Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Characterization of monkey adenoviruses with three fiber genes

Iva I. Podgorski [1], Balázs Harrach, Mária Benkő, Tibor Papp *

Veterinary Medical Research Institute, H-1143 Budapest, Hungary

ARTICLE INFO

Keywords:
Simian adenovirus
Genome
Fiber
Species
Old World monkey

ABSTRACT

Although the occurrence of three fiber genes in monkey adenoviruses had already been described, the relatedness of the "extra" fibers have not yet been discussed. Here we report the genome analysis of two simian adenovirus (SAdV) serotypes from Old World monkeys and the phylogenetic analysis of the multiple fiber genes found in these and related AdVs. One of the newly sequenced serotypes (SAdV-2), isolated from a rhesus macaque (*Macaca mulatta*), was classified into species *Human mastadenovirus G* (HAdV-G), while the other serotype (SAdV-17), originating from a grivet (*Chlorocebus aethiops*), classified to *Simian mastadenovirus F* (SAdV-F). We identified unique features in the gene content of these SAdVs compared to those typical for other members of the genus *Mastadenovirus*. Namely, in the E1B region of SAdV-2, the 19K gene was replaced by an ITR repetition and a copy of the E4 ORF1 gene. Among the 37 genes in both SAdVs, three genes of different lengths, predicted to code for the cellular attachment proteins (the fibers), were found. These proteins exhibit high diversity. Yet, phylogenetic calculations of their conserved parts could reveal the probable evolutionary steps leading to the multiple-fibered contemporary HAdV and SAdV species. Seemingly, there existed (a) common ancestor(s) with two fiber genes for the lineages of the AdVs in species SAdV-B, -E, -F and HAdV-F, alongside a double-fibered ancestor for today's SAdV-C and HAdV-G, which later diverged into descendants forming today's species. Additionally, some HAdV-G members picked up a third fiber gene either to the left-hand or to the in-between position from the existing two. A SAdV-F progenitor also obtained a third copy to the middle, as observed in SAdV-17. The existence of three fiber genes in these contemporary AdVs brings novel possibilities for the design of optimised AdV-based vectors with potential multiple target binding abilities.

1. Introduction

Adenoviruses (AdVs) of primates are classified into the genus *Mastadenovirus* within the family *Adenoviridae* (Benkő et al., 2022), and represent the best-studied AdVs nowadays (Harrach et al., 2019). The knowledge we have about the human [*Homo sapiens*] and most other great ape [Hominidae] (i.e. bonobo [*Pan paniscus*], chimpanzee [*Pan troglodytes*] and gorilla [*Gorilla gorilla* and *G. beringei*]) adenoviruses far surpasses that about the AdVs of the non-hominid primates (e.g. Bots et al., 2022). The traditional classification of non-hominid primates, as Old World monkeys (OWM) [Cercopithecoidea], New World monkeys (NWM) [Platyrrhini] and the paraphyletic group of the prosimians [Strepsirrhini and Tarsiiformes], was supported by genetic data as well (Perelman et al., 2011) and will be used accordingly in this paper. Some of the OWM AdVs were discovered >60 years ago (Hull et al., 1958) and numerous other further OWM, NWM and prosimian AdVs followed, most of them showing a co-evolutionary phylogenetic pattern to that of their hosts (Podgorski et al., 2018). So far, there are seven officially recognised human AdV (HAdV) species: *Human mastadenovirus A* to *G* (their names informally abbreviated as HAdV-A to HAdV-G). While for the OWM AdVs nine species were established with the currently accepted name of *Simian mastadenovirus A* to *I* (SAdV-A to SAdV-I), there is one official species for the NWM AdVs: *Platyrrhini mastadenovirus A* comprising isolates from a titi monkey [*Plecturocebus cupreus*] and a squirrel monkey [*Saimiri boliviensis*] (https://sites.google.com/site/adenoseq/) (Benkő et al., 2022).

Next to the human-derived AdVs and their designated species, majority of the so-far described great ape AdVs (unofficially also called as SAdV types) belong to species HAdV-A, -B, -C, -E or -F. On the other hand, whilst HAdV-D species is comprised of AdVs of human origin only (presumed coevolved AdV species of humans), HAdV-G species has one human-derived member and numerous OWM SAdV members. Interest in

---

* Corresponding author.

E-mail addresses: iskrinj@irb.hr (I.I. Podgorski), harrach.balazs@vmri.hu (B. Harrach), benko.maria@vmri.hu (M. Benkő), papp.tibor@vmri.hu, kabafalvi@gmail.com (T. Papp).

[1] Present address: Ruđer Bošković Institute, Zagreb, Croatia.

https://doi.org/10.1016/j.meegid.2023.105403
Received 14 November 2022; Received in revised form 1 January 2023; Accepted 3 January 2023

SAdVs as possible gene delivery vectors has started to increase recently, as SAdVs are largely similar to HAdVs, yet antigenically far enough to avoid the immunity pre-existing in the human population (Lopez-Gordo et al., 2014).

It is a controversial issue whether zoonotic or reverse zoonotic transmissions of SAdV and HAdV strains/types occur, and a recent review of 74 full papers provided numerous examples on both sides of the argument (Borkenhagen et al., 2019). A subsequent study of AdV infections in African humans and wild non-human primates (NHP) detected indication of cross-species transmission and great genetic diversity, based on limited partial gene sequence data (Medkour et al., 2020). This paper described the anthropozoosis of HAdV-C member from gorillas, and the widest diversity of the HAdV-E genotype variants among NHP species, suggesting lack of host-specificity in the latter virus species.

