



HAL
open science

Phylogenomic analysis of the complete sequence of a gastroenteritis-associated cetacean adenovirus (bottlenose dolphin adenovirus 1) reveals a high degree of genetic divergence

Maja Malmberg, Consuelo Rubio-Guerri, Juliette Hayer, Daniel García-Párraga, Elvira Nieto-Pelegrín, Mar Melero, Teresa Álvaro, Mónica Valls, Jose Manuel Sánchez-Vizcaíno, Sándor Belák, et al.

► To cite this version:

Maja Malmberg, Consuelo Rubio-Guerri, Juliette Hayer, Daniel García-Párraga, Elvira Nieto-Pelegrín, et al.. Phylogenomic analysis of the complete sequence of a gastroenteritis-associated cetacean adenovirus (bottlenose dolphin adenovirus 1) reveals a high degree of genetic divergence. *Infection, Genetics and Evolution*, 2017, 53, pp.47 - 55. 10.1016/j.meegid.2017.05.008 . hal-04409148

HAL Id: hal-04409148

<https://hal.science/hal-04409148>

Submitted on 22 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Research paper

Phylogenomic analysis of the complete sequence of a gastroenteritis-associated cetacean adenovirus (bottlenose dolphin adenovirus 1) reveals a high degree of genetic divergence



Maja Malmberg^{a,b,c,*}, Consuelo Rubio-Guerri^{d,e}, Juliette Hayer^b, Daniel García-Párraga^{e,f}, Elvira Nieto-Pelegrín^d, Mar Melero^d, Teresa Álvaro^f, Mónica Valls^f, Jose Manuel Sánchez-Vizcaíno^d, Sándor Belák^{a,c}, Fredrik Granberg^{a,c}

^a Section of Virology, Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, Box 7028, 750 07 Uppsala, Sweden

^b SLU Global Bioinformatics Centre, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, 750 07 Uppsala, Sweden

^c World Organisation for Animal Health (OIE) Collaborating Centre for the Biotechnology-based Diagnosis of Infectious Diseases in Veterinary Medicine, Box 7028, 750 07 Uppsala, Sweden

^d VISAVET Center and Animal Health Department, Veterinary School, Complutense University of Madrid, Av Puerta del Hierro s/n, 28040 Madrid, Spain

^e Fundación Oceanográfica, Ciudad de las Artes y las Ciencias, Primo Yúfera 1B, 46013, Valencia, Spain

^f Veterinary Services, Avanza Oceanogràfic-Agora, C/Eduardo Primo Yúfera 1B, 46013, Valencia, Spain

ARTICLE INFO

Article history:

Received 19 January 2017

Received in revised form 9 May 2017

Accepted 11 May 2017

Available online 12 May 2017

Keywords:

Adenovirus

Complete genome

E3 region

Bottlenose dolphin

Tursiops truncatus

High-throughput sequencing

Viral metagenomics

ABSTRACT

Adenoviruses are common pathogens in vertebrates, infecting a wide range of hosts, but only having rarely been detected and correlated with disease in cetaceans. This article describes the first complete genomic sequence of a cetacean adenovirus, bottlenose dolphin adenovirus 1 (BdAdV-1), detected in captive bottlenose dolphin population (*Tursiops truncatus*) suffering from self-limiting gastroenteritis. The complete genome sequence of BdAdV-1 was recovered from data generated by high-throughput sequencing and validated by Sanger sequencing. The genome is 34,080 bp long and has 220 nucleotides long inverted terminal repeats. A total of 29 coding sequences were identified, 26 of which were functionally annotated. Among the unusual features of this genome is a remarkably long 4380 bp E3 ORF1, that displays no sequence homology with the corresponding E3 regions of other adenoviruses. In addition, the fiber protein only has 26% identity with fiber proteins described in other adenoviruses. Three hypothetical proteins were predicted. The phylogenetic analysis indicates that the closest known relative to BdAdV-1 is an adenovirus detected in bottlenose dolphin (KR024710), with an amino acid sequence identity between 36 and 79% depending on the protein. Based on the phylogenetic analysis, the BdAdV-1 appears to have co-evolved with its host.

The results indicate that BdAdV-1 belongs to the *Mastadenovirus* genus of the *Adenoviridae* family, however, it is clearly different from other adenoviruses, especially in the 3'-end of the viral genome. The high degree of sequence divergence suggests that BdAdV-1 should be considered as a novel species in the *Mastadenovirus* genus. The study also demonstrates the usefulness of high-throughput sequencing to obtain full-length genomes of genetically divergent viruses.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Since their discovery in tissues from human adenoids in 1953 (Rowe et al., 1953), adenoviruses have been recognized as common pathogens

able to infect a broad range of vertebrate hosts (Davison et al., 2003). The viral family *Adenoviridae* is divided into five genera: *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, *Siadenovirus*, and *Ichtadenovirus*. *Mastadenovirus* is the genus within *Adenoviridae* that is specifically associated with infections and diseases of mammalian host species.

All members of the *Adenoviridae* family have non-segmented linear double stranded (ds) DNA genomes ranging in size from 26 to 45 kbp and are characterized by inverted terminal repeats (ITRs) (Davison et al., 2003). The adenovirus particles are of icosahedral symmetry and represent the largest known non-enveloped virions (90–100 nm). The virion has unique fibers with knobs that are attached to the penton

* Corresponding author at: Section of Virology, Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, Box 7028, 750 07 Uppsala, Sweden.

E-mail addresses: maja.malmberg@slu.se (M. Malmberg), crubio@oceanografic.org (C. Rubio-Guerri), dgarcia@oceanografic.org (D. García-Párraga), mar.melero@ucm.es (M. Melero), talvaro@oceanografic.org (T. Álvaro), mvals@oceanografic.org (M. Valls), jmvizcaino@ucm.es (J.M. Sánchez-Vizcaíno), sandor.belak@slu.se (S. Belák), fredrik.granberg@slu.se (F. Granberg).

base of the capsid. These structures aid the attachment of the virus particle to the host cells via host cell surface receptors.

The adenovirus life cycle is divided by the DNA replication process into two phases: an early and a late phase. Both phases are initiated by a primary transcript that is alternatively spliced to generate monocistronic mRNAs compatible with the host's ribosome, allowing for the products to be translated. The early genes are responsible for mainly expressing non-structural, regulatory proteins. These proteins alter the expression of host proteins necessary for DNA synthesis, activate other viral genes, and avoid premature death of the infected cell by the host-immune defences. The late genes produce structural protein components and proteins needed to pack all the genetic material produced during DNA replication. Once the viral components have been replicated, the virus is assembled into its protein shell and released from the cell as a result of cell lysis induced by the virus.

Adenoviruses are associated with a wide range of clinical symptoms and disease manifestations in various host species. Infections with human adenoviruses have been associated with respiratory disease (HAdV-C), gastroenteritis (HAdV-G and F) (Wold, 2007), and conjunctivitis (HAdV-D) (Wold, 2007). Other mammalian adenoviruses have been associated with for example: hepatitis (Cabasso et al., 1954), encephalitis (Knowles, 2011), and various respiratory and enteric syndromes such as pneumoenteritis in Bovidae and in Ovidae (Aldasy et al., 1965; Belak and Palfi, 1974).

Historically, adenoviruses have played a major role in the understanding of basic molecular biology processes, such as understanding of alternative splicing and DNA replication. Identification of novel animal adenoviruses is, in addition to being interesting from an animal health perspective, also providing important knowledge for possible further use of adenoviruses in therapy (Alonso-Padilla et al., 2016). Modified forms of adenoviruses are being developed to treat genetic diseases, to fight cancer, and to deliver vaccines (Majhen et al., 2014; Wold and Toth, 2013).

