have originated by interspecies transfer of a bat adenovirus (Kohl *et al*. 2012). Moreover, it has even been reported that a titi monkey adenovirus (TMAdV) transmitted to human and caused respiratory illness (Chen *et al*. 2011). Genetic information about adenoviruses is critical to clarify their evolution, transmission and pathogenesis.

Porcine adenoviruses (PAdVs) were first isolated from a rectal swab of a piglet with obvious diarrhea in 1964 (Haig *et al*. 1964). PAdVs belong to the genus *Mastadenovirus* and contain 3 species including *Porcine mastadenovirus A* (PAdV-A), *Porcine mastadenovirus B* (PAdV-B), and *Porcine mastadenovirus C* (PAdV-C) (Harrach B 2011). PAdVs were classified into 5 serotypes (PAdV-1 to 5) based on the cross-neutralization test, among which PAdV types 1–3 are closely related, whereas types 4 and 5 are more divergent at the genomic level (Derbyshire *et al*. 1975; Hayes *et al*. 1990). PAdV 1–3 belong to PAdV-A. PAdV serotype 4 belongs to PAdV-B. PAdV-5 belongs to PAdV-C. Experimental infections of piglets with PAdV-A are subclinical or associated with intestinal diseases (Sanford and Hoover 1983). PAdV-B causes lesions in several organs including lungs, kidney and brain (Kasza 1966; Shadduck *et al*. 1967). PAdV-C mainly causes respiratory symptoms such as sneezing with nasal discharges and coughing (Hirahara *et al*. 1990). Among all the PAdVs, PAdV-B, which was firstly isolated by Kasza *et al*., is considered as the most pathogenic PAdV and is widely spreading in the United States, Canada, Germany, Hungary, Bulgaria, the United Kingdom, China and the Netherlands (Fig. 1) (Kasza 1966; Elazhary *et al*. 1985). The full genomic sequence of PAdV-B is critical for the taxonomic and pathogenic studies of these viruses.

Until now, the complete nucleotide sequences were available for the genomes of PAdV-A, PAdV-C (Reddy *et al*. 1998; Nagy *et al*. 2001). For PAdV-B, partial genomic sequences of PAdV-B strain NADC-1 (PVIII, E3 13.4 k, E1B and fiber) and strain Kasza (5′ITR) have also been clarified (Kasza 1966; Kleiboeker 1994). Recently, two novel PAdV strains, SVN1 and WI, have been reported, but just partial pol and hexon were identified, thus the classification of these two strains is uncertain (Sibley *et al*. 2011; Jerman *et al*. 2014).

The genomic sequence and characteristics, evolutionary status, and tropism of PAdV-B have been the pending problems for a long time. In the present study, we have sequenced the genome of a novel PAdV-B strain, PAdV-B-HNU1, and annotated 31 putative genes. This is the first characterization of the complete genomic sequence for PAdV-B. According to the full genome of PAdV-B-HNU1, PAdV-SVN1 and PAdV-WI, two PAdVs with ambiguous taxonomy previously, can be classified into PAdV-B. In addition, the lymphoid tropism of PAdV-B-HNU1 was also preliminarily revealed by the epidemiological data, which may contribute to further pathogenic studies of PAdV-B.

Materials and Methods

Sample Collection

To investigate porcine viral pathogens, the porcine samples were collected from pigs showing diarrhea in the Hubei Province, China, during 2017–2018. Totally, 180 lymph nodes, 59 oral swabs and 56 rectal swabs were collected. Virus transport medium (VTM) was made by adding 1% fetal bovine serum (Hyclone) and 1% penicillin–streptomycin (10, 000 U/mL, Gibco) to HBSS medium (Hyclone). The oral and rectal samples were sampled using aseptic swabs and stored in the cryopreservation tubes containing 1 mL VTM. Tissue samples were kept in the cryopreservation tube direct. All the samples were immediately put in dry ice, then transported to the laboratory and stored at –80 °C.

Viral Metagenomic Assay

To perform the viral metagenomic analysis, a library was constructed using pooled lymph samples. Briefly, 100 mg tissue of each sample was homogenized using 1 mL phosphate buffer saline, centrifuged, pooled, and filtered. Viral particles in filtrate enriched by ultracentrifugation, and treated with DNase and RNase to digest unprotected nucleic acid. Then viral DNA/RNA was extracted using QIAamp Viral DNA/RNA Mini Kit (Qiagen, Germany) and subjected to random PCR (rPCR) as previously described (Li *et al*. 2010). The purified rPCR products were used to construct the sequencing library and sequenced on HiSeq-PE150 instrument (Illumina platform). The raw reads were debarcoded, trimmed, and *de novo* assembled using Geneious software package (Version10.2.2) (Kearse