The icosahedral capsid of AdVs consists of 240 non-vertex capsomers (hexons), and 12 vertex capsomers (penton bases), each anchoring a cellular attachment protein, the fiber that protrudes from the virion's surface (Benkő et al., 2022). These three structural proteins: the hexon, the penton base, and the fiber are known to mediate the binding of the virion to the host cells (Persson et al., 2021; Stasiak and Stehle, 2020). The majority of mastadenoviruses possess only one fiber gene. However, several primate AdVs have been identified to contain more than one fiber gene. Two fiber genes have been recognised in HAdV-F and -G, SAdV-B, -C, and -E species within the genus *Mastadenovirus* (Alonso-Padilla et al., 2015; Chiu et al., 2013; Jones et al., 2007; Kidd et al., 1993; Podgorski et al., 2016; Roy et al., 2012, 2009; Zeng et al., 2016). The first SAdV described having two fiber genes was SAdV-1 (Kovács et al., 2005), which had been found phylogenetically closer to HAdVs with two fiber genes (in species HAdV-F) than to the previously sequenced OWM SAdV-3 of species SAdV-A (Kovács et al., 2004). Even closer kins of SAdV-1 were described later both from humans (HAdV-52) (Jones et al., 2007) as well as from simian hosts (Pantó et al., 2015; Podgorski et al., 2016; Roy et al., 2012, 2009) resulting in the establishment of the viral species HAdV-G. Moreover, among recently vectorised 17 novel rhesus AdVs (RhAdV), all belonging to this species, 10 RhAdV genomes possessed 3 fiber genes, a unique phenomenon so far in the genus (Abbink et al., 2015, 2018). Regarding other genera, two fiber genes have also been found in many representatives of different aviadenovirus species (Kaján et al., 2010; Marek et al., 2014a, 2014b; Zhao et al., 2015) and in three atadenoviruses (Pénzes et al., 2014; To et al., 2014). The sole member of genus *Ichtadenovirus*, white sturgeon AdV was found to contain four unusually located fiber-like genes at the left end of its genome (Doszpoly et al., 2019), however their function has not been elucidated yet.

The AdV fiber is a homotrimer and consists of an N-terminal tail, a shaft, and a C-terminal knob, and binds to the cellular receptors with its knob domain (Lasswitz et al., 2018). Subsequent binding of the RGD (arginine-glycine- aspartic acid) motif of the penton base to cellular integrins allows virus endocytosis (Descamps and Benihoud, 2009). Cellular receptors described so far to bind the fiber knob are the coxsackie and adenovirus receptor (CAR) (Roelvink et al., 1998; Seiradake et al., 2006), desmoglein-2 (Wang et al., 2011), CD46 (Marttila et al., 2005), sialic acid (Sia)-containing glycans (Nilsson et al., 2011) and others (Lasswitz et al., 2018). The length and flexibility of the fiber shafts seem to play a role in interaction with the receptors, as it was shown that short and sturdy fibers cannot bend to bind CAR on a cell surface (Wu et al., 2003). Cellular entry of HAdV-40 and -41 (HAdV-F), is assumed to be fundamentally different from that of other HAdVs since their penton bases lack the RGD motif (Albinsson and Kidd, 1999), and recently it has been shown that they may use a different repertoire of the laminin-binding integrins as co-receptors for infection (Rajan et al., 2018). Despite the extensive analyses of the AdV genomes of the members of the different genera and species, there is very limited knowledge published about the AdVs with more than one fiber gene, particularly about their receptor usage. So far, it has only been shown that HAdV-52 (HAdV-G) short fiber (SF) knob (SFK) binds to Sia-containing glycans (Lenman et al., 2018, 2015), but no target or function has been described for any of the SFs either in the further HAdVs (HAdV-40 and -41), or in any of the non-human AdVs. The HAdV-41 SFK is the only other structurally characterized knob of SFs (Seiradake and Cusack, 2005).

The main purpose of the present study was the analysis of full genomic sequence of SAdV-2 and SAdV-17. The host species of the isolates were rhesus macaque [*Macaca mulatta*] and grivet [*Chlorocebus aethiops*], respectively. The former OWM is native to South, Central, and Southeast Asia, and it has feral colonies overseas and a well-documented history as model-species in various fields of science (Singh et al., 2020). Whereas grivets are resident only in a few East-African countries in the savanna and shrubland habitats, with the widest area in Ethiopia (Butynski and De Jong, 2022). SAdV-2 and SAdV-17 have been assigned to different serotypes based on conventional serology, i.e., cross-neutralization test (Rapoza, 1967), later they have been classified into two separate species according to phylogenetic analyses of partial sequences (Pantó et al., 2015). For the first time in any serotyped member of the family *Adenoviridae,* we have revealed three fiber genes in two AdVs with apparently diverse evolutionary origins. Furthermore, an unusual reiteration of the genome end of SAdV-2 is also reported, a unique feature in HAdVs and SAdVs so far. The existence of three fiber genes might raise new possibilities for engineering AdV vectors targeting different receptors.

2. Materials and methods

2.1. Cells and virus stocks

The isolates studied originated from the American Type Culture Collection (ATCC). No information on pathogenicity was available on SAdV-2 (ATCC VR-196, strain 646776, isolated from rhesus macaque [*Macaca mulatta*]) and SAdV-17 (ATCC VR-942, strain B-105, from grivet [*Chlorocebus aethiops*]), besides the host species. For virus propagation, Vero E6 cells were infected with the isolates. Eight 175 cm^2 tissue culture flasks were used for each virus type. When full CPE was reached, the flasks were frozen and thawed three times. After centrifugation on 4 °C, at 1300 × *g*, for 30 min, the virions were pelleted by ultracentrifugation (41,000 ×*g*, 4 °C, 90 min, Beckman rotor 60 Ti), and the viral DNA was extracted with phenol-chloroform extraction method for the subsequent multiple parallel sequencing (MPS).