In cetaceans, adenoviruses were first described in 1979, when adenovirus-like particles were isolated from rectal swabs of a sei whale in the Antarctic (Smith and Skilling, 1979). Thereafter, adenoviruses were isolated from colon samples of two bowhead whales harvested at Barrow, Alaska (Smith et al., 1987), and from the intestines of a beluga from the St. Lawrence estuary (De Guise et al., 1995). None of these cases were however, associated with clinically manifested disease symptoms (Van Bresseem et al., 1999). In marine mammals, up to recently, the only clearly established association between adenovirus and disease was in Californian sea lions (*Zalophus californianus*) suffering from gastroenteritis with fatal hepatitis (Britt et al., 1979; Dierauf et al., 1981; Goldstein et al., 2011; Inoshima et al., 2013). However, in 2015 a novel adenovirus was described causing self-limiting gastroenteritis in captive bottlenose dolphins (*Tursiops truncatus*) in Valencia, Spain. This was the first adenovirus identified in bottlenose dolphins. Sequence analysis based on partial DNA polymerase (320 bp) and hexon gene sequences (435 bp) showed less than 80% sequence identity to previously published sequences (Rubio-Guerri et al., 2015). For clarification, the virus name used in the publication by Rubio-Guerri et al. 2015 (*Tursiops adenovirus*) has been changed to bottlenose dolphin adenovirus 1 (BdAdV-1) in line with the adenovirus nomenclature. In September 2015, a partial genome (29,474 bp) of a bottlenose dolphin adenovirus detected in faeces from a bottlenose dolphin in USA became available in GenBank (KR024710). This sequence will hereafter be referred to as bottlenose dolphin adenovirus 2 (BdAdV-2). A partial genome of harbour porpoise adenovirus 1 (HpAdV-1) was published in March 2017 (KY352473). This virus was detected in The Netherlands, in faeces from stranded harbour porpoises (*Phocoena phocoena*) with no adenovirus-associated lesions (van Beurden et al., 2017).

The high degree of sequence divergence found in the initial sequencing efforts of BdAdV-1 spurred interest in the complete genome. To address important research questions regarding phylogenetic classification, evolution, and tropism of this likely disease-causing

adenovirus we sequenced, assembled and annotated the full genome. Traditional methods for sequencing of whole genomes, such as primer walking, can be particularly challenging when working with genomes that have very low identity with previously identified organisms, and where the size of the genome can vary greatly. To circumvent these challenges, we used a high-throughput sequencing approach.

2. Material and methods

2.1. Samples and sample preparation

As previously described by Rubio-Guerri et al., 2015, four adult bottlenose dolphins kept at the Oceanogràfic in Valencia, Spain, developed anorexia, diarrhoea and vomiting. A novel adenovirus was suspected to be the causative agent, based on seropositivity for antibodies against canine adenovirus 1 and 2, as well as PCR and Sanger sequencing of partial DNA polymerase and hexon adenoviral genome regions (Rubio-Guerri et al., 2015). Faecal samples from one of the severely affected bottlenose dolphins (Animal 3) were collected on day 2 (2/10/14) as previously described (Rubio-Guerri et al., 2015). Animal sampling was conducted in accordance with Spanish and European regulations (RD 53/2013, Directive 2010/63/UE) on animal welfare. A fraction of the faecal sample was shipped to the Swedish University of Agricultural Sciences in Uppsala, Sweden, for further analysis. There, the faecal sample was diluted in ice-cold PBS, homogenized using an Omni homogenizer with disposable soft tissue tips (Omni International, USA) and centrifuged at 3400 rpm for 10 min. The supernatant was transferred to a spin filtrate column (0.22 µm) and centrifuged for 4 min at 3400 rpm. The filtrate was treated with 6 U of TURBO™ DNase (Ambion, Thermo Fisher Scientific, Pittsburgh, PA, USA), in reaction volumes of 100 µl, and incubated for 30 min at 37 °C. A total volume of 400 µl was treated with QIAGEN Protease (dissolved in AVE buffer) and Ambion RNase Cocktail™ for 5 min at room temperature. DNA was extracted using QIAamp MinElute kit (Qiagen, Gilden, Germany). The DNA concentration was measured on Qubit 2.0 Fluorometer using dsDNA HS (high sensitivity) kit.

2.2. Library preparation and sequencing

Sequencing libraries were made using Nextera XT Library Preparation Kit (Illumina, San Diego, CA, USA), normalization and pooling of 2 nM sequencing libraries was done based on concentration measurements from Agilent High Sensitivity DNA Kit (2100 Bioanalyzer, Agilent Technologies, Palo Alto, CA, USA). The pool of sequencing libraries was denatured with NaOH and further diluted with hybridization buffer to a final concentration of 10 pM and spiked with 5% PhiX for diversity. Paired-end sequencing was performed with MiSeq Reagent Kit v3 600 cycles on the MiSeq instrument (Illumina, San Diego, CA, USA). The quality of the dataset was assessed using the FastQC software.

2.3. Assembly of sequence reads

All sequence reads were initially assembled using the MIRA assembler (v4.0.2) (Chevreux et al., 1999) with the standard settings for de novo assembly of Illumina data. The obtained contigs were merged into even longer sequences using CodonCode Aligner (v5.1.5) (CodonCode Corporation) and the “assemble from scratch”-function with default settings. Potential adenoviral sequences were identified by using BLASTn searches against the NCBI non-redundant nucleotide library (NCBI NT).

2.4. Estimation of coverage

All reads were trimmed according to FastQC results using the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit). Alignments against sequences of interest were performed using Bowtie2 (v2.2.3)

(Langmead and Salzberg, 2012; Langmead et al., 2009), and the results were converted into the Binary Alignment/Map (BAM) file format using the SAMtools (v0.1.19) software package (Li and Durbin, 2009). Mapping quality and coverage statistics were assessed using QualiMap (v2.2) (Garcia-Alcalde et al., 2012; Okonechnikov et al., 2016).

2.5. Genome verification

To allow verification by PCR and Sanger sequencing, the web-based Primer3 software was used to designed primers according to the obtained consensus sequence. For primer sequences see Supplementary Table S1. PCRs were made using KAPA2G Robust HotStart PCR Kit (Kapa Biosystems, Boston, MA, USA). The PCR products were run on a 1.5% agarose gel stained with GelRed and visualized under UV transillumination (GelDoc, Bio-Rad Laboratories, Inc., Richmond, CA, USA). PCR products were purified and Sanger sequenced commercially (Macrogen Inc. Europe, Amsterdam, NL).

2.6. Genome annotation

As adenovirus genomes contain splice sites, several pipelines were used for the gene prediction and annotation. PROKKA (Seemann, 2014), an annotation pipeline for prokaryotic and viral genomes, was run using a set of Adenoviridae protein sequences for the functional inference (SwissProt release 2015_09). MAKER (Cantarel et al., 2008), more suited for annotating eukaryotic genomes, was run using two protein training sets: Adenoviridae sequences extracted from SwissProt and from TrEMBL (release 2015_09). In addition, gene prediction was also performed using GeneMark.hmm with heuristics models online (Besemer and Borodovsky, 1999). The previously annotated genomes of BdAdV-2 (KR024710) and California sea lion adenovirus 1 strain Zc11-030 (KJ563221) were used as a comparison to manually curate the annotation.

2.7. GC-content and recombination analysis

The ENDMEMO online web tool (www.endmemo.com/bio/gc.php) was used to calculate the GC-content for all genes in BdAdV-1. For spliced genes, the sequence corresponding to the spliced mRNA was used. A representative subset of ten complete or near-complete adenovirus genomes (KR024710.1, KY352473.1, NC_002513.1, NC_030860.1, NC_030874.1, NC_024150.1, NC_001734.1, NC_027708.1, NC_015932.1, NC_002702.1) was recovered for recombination analysis with BdAdV-1 based on homology search using BLASTn (Brudno et al., 2003) against the NCBI nr-database. The sequences were aligned using the ClustalW in CodonCode Aligner (v5.1.5) (CodonCode Corporation) prior to analysis with the RDP4 (v4.84) software program (Martin et al., 2015).