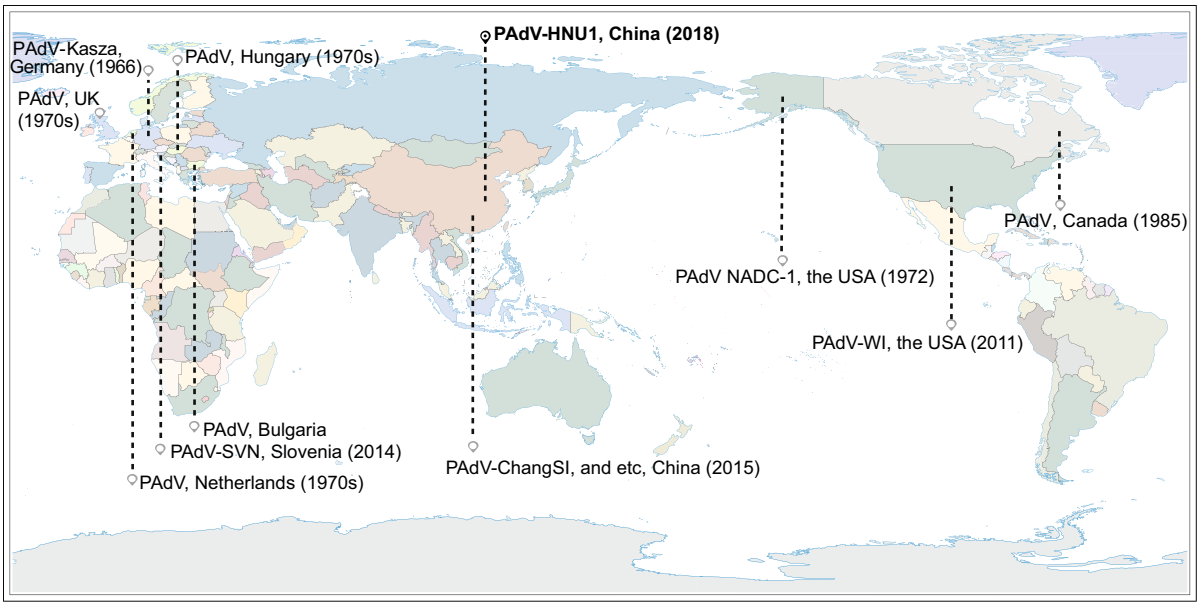

Fig. 1 Epidemics of PAdV-B in the world. The location where PAdV-B-HNU1 was detected is shown in symbol.

et al. 2012). The assembled contigs were aligned to NCBI nr database using BLASTx with an *E*-value less than 10^{-5}.

PCR Screening

For further screening the presence of PAdV-B-HNU1, a set of primers to amplify the partial penton gene (F-5′-GTT AGG CTG TAT GAC TCT GTT GAG-3′ and R-5′-CCT AAA CTC CCC ACC CCA GTT AGA TCT C-3′), were designed according the AdV *penton* gene detected by metagenomic sequencing. Viral DNA was extracted from individual sample with TIANamp Virus DNA Kit (TIANGEN, China) according to the manufacture's instruction. The 15-μL reaction mix contained 2 μL of extracted DNA, 5 pmol each primer (pentonF and R), 7.5 μL 2 × Taq PCR Mastermix (TIANGEN, China). After an initial incubation step at 94 °C for 5 min, 40 cycles of amplification were carried out, consisting of denaturation at 94 °C for 30 s, annealing at 58 °C for 30 s, and extension at 72 °C for 40 s, and a final extension step at 72 °C for 5 min. PCR products were run in a 1.5% agarose gel. The expected products were cut and extracted using a TIANgel Midi Purification Kit (TIANGEN, China), then sequenced with the two primers.

Full-Length Genome Sequencing

One of the lymph samples positive for PAdV-B-HNU1 was used for genome sequencing. The extracted viral DNA was used as a template for PCR amplification with a combination of primer sets (available upon request) designed for sequencing the genome of PAdV-B-HNU1, according to the metagenomic contig sequences. The expected products were cut and extracted using a TIANgel Midi Purification Kit (TIANGEN, China) and sequenced directly. To confirm the terminal sequences of the viral genome, the end of the extracted viral DNA were tailed using SMARTer RACE 5′/3′Kit (Takara, Japan) and amplified according to the manufacturer's instruction (Elsing and Burgert 1998). PCR products were run by electrophoresis and were sequenced with the genome-specific primers. All the sequences were assembled to a full-length genome.

Gene Annotation and Phylogenetic Analysis

By ORF Finder (National Center for Biotechnology Information), ORFs that code for peptides > 50 aa were identified as potential genes. Gene sequences and encoded putative protein aa sequences were compared to that in the NCBI GenBank using the Basic Local Alignment Search Tool (BLASTn and BLASTp) and the homologous ORFs in other adenoviruses were used to confirm potential genes. Protein identity comparisons were performed by using ClustalW and corrected manually (Thompson *et al*. 2002). Models of evolution were evaluated using a corrected Aikakeinformation criteria (AICc) in ProtTest3.4.2 to determine the best amino acid substitution model (Darriba *et al*. 2011). Maximum Likelihood analysis was run in

PhyML 3.0 with the best model of evolution according to AICc, using 1000 bootstrap replicates to test the strength of the tree topology (Guindon *et al*. 2010). Meanwhile, the Neighbor Joining tree was run in MEGA7.0 (Kumar *et al*. 2016). Trees were edited using FigTree v1.4.3.

Nucleotide Sequence Accession Numbers

The following published AdV genome sequences were retrieved from NCBI and included in the analysis in this study: Bat AdV-TJM (NC_016895), Bat AdV-WIV10 (NC_029899), Bat AdV-WIV13 (NC_030874), Bovine AdV-3 (NC_001876), Bovine AdV-2 (NC_002513), California sea lion AdV-1 (NC_024150), Canine AdV-1 (NC_001734), Cervid AdV-1 (NC_030792), Chimpanzee AdV-Y25 (NC_017825), Cynomolgus AdV-1 (NC_034382), Duck AdV-2 (NC_024486), Equine AdV-1 (NC_030792), Fowl AdV-5 (NC_021221), Frog AdV-1 (NC_002501), Goose AdV-4 (NC_017979), Human AdV-1 (AC_000017), Human AdV-54 (NC_012959), Human AdV-7 (NC_004001), Lizard AdV-2 (NC_024684), Murine AdV-3 (NC_012584), Odocoileus AdV-1 (NC_035619), Ovine AdV-7 (NC_004037), Pigeon AdV-2 (NC_031503), Porcine AdV-3 (NC_005869), Porcine AdV-5 (NC_002702), Raptor AdV-1 (NC_015455), Simian AdV-3 (NC_020487), Skunk AdV-PB1 (NC_027708), Squirrel AdV-1 (NC_035207), Titi monkey AdV-ECC-2011 (NC_020487), Tree shrew AdV-1 (NC_004453), Turkey AdV-3 (NC_001958), Turkey AdV-4 (NC_022612), Turkey AdV-1 (NC_014564), Porcine AdV-WI (JF699045), Porcine AdV-SVN (KJ933482 and KJ499459).The genome sequences of PAdV-B-HNU1 sequenced in this study were submitted to GenBank (accession no. MK774519).