2.2. Sequencing, PCR, genome assembly, and analysis

The MPS was performed by a commercial service on an Illumina platform (BaseClear, Leiden, The Netherlands) and the reads were processed as described earlier (Podgorski et al., 2016). Briefly, quality-trimmed paired-end reads were '*de novo* assembled' with stringent conditions using the CLC Genomics Workbench version 7.0.4. When this assembly resulted in multiple contigs for the viral genome of SAdV-17, the gaps were filled by PCR using specific primers designed with the use of the Primer Designer program version 2.0. For the amplification of these fragments, Takara PrimeSTAR® Max DNA polymerase was used according to the manufacturer's recommendations. The PCR products were purified from agarose gel with the use of the MEGA quick-spin Total Fragment DNA Purification Kit (iNtRON Biotechnology, Kyungki-Do, Korea) and sequenced by traditional (Sanger) method with the use of the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Waltham, MA, USA). The genome ends were justified by using primers designed for the inverted terminal repeat (ITR) sequence (Podgorski et al., 2016), with the conditions of the sequencing reactions and nucleotide sequence assembly also described in detail previously (Pénzes et al., 2014; Tarján et al., 2014). The ORFs and presumed genes were identified by the CLC Workbench. The sequences of the genes, known to contain introns in other AdVs, were checked for the presence of putative splice donor and acceptor sites. The VA-RNA gene sequences of SAdV-2 and SAdV-17 were determined by comparison with the available VA-RNA sequences of primate AdVs, with a special focus on the criteria for identification of VA-RNA genes and fully conserved nucleotide positions (Kidd et al., 1995). The sequences of fiber genes were analysed based on the previously reported models and amino acid (aa) distributions in the β-spiral repeats of the shaft domain (Van Raaij et al., 1999), which include three hydrophobic residues (h), two other (hydrophobic or polar) residues (h/p), and either proline (P) or glycine (G), i.e. "a a h a h a *h/p* a…a *P/G* h a *h/p* a", with the possibility of the disruptions in the sequence repeat pattern causing a kink in the shaft. Fiber genes were structurally aligned to fibers with known structures using the PROMALS3D multiple sequence and structure alignment server (Pei et al., 2008).

2.3. Phylogenetic calculations

Phylogenetic calculations were carried out essentially as described earlier (Podgorski et al., 2018), by using the ProtDist and PhyML algorithms of the PHYLIP package (Felsenstein, 1993) and the Montpellier bioinformatics platform (Guindon and Gascuel, 2003), respectively. The JTT model with the invariable sites and gamma distribution options was applied. Analyses were based on full aa sequences of the DNA-dependent DNA polymerase (pol) from primate AdVs sequenced to date. For a detailed comparative study, fiber protein sequences available from HAdVs and SAdVs were retrieved from GenBank and their structural models from the Protein Data Bank (PDB), to match them in a structure-based alignment using the PROMALS3D (see above). "A la Carte" mode of the phylogeny server (http://phylogeny.lirmm.fr) was applied using Gblocks to eliminate poorly aligned positions and divergent regions from the alignments, followed by a Maximum-Likelihood-based construction of phylogenetic trees.