2.8. Comparative genomics and phylogenetic analysis

An online sequence alignment program, mVISTA Limited Area Global Alignment of Nucleotides (LAGAN) (<http://genome.lbl.gov/vista/index.shtml>) was used for global pairwise sequence alignment to compare BdAdV-1 with eight closely related adenovirus species, as revealed by BLASTn (Brudno et al., 2003).

Phylogenetic relationships were assessed for three predicted proteins from the core region (DNA polymerase, hexon and 100 K), both individually and concatenated. Protein alignments were performed using the MAFFT software (Katoh and Toh, 2008) and phylogenetic relationships were inferred using both Bayesian inference and maximum-likelihood methods. The Bayesian inference analysis was conducted using the MrBayes program (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) from within the TOPALi (v2.5) software suite, and maximum-likelihood analysis in the mega 6 software package (Tamura et al., 2013). Best-fit models of molecular evolution were

chosen using the selection tools in TOPALi (v2.5) and Mega 6. The Dendroscope software (v3.5.8) (Huson et al., 2007) and Adobe Illustrator CS (Adobe Systems, San Jose, CA, USA) were used for tree visualization.

2.9. Viral metagenomics data analysis

The cleaned reads were mapped against the bottlenose dolphin (*Tursiops truncatus*) genome available in Ensembl (in October 2015). Due to known problems with between sample contamination within a MiSeq run (Wright and Vetsigian, 2016) mapping towards Equid herpesvirus 5 (GQ325593.1), and Influenza A virus (KM366383.1, KM366384.1, KM366386.1, KM366387.1, KM366388.1, KM366389.1, KM366390.1, GQ495134.1) was also performed in order to remove contaminations from amplicons sequenced in the same run. The reads that did not match were analysed using the KAIJU software for taxonomic classification (Menzel et al., 2016).

3. Results

3.1. Processing of samples

After homogenization, centrifugation, filtration and nuclease treatment the DNA concentration was below the detection limit (1 ng/μl) of the Qubit 2.0 Fluorometer. The sample did however generate a sequencing library and good quality sequence reads.

3.2. Sequence data and assembly

High-throughput sequencing of the sample produced 515,234 paired-end reads with a mean length of 265 bp representing a total of 136.44 Mbp of sequence data. The initial MIRA assembly produced 1531 contigs and three singletons. CodonCode used 1529 of these contigs to produce a 34,120 bp long sequence. Manual inspection and trimming were performed before the sequence was used as the reference genome for mapping of reads and estimation of coverage. According to the results, 99.49% of all the reads mapped to the genome and were evenly distributed along the full length (Supplementary Fig. S1). This yielded an average coverage depth of 3674 (standard deviation 645) with a mean mapping quality of 43.55. However, since the sequence was very divergent to other adenovirus genomes publicly available at the time, the final sequence was also verified using Sanger sequencing.

3.3. The genome

The complete genome of BdAdV-1 (ENA: LT841149.1) is 34,080 bp and has ITR regions of 220 bp at each end of the genome (Fig. 1). BdAdV-1 has a GC-content of 36.2%. The core region of BdAdV-1 is similar to other mastadenoviruses. A total of 29 coding sequences were identified, out of which 26 were functionally annotated. Three proteins were assigned as hypothetical. Table 1 shows an overview of all genes, and their closest homologue according to BLASTp (9th of May 2017), identity percentage, and query coverage.

A global pairwise whole-genome sequence comparison between BdAdV-1 and eight closely related adenovirus species, as revealed by BLASTn, showed overall high sequence divergence (Fig. 2), especially in the 3' end including the E3 ORF1, U exon, fiber, E4 34K, E4 ORF1-2 and dUTPase. At the 5' end of the BdAdV-1 genome there is some sequence homology with BdAdV-2 but limited homology with other adenoviruses. The region in close proximity to the genes of pVII and pV display rather high degree of sequence divergence in comparison to other parts in the more conserved central part of the genome.

The BdAdV-1 contains a highly divergent 4380 bp E3 ORF1. The only similarity found was with the fiber of BdAdV-2 (12% query coverage, E-value of 9e-06, identity 29%). The fiber of BdAdV-2 does not have any

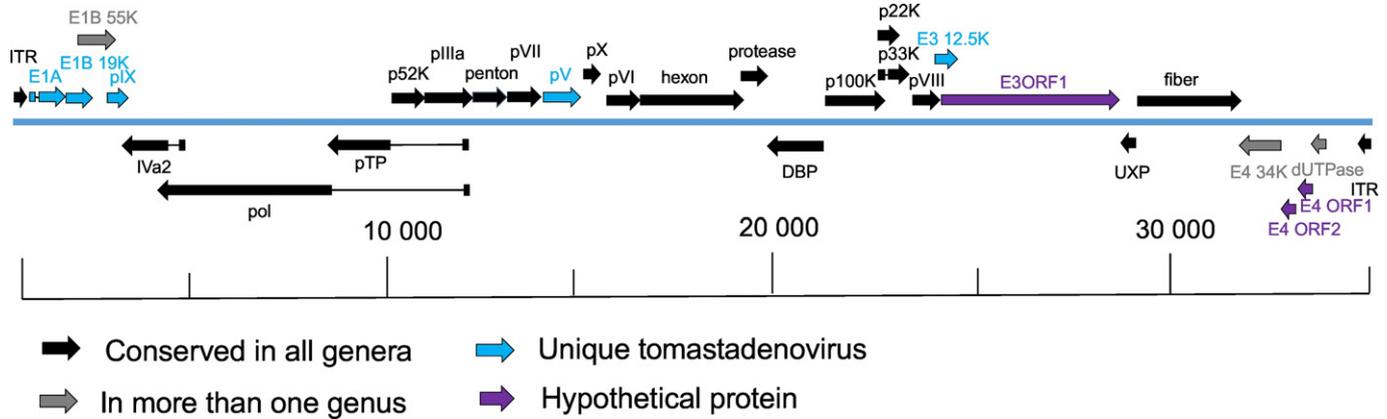


Fig. 1. Bottlenose dolphin adenovirus 1 (BdAdV-1) genome organization. Predicted ORFs are assigned with arrows, exons of spliced genes are connected by lines. Black arrows indicate genes that are conserved in all adenoviruses, grey arrows indicate genes that are present in more than one genus. Turquoise arrows indicate genes that are unique to the genus *Mastadenovirus* and purple arrows indicate genes coding for hypothetical proteins.

similarities to other fibers of adenoviruses therefore we speculate that the sequence/annotation of the 3' end of this genome might be misleading. Another possibility is the unlikely presence of two fiber ORFs in BdAdV-1, something that has been described in human adenovirus 40, human adenovirus 41 and simian adenovirus 1 (Kidd et al., 1993; Kovacs et al., 2005; Pieniasek et al., 1990). In comparison to E3 regions of other mastadenoviruses, which usually contain several smaller ORFs, the complete 1459 amino acids long E3 ORF1 is unusual. In length, it is similar to recently described E3 genes in Bat adenoviruses WIV9–11 (Tan et al., 2016). On a structural level, there appears to be several

transmembrane domains according to predictions made using Interproscan (Fig. 3). The predicted protein E3 12.5 K is located within the 5' end of the E3 region.

The nucleotide sequence of the fiber gene of the BdAdV-1 is homologous to the respective genomic sequences of mastadenoviruses. The C-terminal domain of the fiber matches the knob domain (pfam00541) at position 580–753 with an E-value of 1.91 e-09, which was also confirmed by an Interproscan analysis. The BdAdV-1 fiber is highly divergent with the closest homologue on protein level being bat mastadenovirus WIV13 (99% query coverage, E-value of 9 e-28, identity 26%).