Results

Detection of a Novel HNU1 Strain of PAdV-B by Viral Metagenomics

To investigate the viral infection in the mass lymph node samples collected from pigs, the viral library was constructed with pooled lymph samples of porcine and subsequently sequenced by Illumina platform. The sequencing generated 8,613,350 reads with 150 nt length, with an approximate total coverage of 32 ×. After debarcoding and trimming, a total of 6,673,308 clean reads were obtained. By *de novo* assembly, a long contig was obtained from 6,831 cleaned paired reads. The contig contains genes and encoded proteins that are significantly related to homologues of AdVs in the GenBank. This virus was designated tentatively as PAdV-B-HNU1.

Prevalence of PAdV-B-HNU1 in Porcine

To reveal the prevalence of PAdV-B-HNU1 in clinical porcine samples, a set of primers targeting a 349-bp region within the penton base gene, were used to detect PAdV-B-HNU1 in 295 porcine samples. The result showed that 108 out of 295 samples (36.6%) were positive for PAdV-B-HNU1, including 93 of 180 (51.7%) lymph nodes, 6 of 59 (10.2%) oral swabs and 9 of 56 (16.1%) rectal swabs (Table 1).

Characterization of the Complete Genome of PAdV-B-HNU1

The full-length genome of PAdV-B-HNU1 was reconstructed by PCR amplification and replicon sequencing using primers designed according to the viral metagenomic contig sequence. The terminal sequences of the viral genome were determined by 5′ and 3′ RACE amplification and sequencing. After assembly, a 31,743 bp genome of PAdV-B-HNU1 (GenBank no. MK774519) with 55% GC content was obtained.

Genome Analysis of PAdV-B-HNU1

Totally, 116 ORFs (open reading frames) were predicted in both strands of the PAdV-B-HNU1 genome using a searching strategy defining ATG as the start codon and a cutoff size of 50 amino acids (aa). After annotation of the ORFs, the genomic organization of PAdV-B-HNU1 was similar to that of most known members in the genus *Mastadenovirus* (Fig. 2). The GC contents vary from 46 to 65% in different genes, among which genes in the center of the genome, stretching from IVa2 to pVIII, showed higher GC contents than the genes at the terminal regions. Among all the putative proteins encoded by PAdV-B-HNU1, the aa sequences of 31 products were homologous to those of other AdVs, of which 21 were located at the positive strand and 10 were located at the complementary strand (Table 2, Fig. 2). Next, genes of PAdV-B-HNU1 were classified into two sets, homologous genes in all genera as 'genus-common genes' and other genes as 'genus-specific genes' (Davison *et al*. 2003). Similar to other members of the genus *Mastadenovirus*, the middle part of the PAdV-B-HNU1 genome was predicted to contain 18 genus-common genes. These include genes coding for protein related to DNA replication (Pol, pTP, and DBP), DNA encapsidation (52 K and Iva2), virion formation and capsid structure (pIIIa, penton base, pVII, pX, pVI, hexon, protease, 100 K, 33 K, pVIII, and fiber); 22 K, which originates from a lack of splicing in 33 K; and U exon which has been lost in PAdV-C. And PAdV-B-HNU1 had the longest inverted

Table 1 Prevalence of PAdV-B-HNU1 in porcine samples collected in China.

Sample types	No. of test samples/no. of positive samples (% positive)
Lymph	180/93 (51.7)
Oral swab	59/6 (10.2)
Rectal swab	56/9 (16.1)
Total	295/108 (36.6)

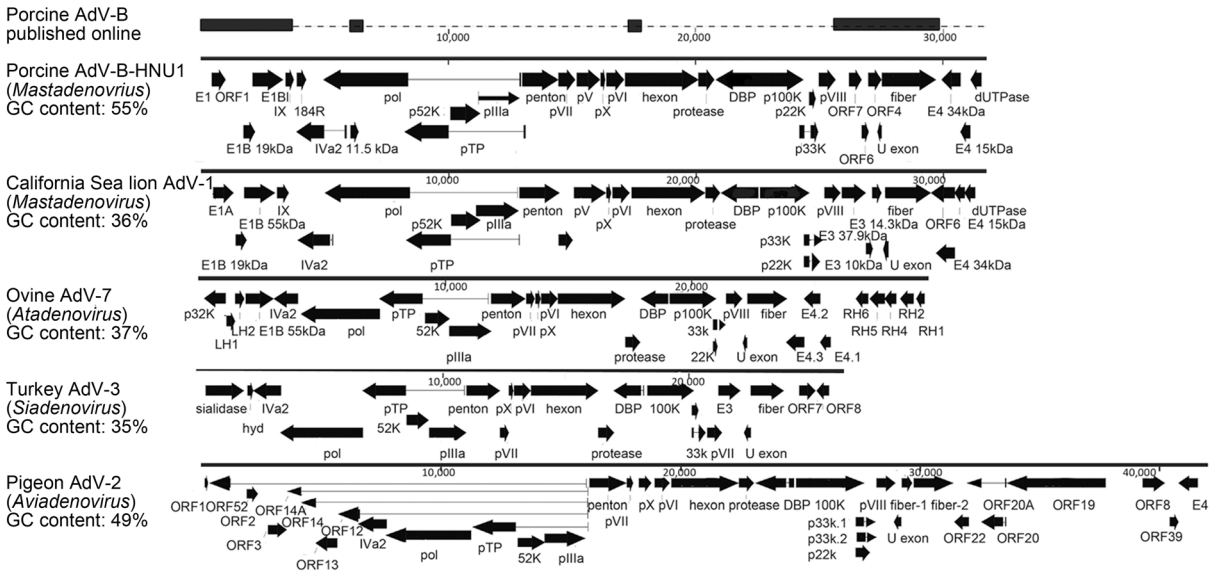

Fig. 2 Genome organization of PAdV-B-HNU1 and other adenoviruses. The viral genome is represented by the thick line in the center marked with 10-kb intervals. Rectangles represent genome regions of PAdV-B published online and the dashed line indicates unavailable region of PAdV-B. Arrows underneath the genome line predicted ORFs of adenoviruses and the thin lines indicate the introns.

terminal repeat compared to other porcine adenoviruses (Fig. 3).