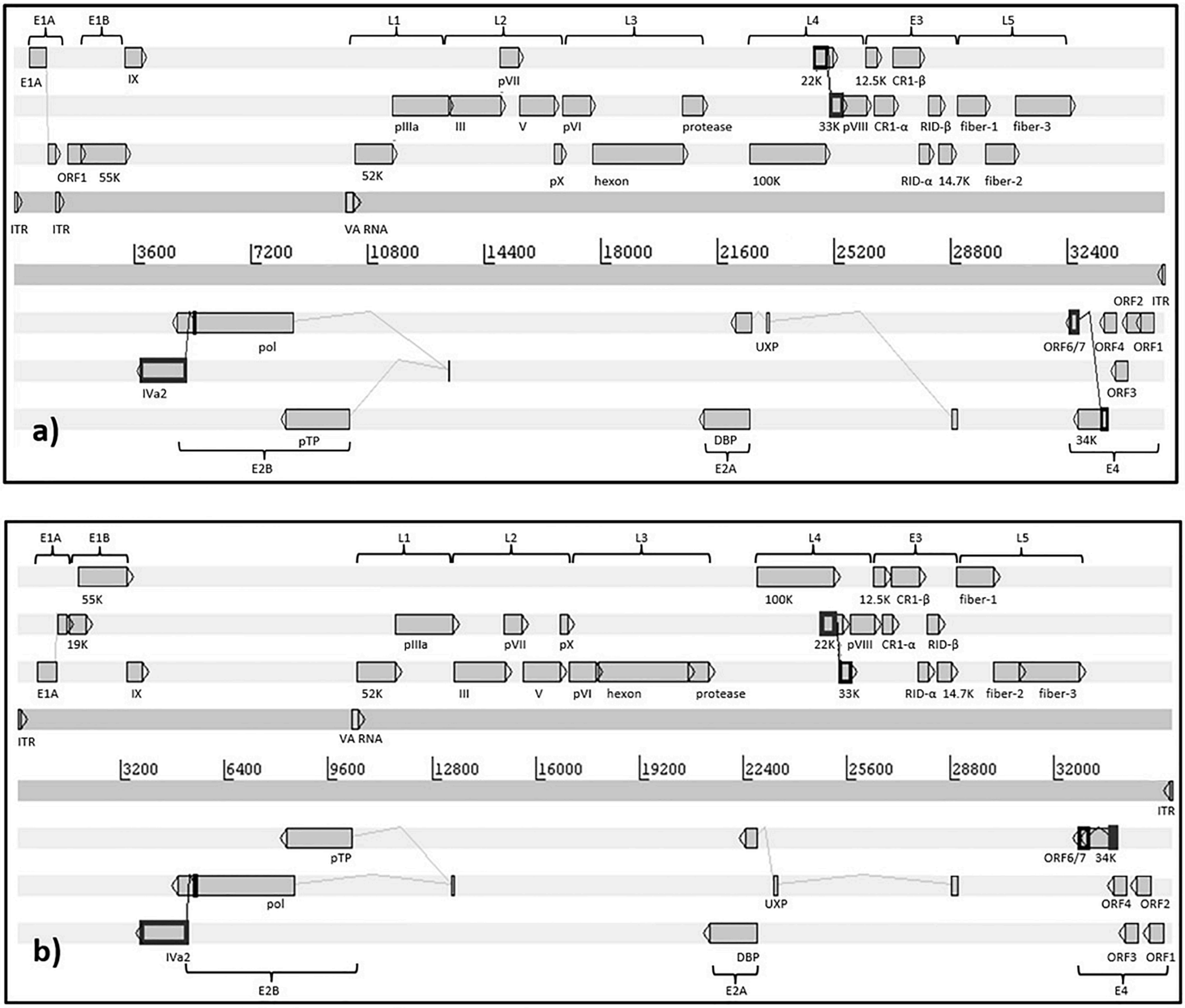

Fig. 1. The annotated genomes of SAdV-2 (a) and SAdV-17 (b), members of the species *Human mastadenovirus G* and *Simian mastadenovirus F*, respectively. The putative encoded proteins are marked with arrows of proportional size. The predicted splicings are depicted by thin lines connecting the exons. Coding sequences partly shared with another coding sequence are highlighted by thick frames. Early (E1A, E1B, E2A, E2B, E3 and E4) and late (L1–L5) transcription units are marked above and under the genomes.

3. Results

3.1. Genome characteristics

From the SAdV-2 isolate, we obtained a single contig with a coverage over 24,400: 35,481 bp, GC content of 54.7%, and 59 bp-long inverted terminal repeats (ITRs). In the case of SAdV-17, three contigs were assembled with lower coverage (from 710 up to 6540) and were later joined and completed by PCR and Sanger sequencing. The length of the full genome complemented with the ITRs was found 35,670 bp with 58.8% GC content and 81 bp-long ITRs. The sequences were deposited to GenBank under accession numbers MG969551 (SAdV-2) and KP329566 (SAdV-17). In both genomes, we identified 37 putative coding regions (ORFs) characteristic of mastadenoviruses (Fig. 1). These included three genes of different lengths predicted to code for the cellular attachment protein, the fiber. Also in both genomes, a single copy of the VA-RNA gene was found. In the SAdV-2 genome, in place of the missing E1B 19K gene, an extra copy of the ITR and the E4 ORF1 were found. The three exons of the so-called U exon protein (UXP) (Podgorski et al., 2016; Tollefson et al., 2007) could also be determined in comparison to their counterparts in members of SAdV-B, -C, -E and HAdV-C and -G (Podgorski et al., 2016; Tollefson et al., 2007). Basic genome characteristics of the newly studied SAdVs and the relevant values in their respective virus species are listed in Table 1. Most of the deduced protein sequences of SAdV-2 and -17 were very similar (with over 95% aa identity) to their counterparts in the representatives of the same species.

The SAdV-2 genome seemed to be closest to that of the recently isolated rhesus macaque AdV-61 (RhAdV-61) (Abbink et al., 2018), except for the aforementioned changes in the E1B region. However, SAdV-2 and the almost identical RhAdV-61 are very different in several proteins from all the other so far described AdVs. E.g., SAdV-2 UXP shows a maximum of 81% identity, its CR1-α 64%, its CR1-β 50%, and the three fiber genes show up to 76%.

SAdV-17 and its closest relative SAdV-18 in the species SAdV-F shared high percent identity in their hexon (90%), DBP (91%), ORF6/7 (87%), ORF4 (87%) and almost the whole E3 region. The three fiber genes of the SAdV-17 showed very different identities: the short fiber 1 (SF1) 46% to HAdV-F members, the SF2 48% to SAdV-C members, and the long fiber (LF) 99% to SAdV-18 (SAdV-F). (Notably, SAdV-18 has only a single fiber gene.) Both studied AdVs contain the integrin-binding RGD motif in their penton base protein sequence.

3.2. Phylogenetic analyses

The phylogeny reconstructions, performed with the aa sequences deduced from the full-length *pol* gene (ML method, JTT + I + G model; Fig. 2), showed that SAdV-2 belongs to HAdV-G, while SAdV-17 clusters with SAdV-F. These species classifications have been already proposed previously based on partial IVa2, penton base, and hexon gene sequences (Pantó et al., 2015). The rightmost-located fiber gene (LF gene or fiber-3 gene) sequences of both SAdV-2 and SAdV-17 showed close identity (up to >99%) to the already described LFs of other members in their corresponding species (Table 2).