Table 1

Coding sequences of the bottlenose dolphin adenovirus (BdAdV-1), their closest homologue according to BLASTp, identity percentage, and query coverage.

| Gene | Location (bp) | Strand | Length (aa) | Description | Closest homologue | GenBank accession no. | Identity (%) | Query cover (%) |
|-----------|---------------|--------|-------------|--|--|-----------------------|--------------|-----------------|
| E1A | 414–1029 | + | 177 | Early E1A protein | BdAdV-2 | ALE15290.1 | 36 | 87 |
| E1B_19K | 1130–1681 | + | 183 | E1B 19 K protein | BdAdV-2 | ALE15291.1 | 48 | 86 |
| E1B_55K | 1498–2817 | + | 439 | E1B protein, large T-antigen homologue | BdAdV-2 | ALE15292.1 | 54 | 89 |
| pIX | 2876–3292 | + | 138 | Hexon-interlacing protein | HpAdV-1 | AQQ73623.1 | 46 | 78 |
| Iva2 | 3302–4917 | – | 443 | Iva2 multifunctional protein | BdAdV-2 | ALE15294.1 | 61 | 100 |
| pol | 4390–12,208 | – | 1130 | DNA-dependent DNA polymerase | HpAdV-1 | AQQ73625.1 | 64 | 99 |
| pTP | 7617–12,208 | – | 616 | DNA terminal protein | HpAdV-1 | AQQ73626.1 | 71 | 99 |
| p52K | 9504–10,487 | + | 327 | Packaging protein 3 | BdAdV-2 | ALE15297.1 | 71 | 96 |
| pIIIa | 10,509–12,170 | + | 553 | Pre-capsid vertex protein | BdAdV-2 | ALE15298.1 | 64 | 99 |
| penton | 12,227–13,711 | + | 494 | Penton protein | HpAdV-1 | AQQ73629.1 | 72 | 100 |
| pVII | 13,716–14,366 | + | 216 | Pre-histone-like nucleoprotein | BdAdV-2 | ALE15300.1 | 62 | 100 |
| pV | 14,439–15,422 | + | 327 | Core-capsid bridging protein | HpAdV-1 | AQQ73631.1 | 58 | 100 |
| pX | 15,413–15,664 | + | 83 | Late L2 mu core protein | Equine adenovirus 1 | ANG08561.1 | 61 | 80 |
| pVI | 15,733–16,467 | + | 244 | Pre-protein VI | BdAdV-2 | ALE15303.1 | 54 | 100 |
| hexon | 16,503–19,334 | + | 943 | Hexon protein | HpAdV-1 | AQQ73634.1 | 80 | 100 |
| protease | 19,336–19,974 | + | 212 | Cysteine endopeptidase (adenain) | Tree shrew adenovirus 1 | YP_068070.1 | 61 | 95 |
| DBP | 19,971–21,278 | – | 435 | Early E2A DNA-binding protein | BdAdV-2 | ALE15306.1 | 72 | 99 |
| p100K | 21,292–23,406 | + | 704 | Shutoff protein | HpAdV-1 | AQQ73637.1 | 64 | 100 |
| p22K | 23,243–23,611 | + | 122 | Packaging protein 2 | HpAdV-1 | AQQ73639.1 | 53 | 91 |
| p33K | 23,243–23,840 | + | 160 | Alternative RNA splicing factor | HpAdV-1 | AQQ73638.1 | 50 | 93 |
| pVIII | 23,859–24,524 | + | 221 | Pre-hexon-linking protein | BdAdV-2 | ALE15310.1 | 66 | 100 |
| E3 12.5 K | 24,517–24,885 | + | 122 | Probable immunomodulatory protein | BdAdV-2 | ALE15311.1 | 60 | 97 |
| E3 ORF1 | 24,581–28,960 | + | 1459 | hypothetical protein | BdAdV-2 (fiber) ^a | ALE15313.1 | 29 | 12 |
| U exon | 29,105–29,272 | – | 55 | U exon protein | Bovine mastadenovirus B | AP_000040.1 | 44 | 100 |
| fiber | 29,282–31,573 | + | 763 | Fiber protein | Bat mastadenovirus WIV13 | YP_009272940 | 26 | 99 |
| E4 34 K | 31,613–32,410 | – | 265 | Early E4 34 kDa protein | Bovine mastadenovirus A (E4 ORFD) | YP_094053.1 | 38 | 88 |
| E4 ORF2 | 32,394–32,789 | – | 131 | hypothetical protein | <i>Sanguangporus baumii</i> (hypothetical protein) | OCB86471.1 | 30 | 54 |
| E4 ORF1 | 32,873–33,253 | – | 126 | hypothetical protein | BdAdV-2 (E4 34 K) ^b | ALE15315.1 | 54 | 88 |
| dUTPase | 33,253–33,624 | – | 123 | dUTPase | BdAdV-2 | ALE15316.1 | 43 | 99 |

^a Except for this hit, the E3 ORF1 does not match anything by BLASTp. Our interpretation is that the fiber in BdAdV-1 show clear characteristics of an adenovirus fiber and we find it more likely that this is an E3 ORF than a second fiber. We therefore annotated it as E3 ORF1.

^b The E4 ORF1 in BdAdV-1 is homologous to the E4 34 K of the BdAdV-2 but as there already is a, by amino acid sequence homology, well supported E4 34 K in the BdAdV-1 genome this ORF was annotated as E4 ORF1.