A summary of the sequence identity among proteins of PAdV-B-HNU1 and other porcine adenoviruses including California sea lion AdV-1 is presented in Table 2. The amino acid sequences (94–100%) of PAdV-B-HNU1 showed the highest similarity to PAdV-B strain NADC-1 compared to other porcine adenoviruses including PAdV-A and PAdV-C, suggesting that PAdV-B-HNU1 belongs to PAdV-B.

Classification of PAdV-SVN1 and PAdV-WI Based on PAdV-B-HNU1 Genomic Sequence

Several PAdV strains, such as PAdV-SVN1 and PAdV-WI, have not been classified into any species, because only part of their genome has been sequenced for which the species-reference sequence is lacking. The PAdV-B-HNU1 provides the first representative complete genomic sequence of PAdV-B, which can be used to classify related PAdV strains. The classification of the PAdV-SVN1 and PAdV-WI is made by comparing the aa sequence identities. According to multiple alignment of partial DNA polymerase (Fig. S1A) and hexon (Fig. S1B) aa sequence, partial DNA polymerase of PAdV-SVN1 shows 99%, 76%, and 74% identity to PAdV-B-HNU1, PAdV-A, and PAdV-C, respectively, while partial hexon of PAdV-WI shows 93%, 93%, 56%, and 56% identity to PAdV-SVN1, PAdV-B-HNU1, PAdV-A, and PAdV-C, respectively. This suggests that PAdV-SVN1 and PAdV-WI, the two PAdVs whose taxonomic position was uncertain, belong to PAdV-B.

Phylogenetics of PAdV-B-HNU1

Phylogenetic analysis was conducted to reveal the genetic and evolutionary status of PAdV-B-HNU1. Phylogenetic trees were constructed based on the amino acid sequences of DNA polymerase, hexon, penton base and pre-terminal protein. In these trees, the PAdV-B-HNU1 strain was clustered with C.sea lion AdV-1 with high bootstrap value, indicating that PAdV-B-HNU1 is a member of the genus *Mastadenovirus* (Fig. 4, Fig. S2). And the genome of PAdV-B-HNU1 shared highest nucleotide sequence identity (54.1%) with C. sea lion AdV-1 (Table S1).

A dUTPase was identified at the 3′ end of PAdV-B-HNU1 genome, which was more similar to dUTPases of bacteria and fungi than that of mastadenoviruses except C.

Table 2 Putative proteins encoded by PAdV-B-HNU1.

Gene product	Location(s) (nt)[a]		Size (aa)	Amino acid identity (%) with:				GC content (%)
				PAdV-NADC1	C.sea lion AdV-1	PAdV-A	PAdV-C	
E1 ORF 1	Termini	485–1039	186	94	29	22	23	55
E1B 19 kDa		1762–2220	152	NI	26	14	NI	54
E1B 55 K		2127–3368	413	95	38	24	28	64
IX		3466–3801	111	100	29	24	26	56
184R		3981–4280	99	NA	NI	50	NI	48
ORF6		26,760–27,041	93	96	NI	NI	NI	46
ORF7		26,242–26,760	172	97	NI	NI	NI	56
ORF4		27,022–27,561	119	99	NI	NI	NI	49
U exon		27,560–27393c	55	100	53	38	NI	48
Fiber		27,575–29,764	729	94	28	14	14	55
E434kDa		30,753–30004c	249	NA	36	28	26	51
dUTPase		31,592–31164c	145	NA	50	NI	40	62
E415kDa		31,151–30750c	133	NA	30	NI	NI	53
IVa2	Center	3891–5218, 5298–5310c	373	NA	69	58	65	50
pol		4982–8410,12,937–12966c	1037	NA	69	65	55	55
11.5 kDa		6087–6407	106	NA	NI	NI	NI	54
p52k		10,128–11,339	403	NA	61	47	48	56
Penton base		13,035–14,483	482	NA	76	60	63	53
pV		15,234–16,175	313	NA	28	14	23	61
pX		16,197–16,391	64	NA	70	48	59	57
pVI		16,428–17,159	244	NA	48	34	44	61
pTP		8263–10,095, 13,097–13162c	595	NA	72	55	58	59
pIIIa		11,266–12,927	554	NA	62	43	48	58
p33K		24,249–24,423, 24,690–25,000	161	NA	52	29	NI	64
p22K		24,625–24,912	95	NA	48	29	NI	55
pVII		14,486–15,178	230	NA	48	27	24	65
Hexon		17,201–20,161	986	NA	77	63	66	57
protease		20,165–20,791	208	NA	64	63	57	54
DBP		22,379–20853c	508	NA	55	41	39	65
p100K		22,394–24,403	669	NA	64	50	54	58
pVIII		25,031–25,696	221	94	57	50	52	62

NA not applicable due to the lack of detectable similarity, *NI* homologous gene not identified.

[a]The letter "c" in this column indicates that those genes are encoded by the complementary strand.

Fig. 3 Comparison of nucleotide sequences of 5′UTR between PAdV-B-HNU1, PAdV-C, PAdV-A and C. sea lion AdV-1. Pairwise alignments were calculated using DNAMAN. Dots indicate gaps.