When inspecting the short fibers (SFs) of SAdV-2, besides the almost complete identity to RhAdV-61, they share up to 55% (SF1) and 68% (SF2) identity with other members in SAdV-C or HAdV-G, respectively. The SFs of SAdV-17 showed very different and low identities to the other so far described SFs (Table 2). Its SF1 shares 46% identity with HAdV-41 (HAdV-F) SF gene, whereas SF2 shares 48% identity with SAdV-19 (SAdV-C) SF. Thus, these two fiber genes are supposed to have different evolutionary origins (and are not a result of a recent gene duplication). To test this hypothesis, we attempted to reconstruct the phylogenetic relatedness among fiber protein sequences from multiple-fibered HAdVs and SAdVs. As the solely sequence-based alignment algorithms (MAFFT, MUSCLE, ClustalW) were inconsistent in their performance with the highly divergent fiber protein sequences, we made a structure-based alignment, using PROMALS3D. This program correctly aligned the conserved motifs of the fiber; thus, it could be used more reliably in the phylogenetic analysis. The phylogenetic tree based on such an alignment (Fig. 3a) shows that the "rightmost fibers" (fiber-2 or -3, in cases of in HAdVs and SAdVs with double and triple-fiber genes) form a monophyletic group. This group contains the SAdV-B, -C, -E, -F and HAdV-F, -G branches. The pol-based calculation resulted in a similar tree topology (Fig. 2). The fibers encoded upstream from these presumed ancient "right-hand fiber" genes do not form a single monophyletic group in the phylogeny inference. (Please note that in the case of SAdV-B and -E species, these "left-hand fibers" are the longer ones, but in all other HAdV and SAdV species the upstream encoded ones are rightfully often referred to as short fibers [SFs]; see Fig. 3b). In many cases, the phylogenetic branching among the left-hand fibers reflects those seen among the right-hand ones (Fig. 3a) and subsequently those of the pol tree (Fig. 2), affirming that during the evolutionary diversification of these AdVs, there has not been any recombination in these parts of their genomes. On the other hand, there are several exceptions as well, when SFs deviate from this coevolution pattern. For example, SF1 of SAdV-17 clusters together with HAdV-F fiber-1 in a larger monophyletic group ("Clade1") that contains fiber-1 proteins of SAdV-B, -C and -E, branched very similar to that seen among their "ancient" fibers (Fig. 3a). Whereas SF2 of SAdV-17 clusters into "Clade2", most closely related to an outlying SF of a SAdV-C member (SAdV-19). This apparent polyphyly of SAdV-C fiber-1 (SF) sequences suggests that either a parallel evolutionary event with two separate SF gene acquisitions had taken place within this species, or as proposed in Fig. 3b, a recombination happened. Next closest sisters to SF2 of SAdV-17 include SF1 of SAdV-2 and three other HAdV-G members (RhAdV-55, -59, -61). Altogether the aforementioned SFs form a strongly supported monophyletic upper branch in "Clade 2" (Fig. 3a), with further unparalleled kinship patterns to that seen among the "ancient" fibers as signs of non-coevolution processes.

Concerning the SF2 protein of SAdV-2, it seems to be very distantly connected to all other fibers, except for the SF2s of the before mentioned three HAdV-G members (RhAdV-55, -59, -61), and presumably also to SF1s from other three rhesus AdV genomes (RhAdV-60, -65, -66 of HAdV-G) ("Clades 3 and 4" on Fig. 3a). Kinship between "Clade 3 and 4" members was apparent on a more comprehensive phylogenetic tree, including all (also the single fibered) HAdV and SAdV fibers (data not shown) or can be extrapolated from a BLASTp search. E.g., testing SAdV-2 SF2 in a BLASTp search results first hits of "Clade 4" members with 68% identity (ID) and 86% similarity (SIM) values. Yet, the second-best hits are "Clade 3" members with 47/61% (ID/SIM) and two stretches of insertions (17% gaps) in the query. Omitting these gaps of 47 amino acids from the comparison, ID/SIM values are as high as 55/74%.

Table 1
Genome characteristics of the simian adenovirus (SAdV) types studied in this work.

Virus type	Virus species	ATCC No. of isolates	Host monkey species	Genome length (bp)	ITRs length (bp)	ITRs in the species (bp)	GC content (%)	GC content in the species (%)	Accession number
SAdV-2	HAdV-G	VR-196, strain 646,776	rhesus macaque (*Macaca mulatta*)	35,481	81	60–133	54.7	55.1–56.3	MG969551
SAdV-17	SAdV-F	VR-942, strain B-105	grivet (*Chlorocebus aethiops*)	35,670	59	180	58.8	61.4	KP329566