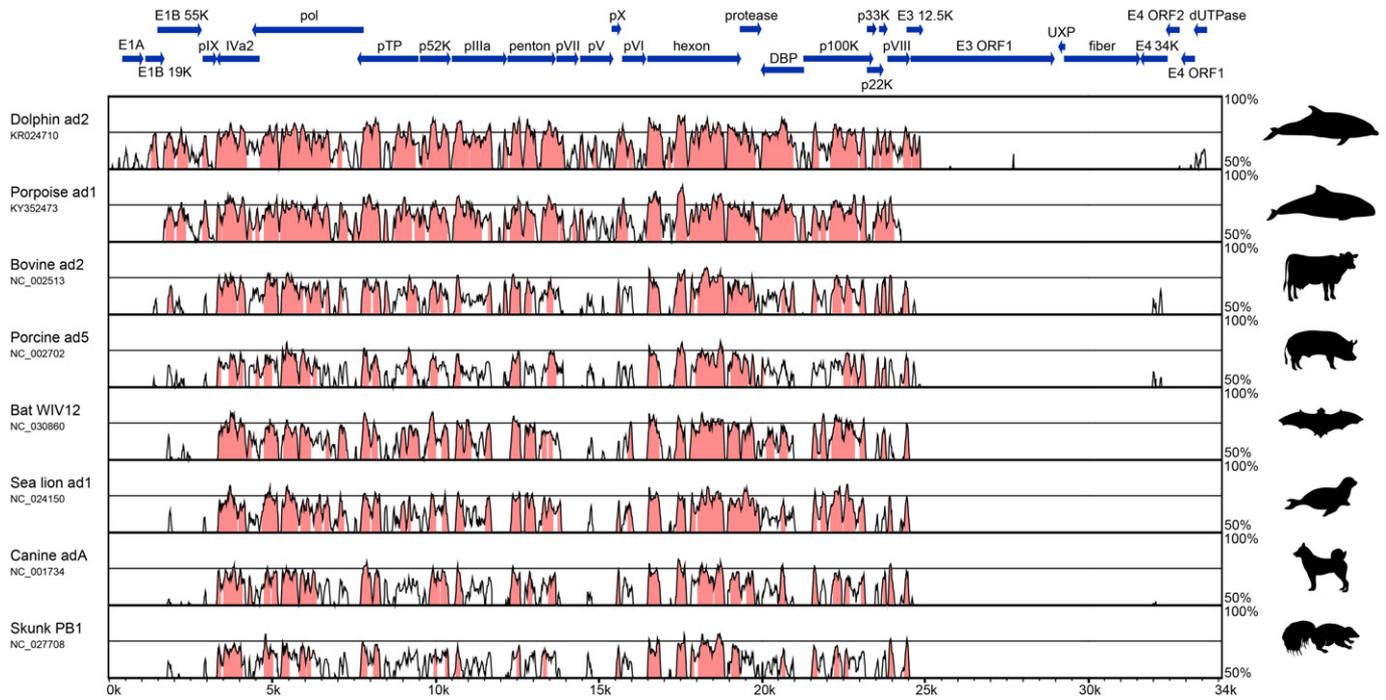


Fig. 2. Global pairwise whole-genome sequence comparison between BdAdV-1 and eight closely related adenovirus species, as revealed by BLASTn, using the online mVISTA Limited Area Global Alignment of Nucleotides (LAGAN) tool. The BdAdV-2 and HpAV-1 are only partial genomes with the 3' end missing. The height along the y-axis reflects the percentage of sequence conservation. All alignments display a general low level of sequence similarity. The regions towards the 3' end, including the E3 ORF1, U exon, fiber, E4 34K, dUTPase and E4 ORF1-2, seem to be highly diverged from all other known adenoviruses.

The E4 ORF2, located in the 3' end, has only one homologous protein in the gene bank database, a hypothetical protein in *Sanguangporus baumii*. The E4 ORF1, also located at the 3' end, matched with what has been annotated as the E4 34 K of BdAdV-2. However, as the E4 34 K in BdAdV-1 is homologous to many other E4 34K we choose to annotate the E4 ORF1 as a hypothetical protein. The only other BLAST hit for E4 ORF1 is a inaD-like protein (*Castor canadensis*) (30% query coverage, E-value of 4.3, identity 43%).

The BdAdV-1 dUTPase matches the dUTPase superfamily domain (pfam00692) at position 26–110 with an E-value of 7.72 e-04. The closest homologue is the dUTPase2 of BdAdV-2, followed by dUTPase in “*Candidatus Saccharibacteria bacterium*” GW2011_GWA2_46_10 (KKU18896.1), indicating a dUTPase activity.

The U exon of BdAdV-1 is most closely related to *Bovine mastadenovirus B*.

Virus associated RNAs (VA-RNAs) are non-protein coding sequences, which can antagonize the antiviral activity of interferons (Mathews and Shenk, 1991). There is to date no evidence of VA-RNA in non-primate mastadenoviruses (Venkatesh et al., 1998). In line with previous research the BdAdV-1 appears to lack VA-RNAs.

3.4. Splicing sites in the BdAdV-1 genome

The genome of BdAdV-1 has splicing sites in several locations found in previous adenovirus studies (Davison et al., 2003). To confirm if these sites are the correct ones, further experimental evidence is needed.

Unintegrated signatures



Fig. 3. A schematic representation of domains in the E3 ORF1 as predicted by an Interproscan analysis. Left to the right in the picture corresponds to the 5' to 3' of the E3 ORF1.

3.5. Phylogenetic analysis

3.5.1. Recombination

Average GC-content per gene across the BdAdV-1 genome varied from 29.8–51.5% (Fig. 4) with the lowest figures seen in the 3' end of the genome. The three genes used for phylogenetic analysis had GC contents that were close to the average of the BdAdV-1 genome, i.e. 36.2%, indicating that they have not been subjected to recombination recently. In addition, recombination analysis of complete and near-complete adenovirus genomes revealed no significant recombination events within the BdAdV-1 genome.

3.5.2. Homology search

A homology search in BLASTn (9th of May 2017) indicated that the most similar genome is the partial genome of BdAdV-2 (KR024710) (60% query coverage, E-value of 0, identity 71%), followed by the partial genome of harbour porpoise adenovirus 1 (KY352473) (56% query coverage, E-value of 0, identity 69%), the complete genome of bat mastadenovirus WIV12 (KT698856) (44% query coverage, E-value of 0, identity 67%), and the bovine adenovirus type 2 (AF252854) (38% query coverage, E-value of 0, identity 68%) (Salmon et al., 1993).

3.5.3. Phylogenetic analysis on protein level

Individual phylogenetic analyses of the protein sequences for the three genes DNA polymerase, hexon and 100 K, using both maximum-likelihood and Bayesian inference methods with best suggested

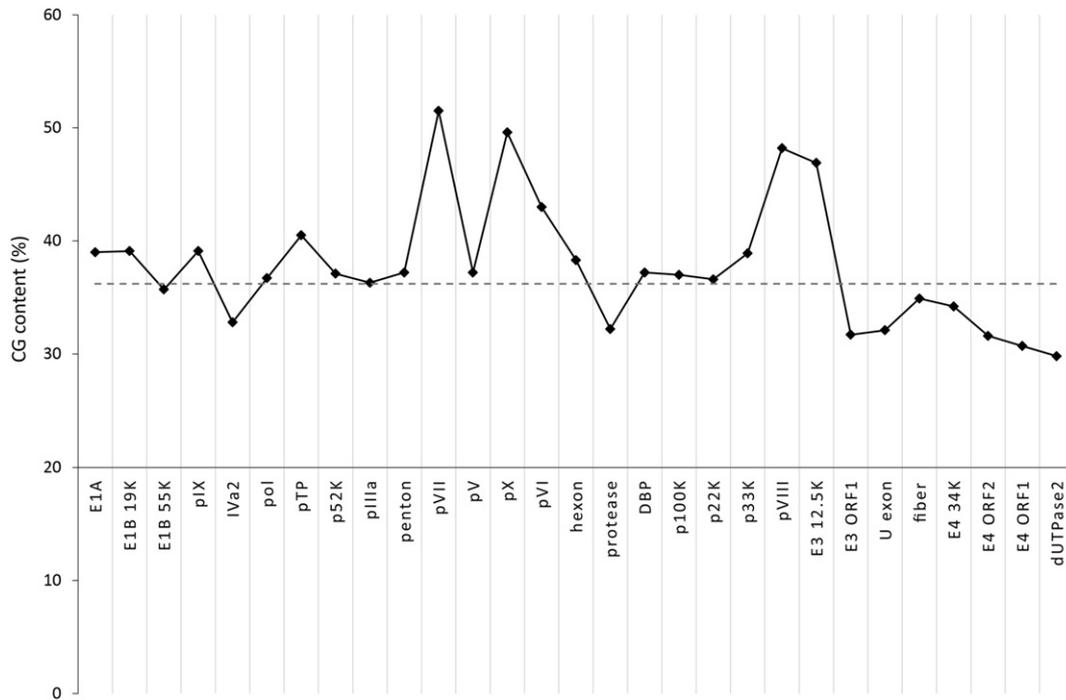


Fig. 4. The GC-content of the different genes of the BdAdV-1 genome. The average GC-content is displayed as a dotted line.

evolution models, showed similar evolutionary patterns (Supplementary Fig. S2). A concatenated analysis was therefore performed (Fig. 5). Detailed information about the 69 adenoviruses included in the phylogenetic analysis can be found in Supplementary Table S2. The best-fit substitution model used for the majority of the maximum-likelihood analyses was LG + G + I. The only exception was the hexon dataset for which the LG + G + I + F amino acid substitution model was deemed superior. The model used for all the Bayesian inference analyses was WAG + I + G as suggested by the model selection feature within TOPALI.