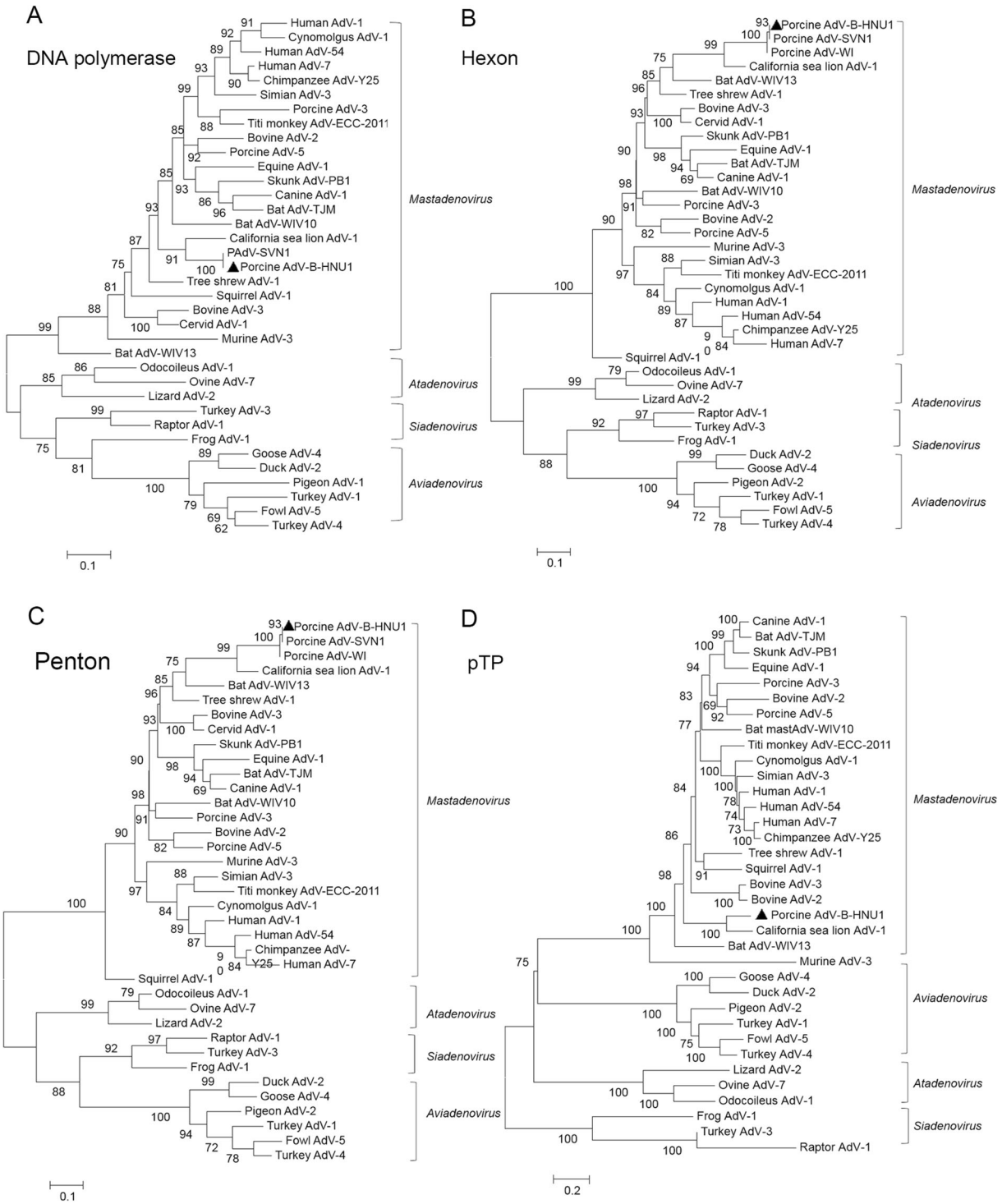

Fig. 4 The evolutionary position of PAdV-B-HNU1 in the phylogenetic tree of adenoviruses. The phylogenetic trees were constructed based on the amino acid sequences of DNA polymerase (**A**), hexon (**B**), penton base (**C**) and pTP (**D**) of porcine adenoviruses by using the maximum likehood (ML) method with rtREV + F + I, LG + F + R5, LG + I + G + F, and LG + G as the model of protein evolution, respectively. PAdV-B-HNU1 detected in this study was indicated with solid triangle.

sea lion AdV-1. Phylogenetic analysis further confirmed that PAdV-B-HNU1 dUTPase does not cluster with other mastadenoviral dUTPases and was more related to dUTPases of baculovirus, poxvirus and eukaryotes (Fig. 5). This suggested that PAdV-B-HNU1 is more closed to C. sea lion AdV-1 than other PAdVs in evolutionary positions.

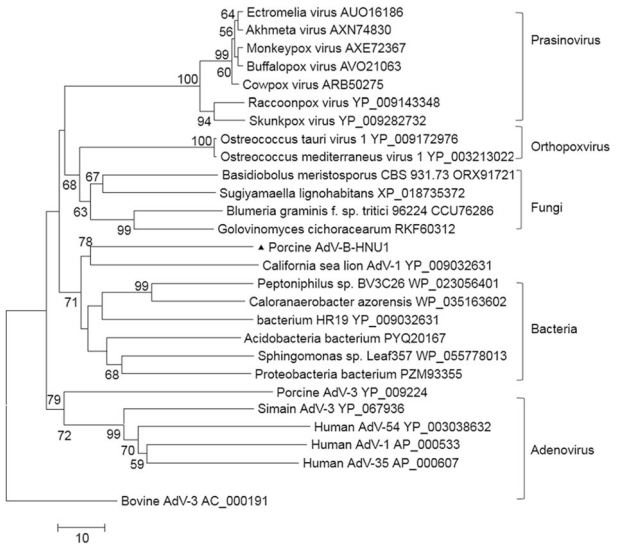

Fig. 5 Phylogenetic tree constructed on the amino acid sequences of dUTPase. Phylogenetic analysis of PAdV-B-HNU1 is based on the partial conserved amino acid sequences of dUTPase from selected adenoviruses, bacteria, fungi, prasinovirus and poxviruses by using the Maximum Likelihood method based on the JTT matrix-based model. PAdV-B-HNU1 detected in this study was indicated with solid triangle.