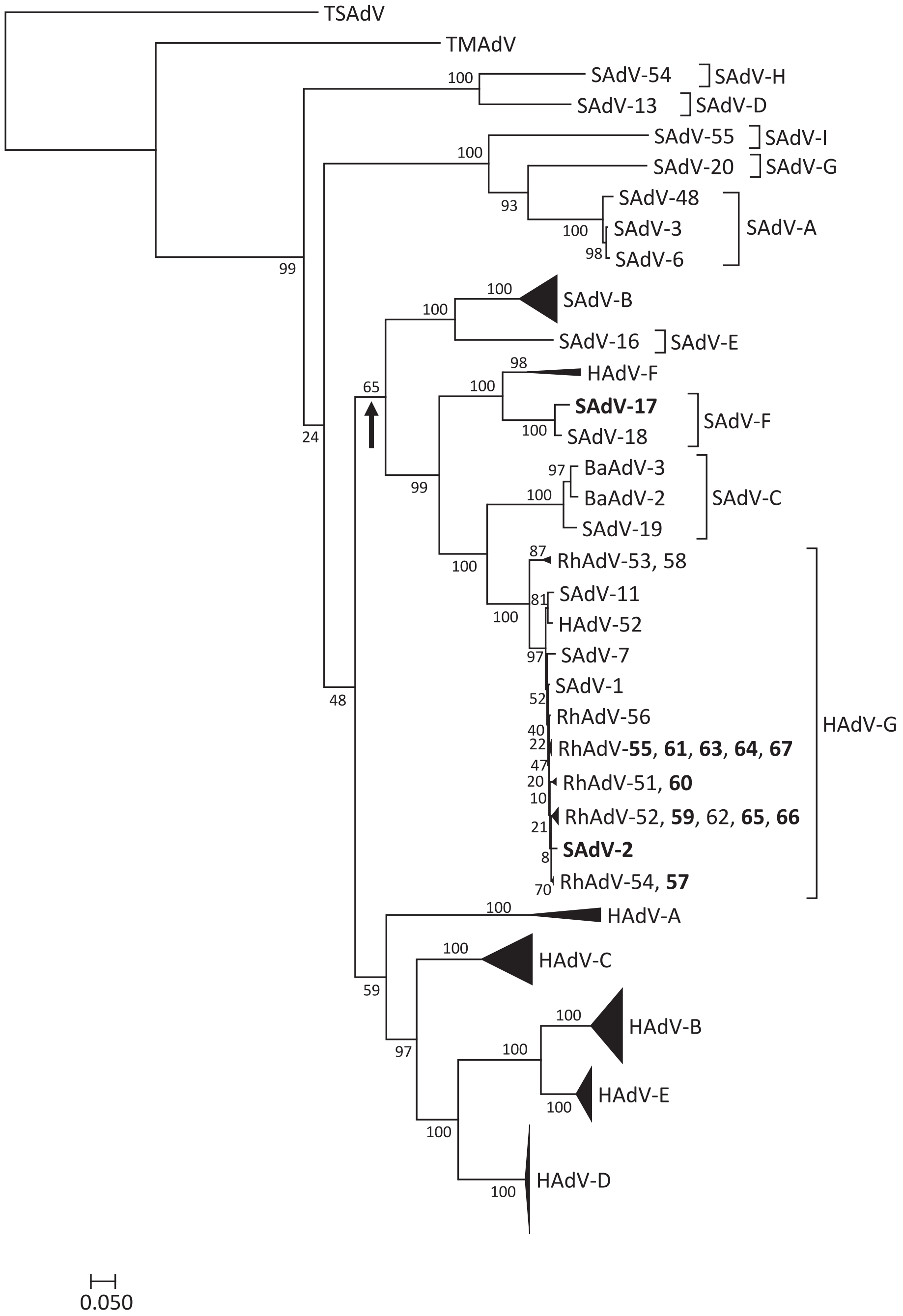

Fig. 2. The phylogeny reconstruction performed by PhyML with the aa sequences deduced from the DNA-dependent DNA polymerase (JTT model with the invariable sites and gamma distribution options used). Bootstrap values are shown. Branches of members of some simian and human adenovirus species were collapsed for clarity. Black arrow shows the AdV lineage with at least two fiber genes (except SAdV-18). Bold numbers depict the RhAdV isolates with three fiber genes. SAdV-55 corresponds to strain WIV19. RhAdV - rhesus macaque AdV; HAdV - human AdV; TSAdV - tree shrew AdV; TMAdV - titi monkey AdV. For the accession numbers of the referred sequences please visit site: https://sites.google.com/site/adenoseq/

Table 2
Sequence identity comparisons of the short (SF) and long fibers (LF) and the encoded protein parts (tail, shaft, knob) of the SAdVs studied in this work to other AdVs available in the GenBank. The species names of the most similar AdVs belonging to are shown in bold if differs from the DNA polymerase-based classification.

	Fiber identity	Compared to	Tail identity	Compared to	Shaft identity	Compared to	Knob identity	Compared to
[*]SAdV-2 SF1 (HAdV-G)	99%	RhAdV-61 (HAdV-G)	100%	RhAdV-61	97%	RhAdV-61	100%	RhAdV-61
	55%	SAdV-19 (**SAdV-C**)	88%	RhAdV-67 (HAdV-G)	44%	SAdV-19	66%	RhAdV-66
[*]SAdV-2 SF2	99%	RhAdV-61	100%	RhAdV-61	100%	RhAdV-61	99%	RhAdV-61
	68%	RhAdV-59 (HAdV-G)	81%	RhAdV-51 (HAdV-G)	76%	SAdV-11	66%	RhAdV-59
[*]SAdV-2 LF	99%	RhAdV-61	100%	RhAdV-61	99%	RhAdV-61	99%	RhAdV-61
	77%	RhAdV-53 (HAdV-G)	92%	RhAdV-53	73%	RhAdV-53	81%	RhAdV-53
SAdV-17 SF1 (SAdV-F)	46%	HAdV-41 (**HAdV-F**) [#]	81%	HAdV-40 (HAdV-F) [#]	49%	HAdV-41 [#]	38%	HAdV-41
SAdV-17 SF2	48%	SAdV-19 (**SAdV-C**) [#]	64%	SAdV-19 (**SAdV-C**) [#]	58%	SAdV-19 [#]	50%	SAdV-19
SAdV-17 LF	99%	SAdV-18 (SAdV-F)	83%	SAdV-18	99%	SAdV-18	100%	SAdV-18

[*] As SAdV-2 is almost the same as RhAdV-61, the second highest identity score is shown, too.
[#] These fiber genes are not present in SAdV-18, the single other known SAdV-F member.