Both the maximum-likelihood and Bayesian analyses grouped BdAdV-1 within the *Mastadenovirus* genus, and more specifically

among mastadenoviruses infecting *Cetartiodactyla* (even-toed ungulate). The most closely related virus was the BdAdV-2, followed by HpAdV-1, and then porcine and bovine adenoviruses. The location of BdAdV-1 among mastadenoviruses infecting *Cetartiodactyla* indicates co-evolution with its host.

Individual analyses of the DNA polymerase, hexon and 100 K, on an amino acid level, indicate 64–80% identity (99–100% query coverage) to the HpAdV-1 (Table 1), and 62–79% identity (97–99% query coverage) to the BdAdV-2. The fiber of BdAdV-1 was shown to be highly divergent from other fibers, with the closest fiber having 26% identity (in Bat mastadenovirus), phylogenetic analysis of the fiber could therefore not be conducted.

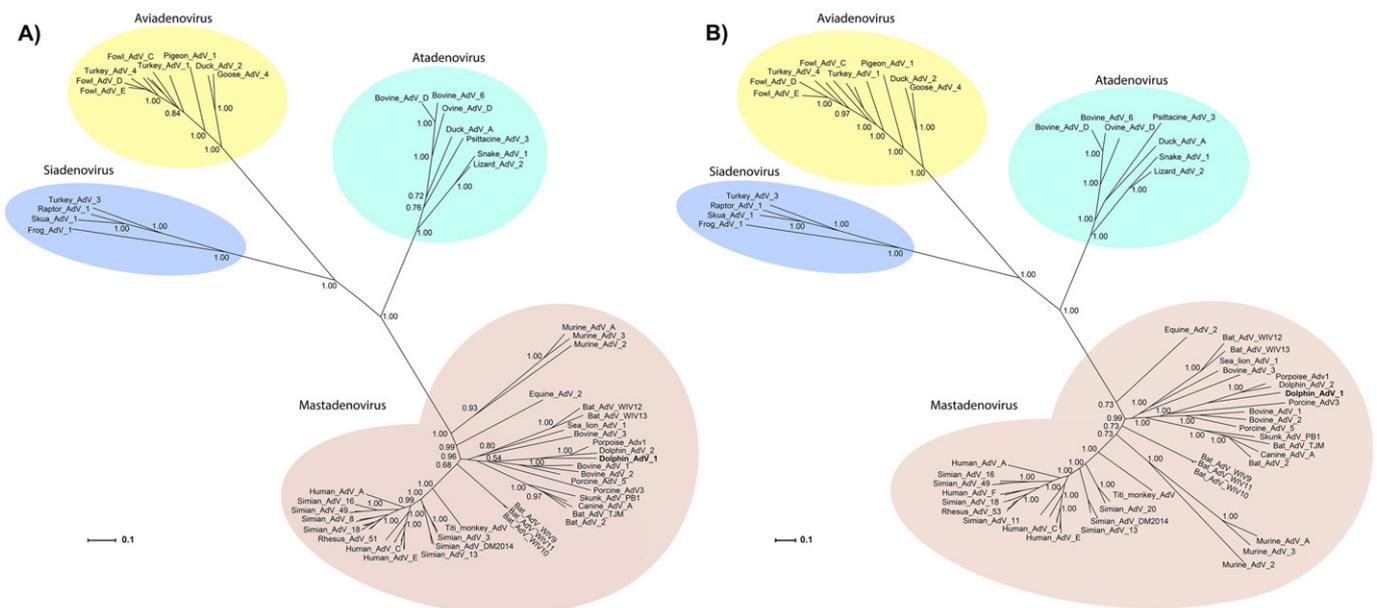


Fig. 5. A) A Maximum-likelihood tree of representatives from each of the genera (except Ichtadenovirus) in the family *Adenoviridae*, based on the deduced amino acid sequences for concatenated DNA polymerase, hexon, and 100 K. B) Bayesian analysis phylogram depicting the relationship of the BdAdV-1 to representatives from each of the genera (except Ichtadenovirus) in the family *Adenoviridae*, based on the deduced amino acid sequences for concatenated DNA polymerase, hexon and 100 K.

3.6. Viral metagenomics

In total 0.09% of the cleaned reads mapped against the bottlenose dolphin genome. The majority of the remaining 99.91% of the reads were classified as adenovirus, in particular as mastadenovirus. The full table of the classified reads is available in Supplementary Table S3. The viral metagenomics analysis shows no indication of DNA viruses other than the adenovirus present in the studied faecal sample. The raw sequence data from the MiSeq is deposited in ENA with accession number ERX1999096.

4. Discussion

We describe the complete genomic sequence of an adenovirus identified during an investigation of an outbreak of gastroenteritis among bottlenose dolphins living under human care (Rubio-Guerri et al., 2015). According to our results, the identified bottlenose dolphin adenovirus 1 (BdAdV-1) is distinctly different from other adenoviruses and we therefore propose that it should be considered as a novel species in the *Mastadenovirus* genus. Similarity searches against databases of publicly available sequences revealed that the most closely related virus is an undescribed adenovirus found in the faeces of a bottlenose dolphin from the United States (GenBank Acc. no. KR024710), here referred to as BdAdV-2. Despite the partial sequence of BdAdV-2, large parts of the genomes could still be compared. Amino-acid-level comparisons of individual genes revealed a sequence identity of maximum 79%, suggesting that there are profound differences even though the two adenoviruses were detected in faecal samples from the same host animal. Interestingly, nine proteins in the HpAdV-1, which was detected in harbour porpoise stranded in the Netherlands, were given a higher score in BLASTp than BdAdV-2 when compared to BdAdV-1 (Table 1). There is a major lack of adenovirus sequences from cetaceans in general, and delphinids in particular. Thus, by revealing the complete genetic characteristics of a dolphin-associated adenovirus, this work contributes to the growing body of knowledge concerning the molecular biology of adenoviruses. Furthermore, the genome has been carefully annotated and curated by combining predictions made by annotation tools designed for eukaryotes and prokaryotes, as well as tools for protein domain prediction. This has allowed us to investigate important aspects such as genome structure, function, and evolution.

BdAdV-1 has a linear, double-stranded DNA genome of 34,080 bp in length, flanked on each side by 220 bp ITRs. A total of 29 coding sequences were identified, of which 26 were functionally annotated. Two short hypothetical proteins E4 ORF1 and E4 ORF2, as well as a longer hypothetical protein named E3 ORF1 were identified. This unusually long E3 ORF1 is an interesting finding in the BdAdV-1 genome. E3 regions are present in all mastadenoviruses and contains from one to nine genes (Davison et al., 2003). The E3 region is involved in immune modulating activities and is known to display genetic variation (Ginsberg et al., 1989). Genes located in the E3 region encode membrane proteins which may have conserved regions involved in manipulating the immune response (Deryckere and Burgert, 1996; Hofmayer et al., 2009). An extremely large E3 gene, 1482–1648 amino acid long, was recently described to encode a mRNA transcribed 4–36-hour post-infection in Bat adenoviruses WIV 9–11 (Tan et al., 2016). So, although the 1458 amino acid long E3 ORF1 of BdAdV-1 is not homologous to any known E3 proteins with respect to protein sequence, we hypothesize that the E3 protein is expressed and that its role is related to immune evasion.

Based on phylogenetic analysis BdAdV-1 clearly belongs to the *Mastadenovirus* genus in the *Adenoviridae* family. Besides BdAdV-2, the most genetically similar sequences are HpAdV-1, bovine and porcine adenoviruses. This corresponds well with the current view of mammalian evolution, and is further supported by a recent phylogenetic study of eutherian mammalian genomes (Foote et al., 2015), in which dolphins and whales were more closely related to bovines and porcine

while being more distantly related to other aquatic mammals such as sea lion, walrus and manatee. This strongly suggests that BdAdV-1 is a dolphin-, or possibly cetacean-specific virus that has been co-evolving with its host. In agreement with this speculation is the finding that the fiber of BdAdV-1 displays clear characteristics of an adenoviral fiber, including the knob domain, but is highly divergent by protein sequence. The closest homologue, the bat mastadenovirus, has only 26% identity on a protein level. The exposed location of the fiber protein and its involvement in invasion could in part explain its increased protein sequence variation. However, further studies are needed in order to truly be able to evaluate species and tissue tropism. Based on the gastroenteritis symptoms, and similarity to other adenoviruses, it is possible to assume that the cells infected are located in the gastrointestinal tract.