Discussion

Three PAdVs species, PAdV-A, PAdV-B and PAdV-C, have been established based on genomic features and phylogeny. Currently, the genomes of PAdV-A and PAdV-C have been completely sequenced and well annotated, but the genome of PAdV-B has not been identified yet. Here we presented the first complete genomic sequence of PAdV-B-HNU1, a novel strain of PAdV-B.

Sequence analysis of PAdV-B-HNU1 genome indicated that the overall genomic structure of PAdV-B-HNU1 is similar to that of most known AdVs. Phylogenetic analysis based on the DNA polymerase sequence and the penton base protein sequence classified PAdV-B-HNU1 into the genus *Mastadenovirus*, and showed that PAdV-B-HNU1 is most closely related to C. sea lion AdV-1. However, the average GC content of PAdV-B-HNU1 was 55%, higher than that of C. sea lion AdV-1 (36%). In different genes of PAdV-B-HNU1, the GC content ranged from 46% to 65%, with high rates in the center of the genome and low rates at terminal regions. High abundance of 5′-C-phosphate-G-3′ (CpG) in the genomes was responsible for GC contents (Tan *et al*. 2017). The center of adenoviruses genome contained structural proteins and enzymes of which high expression levels were required for DNA replication, DNA encapsidation and the formation of the virion (Hoeben and Uil 2013). High GC content in certain ORFs may increase CpG islands, affect viral DNA methylation and subsequently modulate gene expression. However, this speculation lacked any evidence currently and we would focus on it in our future studies. The ITRs played an important role in the initiation of replication by providing binding sites for viral and cellular regulatory proteins in adenoviral infection (Dán *et al*. 2001). The first 17 nucleotides of this region were identical in PAdV-5 and PAdV-3, but they are less conserved in PAdV-B-HNU1.

Fiber protein exists in all adenoviruses for tethering of the viral capsid to the cell surface via its interaction with the cellular receptor such as CAR (Roelvink *et al*. 1998), or CD46/80/86 (Short *et al*. 2006), or sialic acid (Arnberg *et al*. 2000). The carboxy-terminal end of the predicted fiber of PAdV-B-HNU1 showed similarity to galectin-4 and galectin-9 of animals, which was similar to PAdV NADC-1 having a fiber containing a C-terminal galectin domain (Guardado-Calvo *et al*. 2010). Galectin-4 plays an important role in lipid raft stabilization, protein apical trafficking, cell adhesion, wound healing, intestinal inflammation, tumor progression (Cao and Guo 2016). Galectin-9 can interact with CD40 on T-cells induced their proliferation inhibition and cell death (Vaitaitis and Wagner 2012). Whether the PAdV-B-HNU1 fiber gene plays a similar role remains to be determined.

dUTPases were ubiquitously expressed enzymes in eukaryotic and prokaryotic cells as well as viruses. dUTPase catalyzes the conversion of dUTP to dUWP and PPi, reducing dUTP/dTTP ratio to prevent misincorporation of uracil into DNA (Koonin 1996). Viral dUTPases were captured by some viruses via horizontal gene transfer (Baldo and Mcclure 1999). Viral dUTPases showed species specificity and have been discovered in herpesviruses (Glaser *et al*. 2006), poxviruses (Cottone *et al*. 2002) and retroviruses (Payne and Elder 2001). The dUTPase activity has been confirmed in the case of fowl adenovirus 1, and fowl adenovirus 9 dUTPase upregulated the expression of type I Interferons (Weiss *et al*. 1997; Deng *et al*. 2016). PAdV-B-HNU1 dUTPase was more divergent compared to those of other PAdVs. Similar to C. sea lion AdV-1, the predicted PAdV-B-HNU1 dUTPase was more closely related to dUTPase of bacteria and fungi than those of mastadenoviruses. This indicated a potential independent horizontal gene transfer event that bacterial dUTPase may be an endogenous viral element derived from adenovirus. Whether the PAdV-B-HNU1 dUTPase shares similar biological activity like other adenoviruses remains to be determined.

The genomic information for fiber and dUTPase is not available for all porcine adenoviruses. We proposed the HNU1 strain as the reference strain of PAdV-B. Porcine adenovirus strain PAdV-WI which was detected in pen wash water of newborn to finisher pigs and was only available for a fragment of hexon gene, has been proposed as a prototype of a new species in the *Mastadenovirus*

genus (Sibley *et al*. 2011). In addition, PAdV-SVN1 was detected in urinary bladder urothelial cell culture and was available for partial polymerase and hexon (Jerman *et al*. 2014). Limited phylogenetic analysis showed that PAdV-SVN1 clustered together with PAdV-WI, that was clearly separated from other *Mastadenovirus* representative strains (Jerman *et al*. 2014). The clear classification of PAdV-SVN1 as well as PAdV-WI is not possible until the whole genome sequence of PAdV-B-HNU1 is available. According to the amino acid sequence of the DNA polymerase and hexon in this study, the PAdV-SVN1 and PAdV-WI are closely related to HNU1 strain. We proposed PAdV-SVN1 and PAdV-WI can be classified into PAdV-B. With the completion of full genomes from all porcine adenovirus species, additional insight into the classification and evolution were obtained.

C. sea lion AdV-1 caused acute hepatitis in sea lions and other fins (Goldstein *et al*. 2011; Inoshima *et al*. 2013), which was predicted to jumped to California sea lions from an unknown mammalian endemic host (Cortéshinojosa *et al*. 2015). Similarities between PAdV-B-HNU1 and C. sea lion AdV-1 were apparent not only in their genome organization, but also in their phylogenetic relationship. Phylogenetic analyses showed that members of PAdV-B were closely related to C. sea lion AdV-1 with the similar dUTPase gene. Further investigation of infectious characters of C. sea lion AdV-1 and PAdV-B-HNU1 was blocked by the failure of virus isolation. More studies, such as infecting swine cells and infecting sea lion cells with C. sea lion AdV-1 and PAdV-B-HNU1, respectively, will contribute to better understanding of the and transmission of these viruses.