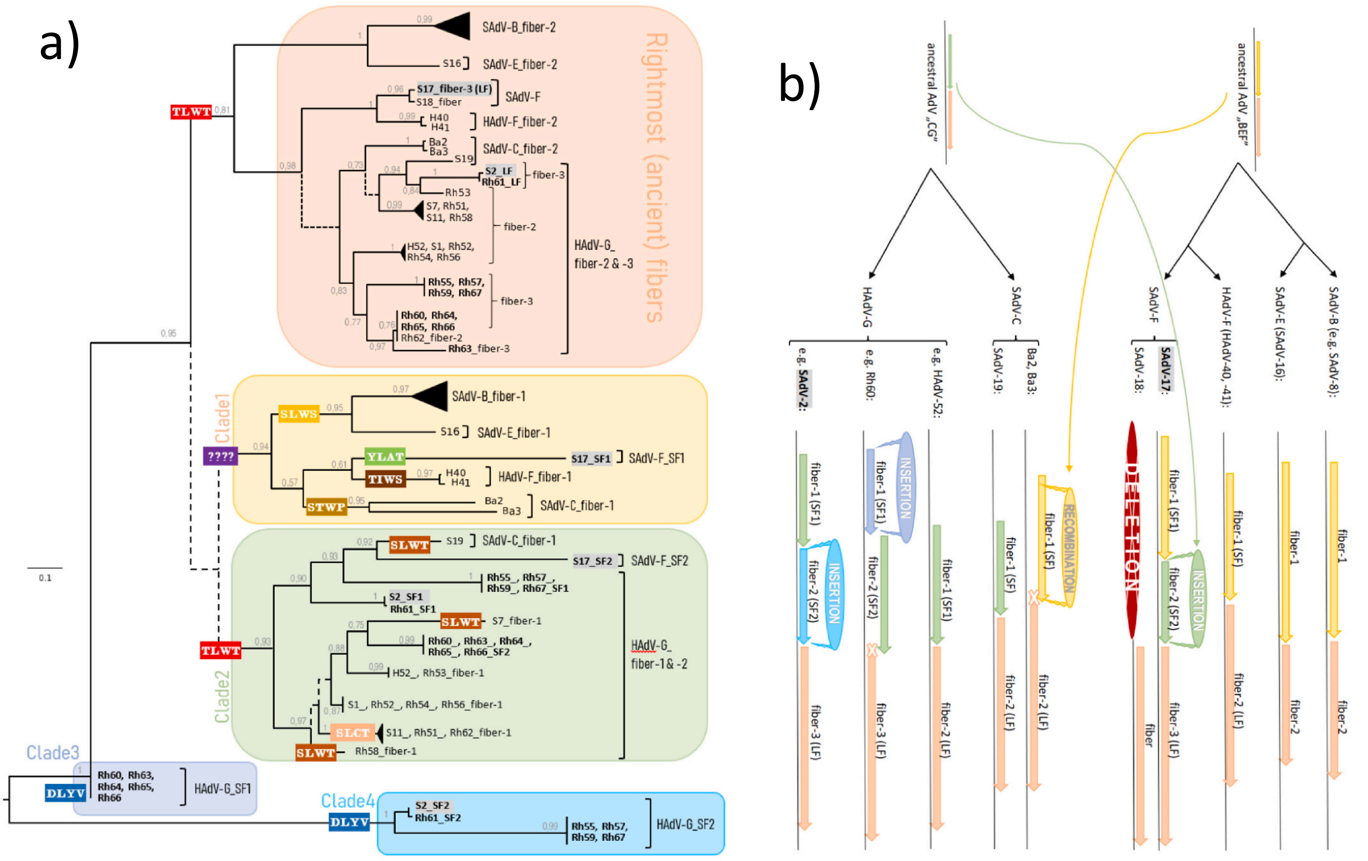

Fig. 3. Phylogenetic tree of fibers and the arrangement of the coding genes in the HAdV/SAdV lineages with two or three fiber genes. **A)** For the phylogenetic tree reconstruction, full fiber protein sequences were retrieved from the GenBank and alongside the HAdV-5 PDB sequence (3IZO) they were PromalS3D aligned. The resulting 729 positions were GBlock edited to a 181 character long compact alignment, which was used for the PhyML analysis to reconstruct phylogenetic relations. One hundred bootstraps were performed, and branches with <0.5 (50%) were dashed on the consensus tree, higher bootstrap values are indicated. Some simian and human adenovirus species were collapsed for clarity. Bold letters were used for the labels of serotypes/strains with 3 fiber genes. Bold numbers depict the RhAdV isolates with three fiber genes. Background colours of the various clades of fiber sequences refer to those used for fiber gene arrangement diagrams in 'section B' of this figure. White lettered rectangular tags show the modification of knob-start motifs for the progeny types/strains, same colours represent identical motifs. **B)** Illustration of fiber arrangements and similarities in HAdV and SAdV species containing multiple homologs of this gene. Colours of fiber genes match those of the clades on 'part A' of this figure. Fat arrows representing the different fiber genes are to scale both in length and overlap. Presumably misidentified upstream LF starts, resulting in overlaps of two fiber genes, are crossed out. Hypothetical arrangements in ancestral AdVs are illustrated in smaller size, using the same colours for the homologous ORFs. Slender curved arrows starting from these genes represent the presumptive origin of the indicated insertion and recombination events. H - human AdV; Ba – baboon AdV; Rh - rhesus macaque AdV; S - simian AdV. For the accession numbers of the referred sequences please visit site: https://sites.google.com/site/adenoseq/

3.3. Characterization of fiber genes

The sequences of the fibers were analysed based on the previously reported model of fiber structure (Van Raaij et al., 1999). Parts of the fibers, the N-terminal tail, the shaft and the C-terminal knob (head), were predicted, as well as the number of pseudo repeats in the shaft. These were compared to the most similar SFs belonging to the HAdV-F, -G and SAdV-C species (Fig. 4).