The presence of dUTPase in viruses has primarily been explained by horizontal gene transfer (Baldo and McClure, 1999). dUTPase hydrolyses dUTP into dUMP, a substrate for synthesis of dTTP. In this way, the dUTP:dTTP ratio are kept low which prevents misincorporation of uracil into DNA. dUTPase play an important role in some viruses in virus replication in non-dividing cells (Chen et al., 2002). dUTPase has also been associated with an oncogenic potential (Baldo and McClure, 1999). The dUTPase in BdAdV-1 is most closely related to the dUTPase2 in BdAdV-2, with the main difference being that there is no evidence for an additional dUTPase in BdAdV-1. The second closest homologue is a dUTPase in "*Candidatus Saccharibacteria bacterium*", which is indicative of a horizontal gene transfer event.

The GC-content of BdAdV-1 is very low (36.2%) in comparison with other mastadenoviruses, such as bovine adenovirus (43.6–54%), human adenoviruses (46.5–55.3%), and porcine adenovirus (50.5–63.6%). The only other two currently genetically characterized mastadenoviruses with as low GC-content are California sea lion adenovirus (36%) and BdAdV-2 (36.4%). Lower GC-content and more acute pathogenicity has been associated with recent host jumps (Benko and Harrach, 2003; Poss et al., 2006; van Hemert et al., 2007), but further research is needed to know if this is the case for BdAdV-1. The GC-content in BdAdV-1 is lowest in the 3' genes dUTPase, E4 ORF1–2, U exon and E3 ORF1. This could be indicative of greater genetic flexibility in this region, as is generally seen in other adenoviruses.

Traditional methods for sequencing of whole genomes, such as primer walking, can be particularly challenging when working with genomes that have very low similarity with previously identified organisms. This study demonstrates the usefulness of high-throughput sequencing to obtain full-length genomes of genetically divergent viruses. An additional benefit with applying a viral metagenomics approach is the ability to exclude, in this case, the involvement of other DNA viruses in the etiology of the disease.

Modified forms of adenoviruses are being developed to treat genetic diseases, to fight cancer, and to deliver vaccines. However, many human based adenoviruses used for therapy cause problems such as toxicity due to their liver tropism (Raper et al., 2003). To our knowledge BdAdV-1 does not seem to display any liver tropism, which might make it an interesting candidate for evaluation of some of the above listed applications.

Future research should focus on investigating expression of BdAdV-1 mRNA/proteins and determination of tissue tropism, in order to clarify important questions regarding the infection biology of this newly detected adenovirus.

5. Conclusion

The complete sequence of a recently identified adenovirus associated with gastroenteritis in dolphins is described. This novel adenovirus, termed BdAdV-1, has a viral genome of 34,080 bp. The present study revealed that BdAdV-1 belongs to the *Mastadenovirus* genus in the *Adenoviridae* family, however, it is clearly different from previously characterized adenoviruses, especially in the 3' end of the viral genome. The BdAdV-1 genome is most similar to the complete genome of

BdAdV-2 (KR024710). The study also demonstrates the usefulness of high-throughput sequencing to obtain full-length genomes of genetically divergent viruses.

Supplementary data associated with this article can be found in the online version. Fig. S1 presents the coverage of reads across the BdAdV-1 genome. Fig. S2 displays individual phylogenetic trees. Table S1 contains primer sequences. Table S2 contains information on all the adenoviruses used for the phylogenetic analysis. Table S3 contains the classified reads. Supplementary data associated with this article can be found in the online version, at doi: [10.1016/j.meegid.2017.05.008](https://doi.org/10.1016/j.meegid.2017.05.008).

Competing interests

The authors have declared that no competing interests exist.

Author contributions

M.Ma, J.H, F.G: design of the study. D.G-P, T.Á, M.V: animal handling and sampling of animals. C.R-G, E.N-P, M.Me, J.M.S.-V: PCR diagnosis. M.Ma: sample processing. M.Ma, J.H, F.G: bioinformatics. M.Ma, C.R-G, J.H, F.G, S-B: data analysis and interpretation. M.Ma drafted the manuscript. All authors revised the manuscript critically and approved the final manuscript.

Acknowledgements

This work was supported by Epi-SEQ: a research project supported under the 2nd joint call for transnational research projects by EMIDA ERA-NET (FP7 project nr 219235). Financial support was also obtained from The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, Formas (grant number 221-2012-586). Karin Ullman at the National Veterinary Institute in Sweden provided valuable assistance with the MiSeq sequencing. We acknowledge Dr. Ulrika Morris at Karolinska Institutet for language revision. This work has been carried out under the auspices of a collaborative agreement on virology studies in marine mammals between The Oceanográfico de la Ciudad de las Artes y las Ciencias de Valencia and the VISAVET Center of Complutense University of Madrid. We acknowledge Belén Rivera and Rocío Sánchez for technical assistance. C.R-G is the recipient of a predoctoral fellowship from the FPU programme of the Spanish Ministry of Education. M.Me. is the recipient of a predoctoral fellowship from the PhD student grant programme of Complutense University of Madrid.

References

Aldasy, P., Csontos, L., Bartha, A., 1965. Pneumo-enteritis in calves caused by adenoviruses. *Acta Vet. Acad. Sci. Hung.* 15, 167–175.

Alonso-Padilla, J., Papp, T., Kajan, G.L., Benko, M., Havenga, M., Lemckert, A., Harrach, B., Baker, A.H., 2016. Development of novel adenoviral vectors to overcome challenges observed with HAAdV-5-based constructs. *Mol. Ther.* 24, 6–16.

Baldo, A.M., McClure, M.A., 1999. Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts. *J. Virol.* 73, 7710–7721.

Belak, S., Palfi, V., 1974. Pneumoenteritis of lambs caused by adenovirus. Preliminary report. *Acta Vet. Acad. Sci. Hung.* 24, 327–328.

Benko, M., Harrach, B., 2003. Molecular evolution of adenoviruses. *Curr. Top. Microbiol. Immunol.* 272, 3–35.

Besemer, J., Borodovsky, M., 1999. Heuristic approach to deriving models for gene finding. *Nucleic Acids Res.* 27, 3911–3920.

Britt Jr., J.O., Nagy, A.Z., Howard, E.B., 1979. Acute viral hepatitis in California sea lions. *J. Am. Vet. Med. Assoc.* 175, 921–923.

Brudno, M., Do, C.B., Cooper, G.M., Kim, M.F., Davydov, E., Program, N.C.S., Green, E.D., Sidow, A., Batzoglou, S., 2003. LAGAN and multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* 13, 721–731.

Cabasso, V.J., Stebbins, M.R., Norton, T.W., Cox, H.R., 1954. Propagation of infectious canine hepatitis virus in tissue culture. *Proc. Soc. Exp. Biol. Med.* 85, 239–245.

Cantarel, B.L., Korf, I., Robb, S.M., Parra, G., Ross, E., Moore, B., Holt, C., Sanchez Alvarado, A., Yandell, M., 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18, 188–196.

Chen, R., Wang, H., Mansky, L.M., 2002. Roles of uracil-DNA glycosylase and dUTPase in virus replication. *J. Gen. Virol.* 83, 2339–2345.

Chevreaux, B., Wetter, T., Suhai, S., 1999. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information German Conference on Bioinformatics. pp. 45–56.