In all, our study on PAdV-B-HNU1 has provided the first representative sequence and annotation of PAdV-B genome, contributing to the classification of PAdVs. For instance, PAdV-SVN1 and PAdV-WI, two unclassified PAdV strains, can be classified into PAdV-B species comparing to the genomic sequence of PAdV-B-HNU1. In addition, the PAdV-B-HNU1 was probably a lymphotropic adenovirus, which may explain the specific pathogenic pattern of PAdV-B-HNU1 infection. Considering that PAdV-Bs widely spread in different countries around the world, our findings would contribute to the future epidemic studies of PAdV-B.

Acknowledgements This work was jointly funded by the National Key Research and Development Program of China (grant number 2017YFD0500104), the Hu-Xiang Youth Talents Scholar Program of Hunan Province (grant number 2017RS3017), the Science Fund for Distinguished Young Scholars of Hunan Province (grant number 2019JJ20004), National Natural Science Foundation of China (grant number 81902070), the Provincial Natural Science Foundation of Hunan Province (grant number 2019JJ50035) and the Fundamental Research Funds for the Central Universities of China (grant number 531107051162).

Author Contributions Conceptualization, SL and XG; Methodology, SL, QW, TL, SZ, JL, YQ and XG; Investigation, SL, QW, TL, SZ, JL, LW, YQ and XG; Writing—Original Draft, SL; Writing—Review & Editing, SL, YQ and XG; Funding Acquisition, YQ and XG; Resources, TL, SZ, LW and XG; Supervision, YQ and XG.

Compliance with Ethical Standards

Conflict of interest The authors declare that there are no conflicts of interest.

Animal and Human Rights Statement All institutional and national guidelines for the care and use of laboratory animals were followed.

References

Arnberg N, Kidd AH, Edlund K, Olfat F, Wadell G (2000) Initial interactions of subgenus D adenoviruses with A549 cellular receptors: sialic acid versus alpha(v) integrins. J Virol 74:7691

Baldo AM, Mcclure MA (1999) Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts. J Virol 73:7710–7721

Benkö M, Harrach B (2003) Molecular evolution of adenoviruses. Curr Top Microbiol Immunol 272:3

Cao ZQ, Guo XL (2016) The role of galectin-4 in physiology and diseases. Protein Cell 7:314–324

Chen EC, Yagi S, Kelly KR, Mendoza SP, Tarara RP, Canfield DR, Maninger N, Rosenthal A, Spinner A, Bales KL (2011) Cross-species transmission of a novel adenovirus associated with a fulminant pneumonia outbreak in a new world monkey colony. PLoS Pathog 7:e1002155

Cortéshinojosa G, Gulland FM, Goldstein T, Vennwatson S, Rivera R, Waltzek TB, Salemi M Jr., Wellehan JF (2015) Phylogenomic characterization of California sea lion adenovirus-1. Infect Genet Evol 31:270–276

Cottone R, Büttner M, Mcinnes CJ, Wood AR, Rziha HJ (2002) Orf virus encodes a functional dUTPase gene. J Gen Virol 83:1043–1048

Dán A, Elo P, Harrach B, Zádori Z, Benko M (2001) Four new inverted terminal repeat sequences from bovine adenoviruses reveal striking differences in the length and content of the ITRs. Virus Genes 22:175

Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27:1164–1165

Davison AJ, Wright KM, Harrach B (2000) DNA sequence of frog adenovirus. J Gen Virol 81:2431

Davison AJ, Benko M, Harrach B (2003) Genetic content and evolution of adenoviruses. J Gen Virol 84:2895–2908

Deng L, Qin XB, Krell P, Lu R, Sharif S, Nagy E (2016) Characterization and functional studies of fowl adenovirus 9 dUTPase. Virology 497:251–261

Derbyshire JB, Clarke MC, Collins AP (1975) Serological and pathogenicity studies with some unclassified porcine adenoviruses. J Comp Pathol 85:437–443

Elazhary MA, Dea S, Mittal KR, Higgins R (1985) Prevalence of antibodies to swine influenza virus, porcine adenovirus type 4 and haemophilus pleuropneumoniae in quebec pig farms with respiratory problems. Can Vet J 26:190–192

Elsing A, Burgert HG (1998) The adenovirus E3/10.4K–14.5K proteins down-modulate the apoptosis receptor Fas/Apo-1 by

inducing its internalization. Proc Natl Acad Sci USA 95:10072–10077
Ghebremedhin B (2014) Human adenovirus: viral pathogen with increasing importance. Eur J Microbiol Immunol 4:26–33
Glaser R, Litsky ML, Padgett DA, Baiocchi RA, Yang EV, Chen M, Yeh PE, Green-Church KB, Caligiuri MA, Williams MV (2006) EBV-encoded dUTPase induces immune dysregulation: Implications for the pathophysiology of EBV-associated disease. Virology 346:205–218
Goldstein T, Colegrove KM, Hanson M, Gulland FM (2011) Isolation of a novel adenovirus from California sea lions *Zalophus californianus*. Dis Aquat Organ 94:243–248
Guardado-Calvo P, Llamas-Saiz AL, Fox GC, Glasgow JN, van Raaij MJ (2010) Crystallization of the head and galectin-like domains of porcine adenovirus isolate NADC-1 fibre. Acta Crystallogr Sect F Struct Biol Cryst Commun 65:1149–1152
Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321
Haig DA, Clarke MC, Pereira MS (1964) Isolation of an adenovirus from a pig. J Comp Pathol 74:81
Harrach BBM, Both GW, Brown M, Davison AJ, Echavarría M, Hess M, Jones MS, Kajon A, Lehmkuhl HD, Mautner V, Mittal SK, Wadell G (2011) Family *Adenoviridae*. In: King A, Adams M, Lefkowitz E (eds) Virus taxonomy: classification and nomenclature of viruses—ninth report of the international committee on taxonomy of Viruses. Elsevier, San Diego, pp 25–41
Hayes BW, Telling GC, Myat MM, Williams JF, Flint SJ (1990) The adenovirus L4 100-kilodalton protein is necessary for efficient translation of viral late mRNA species. J Virol 64:2732
Hess M, Blöcker H, Brandt P (1997) The complete nucleotide sequence of the egg drop syndrome virus: an intermediate between mastadenoviruses and aviadenoviruses. Virology 238:145
Hirahara T, Yasuhara H, Matsui O, Yamanaka M, Tanaka M, Fukuyama S, Izumida A, Yoshiki K, Kodama K, Nakai M (1990) Isolation of porcine adenovirus from the respiratory tract of pigs in Japan. Nihon Juigaku Zasshi 52:407
Hiroaki I, Yasushi S, Tsunetada K, Akio H, Takeshi O, Yoshitsugu T, Koki A, Shigeaki O, Shudo Y (2008) Novel human adenovirus causing nosocomial epidemic keratoconjunctivitis. J Clin Microbiol 46:2002
Hoeben RC, Uil TG (2013) Adenovirus DNA replication. Cold Spring Harb Perspect Biol 5:a013003
Inoshima Y, Murakami T, Ishiguro N, Hasegawa K, Kasamatsu M (2013) An outbreak of lethal adenovirus infection among different otariid species. Vet Microbiol 165:455–459
Jerman UD, Kolenc M, Steyer A, Veranic P, Prijatelj MP, Kreft ME (2014) A novel strain of porcine adenovirus detected in urinary bladder urothelial cell culture. Viruses 6:2505–2518
Kasza L (1966) Isolation of an adenovirus from the brain of a pig. Am J Vet Res 27:751–758
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649
Kleiboeker SB (1994) Sequence analysis of putative E3, pVIII, and fiber genomic regions of a porcine adenovirus. Virus Res 31:17–25
Kohl C, Vidovszky MZ, Muhldorfer K, Dabrowski PW, Radonic A, Nitsche A, Wibbelt G, Kurth A, Harrach B (2012) Genome analysis of bat adenovirus 2: indications of interspecies transmission. J Virol 86:1888–1892
Koonin EV (1996) Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminases. Nucleic Acids Res 24:2411–2415
Kovács GM, Lapatra SE, D'Halluin JC, Benko M (2003) Phylogenetic analysis of the hexon and protease genes of a fish adenovirus isolated from white sturgeon (Acipenser transmontanus) supports the proposal for a new adenovirus genus. Virus Res 98:27–34
Kovács ER, Jánoska M, Dán Á, Harrach B, Benkő M (2010) Recognition and partial genome characterization by non-specific DNA amplification and PCR of a new siadenovirus species in a sample originating from Parus major, a great tit. J Virol Methods 163:262–268
Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874
Li Y, Ge X, Zhang H, Zhou P, Zhu Y, Zhang Y, Yuan J, Wang LF, Shi Z (2010) Host range, prevalence, and genetic diversity of adenoviruses in bats. J Virol 84:3889–3897
Matsushima Y, Shimizu H, Kano A, Nakajima E, Ishimaru Y, Dey SK, Watanabe Y, Adachi F, Mitani K, Fujimoto T, Phan TG, Ushijima H (2013) Genome sequence of a novel virus of the species human adenovirus d associated with acute gastroenteritis. Genome Announc 1:e00068–00012
Nagy M, Nagy E, Tuboly T (2001) The complete nucleotide sequence of porcine adenovirus serotype 5. J Gen Virol 82:525–529
Payne SL, Elder JH (2001) The role of retroviral dUTPases in replication and virulence. Curr Protein Pept Sci 2:381–388
Reddy PS, Idamakanti N, Song JY, Lee JB, Hyun BH, Park JH, Cha SH, Bae YT, Tikoo SK, Babiuk LA (1998) Nucleotide sequence and transcription map of porcine adenovirus type 3. Virology 251:414–426
Roelvink P, Lizonova A, Lee J, Li Y, Bergelson J, Finberg R, Brough D, Kovesdi I, Wickham T (1998) The coxsackievirus-adenovirus receptor protein can function as a cellular attachment protein for adenovirus serotypes from subgroups A, C, D, E, and F. J Virol 72:7909–7915
Sanford SE, Hoover DM (1983) Enteric adenovirus infection in pigs. Can J Comp Med 47:396–400
Shadduck JA, Koestner A, Kasza L (1967) The lesions of porcine adenoviral infection in germfree and pathogen-free pigs. Pathol Vet 4:537
Short JJ, Vasu C, Holterman MJ, Curiel DT, Pereboev A (2006) Members of adenovirus species B utilize CD80 and CD86 as cellular attachment receptors. Virus Res 122:144–153
Sibley SD, Goldberg TL, Pedersen JA (2011) Detection of known and novel adenoviruses in cattle wastes via broad-spectrum primers. Appl Environ Microbiol 77:5001–5008
Tan B, Yang XL, Ge XY, Peng C, Liu HZ, Zhang YZ, Zhang LB, Shi ZL (2017) Novel bat adenoviruses with low G + C content shed new light on the evolution of adenoviruses. J Gen Virol 98:739–748
The HAdV Working Group (2019) http://hadvwg.emu.edu. Accessed 24 July 2019
Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinform Chapter 2:Unit 2.3
Vaitaitis GM, Wagner DH (2012) Galectin-9 controls CD40 signaling through a Tim-3 independent mechanism and redirects the cytokine profile of pathogenic T cells in autoimmunity. PLoS ONE 7:e38708
Weiss RS, Lee SS, Prasad BV, Javier RT (1997) Human adenovirus early region 4 open reading frame 1 genes encode growth-transforming proteins that may be distantly related to dUTP pyrophosphatase enzymes. J Virol 71:1857