Davison, A.J., Benko, M., Harrach, B., 2003. Genetic content and evolution of adenoviruses. *J. Gen. Virol.* 84, 2895–2908.

De Guise, S., Lagace, A., Beland, P., Girard, C., Higgins, R., 1995. Non-neoplastic lesions in beluga whales (*Delphinapterus leucas*) and other marine mammals from the St Lawrence Estuary. *J. Comp. Pathol.* 112, 257–271.

Deryckere, F., Burgert, H.G., 1996. Early region 3 of adenovirus type 19 (subgroup D) encodes an HLA-binding protein distinct from that of subgroups B and C. *J. Virol.* 70, 2832–2841.

Dierauf, L.A., Lowenstine, L.J., Jerome, C., 1981. Viral hepatitis (adenovirus) in a California sea lion. *J. Am. Vet. Med. Assoc.* 179, 1194–1197.

Footo, A.D., Liu, Y., Thomas, G.W., Vinar, T., Alfoldi, J., Deng, J., Dugan, S., van Elk, C.E., Hunter, M.E., Joshi, V., Khan, Z., Kovar, C., Lee, S.L., Lindblad-Toh, K., Mancia, A., Nielsen, R., Qin, X., Qu, J., Raney, B.J., Vijay, N., Wolf, J.B., Hahn, M.W., Muzny, D.M., Worley, K.C., Gilbert, M.T., Gibbs, R.A., 2015. Convergent evolution of the genomes of marine mammals. *Nat. Genet.* 47, 272–275.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Gotz, S., Tarazona, S., Dopazo, J., Meyer, T.F., Conesa, A., 2012. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28, 2678–2679.

Ginsberg, H.S., Lundholm-Beauchamp, U., Horswood, R.L., Pernis, B., Wold, W.S., Chanock, R.M., Prince, G.A., 1989. Role of early region 3 (E3) in pathogenesis of adenovirus disease. *Proc. Natl. Acad. Sci. U. S. A.* 86, 3823–3827.

Goldstein, T., Colegrove, K.M., Hanson, M., Gulland, F.M., 2011. Isolation of a novel adenovirus from California sea lions *Zalophus californianus*. *Dis. Aquat. Org.* 94, 243–248.

Hofmayer, S., Madisch, I., Darr, S., Rehren, F., Heim, A., 2009. Unique sequence features of the human adenovirus 31 complete genomic sequence are conserved in clinical isolates. *BMC Genomics* 10, 557.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755.

Huson, D.H., Richter, D.C., Rausch, C., DeZulian, T., Franz, M., Rupp, R., 2007. Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinform.* 8, 460.

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.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* 9, 286–298.

Kidd, A.H., Chroboczek, J., Cusack, S., Ruigrok, R.W., 1993. Adenovirus type 40 virions contain two distinct fibers. *Virology* 192, 73–84.

Knowles, D.P., 2011. Adenoviridae. *Fenner's Veterinary Virology*. fourth ed. Elsevier/Academic Press, San Francisco.

Kovacs, G.M., Harrach, B., Zakhartchouk, A.N., Davison, A.J., 2005. Complete genome sequence of simian adenovirus 1: an Old World monkey adenovirus with two fiber genes. *J. Gen. Virol.* 86, 1681–1686.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760.

Majhen, D., Calderon, H., Chandra, N., Fajardo, C.A., Rajan, A., Alemany, R., Custers, J., 2014. Adenovirus-based vaccines for fighting infectious diseases and cancer: progress in the field. *Hum. Gene Ther.* 25, 301–317.

Martin, D.P., Murrell, B., Golden, M., Khosal, A., Muhire, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 1, vev003.

Mathews, M.B., Shenk, T., 1991. Adenovirus virus-associated RNA and translation control. *J. Virol.* 65, 5657–5662.

Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7, 11257.

Okonechnikov, K., Conesa, A., García-Alcalde, F., 2016. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32, 292–294.

Pieniazek, N.J., Slemenda, S.B., Pieniazek, D., Velarde Jr., J., Luftig, R.B., 1990. Human enteric adenovirus type 41 (Tak) contains a second fiber protein gene. *Nucleic Acids Res.* 18, 1901.

Poss, M., Ross, H.A., Painter, S.L., Holley, D.C., Terwee, J.A., Vandewoude, S., Rodrigo, A., 2006. Feline lentivirus evolution in cross-species infection reveals extensive G-to-A mutation and selection on key residues in the viral polymerase. *J. Virol.* 80, 2728–2737.

Raper, S.E., Chirmule, N., Lee, F.S., Wivel, N.A., Bagg, A., Gao, G.P., Wilson, J.M., Batshaw, M.L., 2003. Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol. Genet. Metab.* 80, 148–158.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.

Rowe, W.P., Huebner, R.J., Gilmore, L.K., Parrott, R.H., Ward, T.G., 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc. Soc. Exp. Biol. Med.* 84, 570–573.

Rubio-Guerri, C., Garcia-Parraga, D., Nieto-Pelegrin, E., Melero, M., Alvaro, T., Valls, M., Crespo, J.L., Sanchez-Vizcaino, J.M., 2015. Novel adenovirus detected in captive bottlenose dolphins (*Tursiops truncatus*) suffering from self-limiting gastroenteritis. *BMC Vet. Res.* 11, 367.

Salmon, K., Esford, L., Haj-Ahmad, Y., 1993. Subcloning and restriction enzyme mapping of bovine adenovirus type 2. *Intervirology* 36, 72–78.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069.

- Smith, A.W., Skilling, D.E., 1979. Viruses and virus diseases of marine mammals. *J. Am. Vet. Med. Assoc.* 175, 918–920.
- Smith, A.W., Skilling, D.E., Benirschke, K., Albert, T.F., Barlough, J.E., 1987. Serology and virology of the bowhead whale (*Balaena mysticetus* L.). *J. Wildl. Dis.* 23, 92–98.
- 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.
- Tan, B., Yang, X.L., Ge, X.Y., Peng, C., Zhang, Y.Z., Zhang, L.B., Shi, Z.L., 2016. Novel bat adenoviruses with an extremely large E3 gene. *J. Gen. Virol.* 97, 1625–1635.
- van Beurden, S.J., LL, I.J., van de Bildt, M.W., Begeman, L., Wellehan Jr., J.F., Waltzek, T.B., de Vrieze, G., Grone, A., Kuiken, T., Verheije, M.H., Penzes, J.J., 2017. A novel cetacean adenovirus in stranded harbour porpoises from the North Sea: detection and molecular characterization. *Arch. Virol.* <http://dx.doi.org/10.1007/s00705-017-3310-8> (Epub ahead of print).
- Van Bresseem, M.F., Van Waerebeek, K., Raga, J.A., 1999. A review of virus infections of cetaceans and the potential impact of morbilliviruses, poxviruses and papillomaviruses on host population dynamics. *Dis. Aquat. Org.* 38, 53–65.
- van Hemert, F.J., Berkhout, B., Lukashov, V.V., 2007. Host-related nucleotide composition and codon usage as driving forces in the recent evolution of the Astroviridae. *Virology* 361, 447–454.
- Venktesh, A., Watt, F., Xu, Z.Z., Both, G.W., 1998. Ovine adenovirus (OAV287) lacks a virus-associated RNA gene. *J. Gen. Virol.* 79 (Pt 3), 509–516.
- Wold, W.S., Toth, K., 2013. Adenovirus vectors for gene therapy, vaccination and cancer gene therapy. *Curr. Gene Ther.* 13, 421–433.
- Wold, W.S.M.H.M.S., 2007. Adenoviruses. *Fields Virology*, fifth ed. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia (5th ed.).
- Wright, E.S., Vetsigian, K.H., 2016. Quality filtering of Illumina index reads mitigates sample cross-talk. *BMC Genomics* 17, 876.