All newly described SFs showed to have different numbers (from 2 to 11) of the predicted repeats within the shaft (Fig. 4). Disruptions of the pattern sequence were observed in the third repeat of the SAdV-17 SF1 and LF, as well as in the 17th and 18th repeat of the SAdV-2 LF (Fig. 5). SAdV-2 SF1 and SAdV-17 SF2 are the only newly studied short fibers, the knob of which contains the polySia-binding RGN (arginine-glycine-asparagine) motif and conserved TLWT (threonine-leucine-tryptophan-threonine) motif that marks the beginning of the fiber knob domain of the mastadenoviruses (Fig. 4).

We separately compared the sequence identity of the parts of the SFs to those of the so far described fibers (Table 2). The results showed that the tails of SF1 and SF2 of SAdV-17 have >64% identity to either species HAdV-F or SAdV-C members, respectively. However, shafts and knobs have only up to 58% identity. On the other hand, LF is almost identical to the only other type in its species, SAdV-18, which lacks the two other fiber genes.

In the case of SAdV-2, besides RhAdV-61, with which it is almost identical, the highest identity to some other members of its species, HAdV-G, was observed, but sometimes also to the members of the species SAdV-C (shaft of the SF1).

4. Discussion

Here we report, for the first time in any serotyped AdV, three fiber genes (as SAdV-1 to SAdV-25 were differentiated by classical cross-neutralization (Kidd et al., 1995)). The existence of three different fibers in a SAdV might give great potential for vectorising purposes in human medicine: e.g., the fibers could be switched to use up to three different receptors on the host cells. Parallel to our investigations, a recent study detected three fiber genes in ten untyped SAdVs, isolated from stool filtrates of rhesus macaques, and used as vector backbones (Abbink et al., 2018). Nevertheless, hardly any detail was published

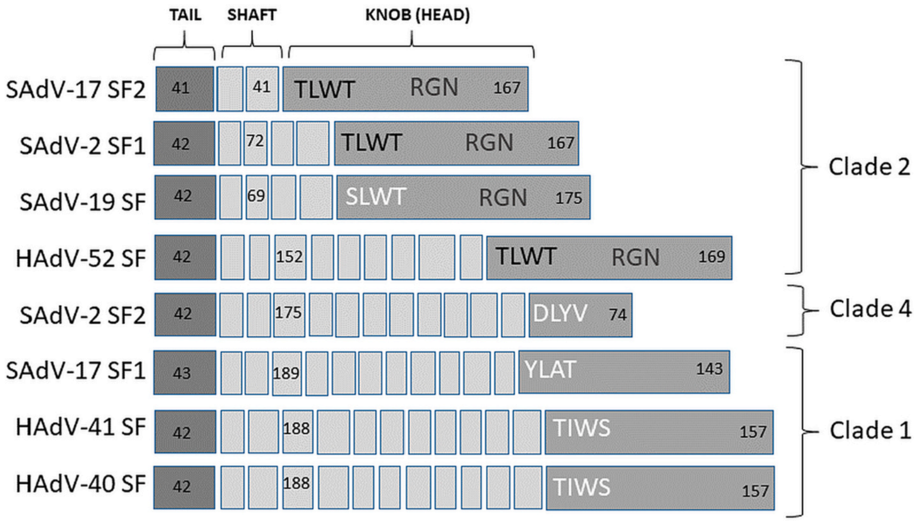

Fig. 4. Schematic representation of the short fibers (SFs) of the newly described AdVs compared to the most similar SFs from the members of the species HAdV-G (HAdV-52), HAdV-F (HAdV-40 and -41), and SAdV-C (SAdV-19). (Notes: SAdV-2 and RhAdV-61 have almost identical SFs, thus only the ones from the former are presented. SAdV-18, of species SAdV-F, does not have a short fiber, therefore it is not included in the scheme). The numbers refer to the aa length of each predicted fiber part (tail, shaft, knob). Short light grey sections depict predicted repeating patterns in each shaft. Conserved TLWT motif (or the corresponding sequence in white) which marks the beginning of the knob part and polySia-binding RGN sequence are marked where they are present. Clades of Fig. 3 are indicated on the right.

there about the “extra” fibers in these isolates. Here we attempted to give an overview of these fibers and their presumed evolutionary origin. Another recent study demonstrated that generating the phylogenetic trees based on fiber knob sequences, rather than whole genomes, can show additional diversity in the HAdV-D species, due to presumed homologous recombination in this part of the genome, as a result of adaptation to novel receptors (Baker et al., 2019). This phylogenetic approach was elaborated further in the present study to delineate the connections between the multiple fibers of SAdV and HAdV and to draw an evolutionary roadmap for the acquisition of second and third fiber genes in these AdVs.

Besides the full genome sequence analyses of two OWM adenoviruses, in this paper we also discuss their genetic characteristics and taxonomy. Of the newly sequenced viruses, SAdV-2 was confirmed to belong to the previously established species *Human mastadenovirus G*, whereas SAdV-17 belongs to the species *Simian mastadenovirus F*, thus supporting the establishment of this novel species. The GC content, an important AdV species demarcation criterion (Benkő et al., 2022) among SAdVs, varies between 47.8% (SAdV-20; (Roy et al., 2009)) and 65.7% (SAdV strain A1139; (Roy et al., 2012)). However, within each of the species, the differences in the base composition do not exceed 3% (Pantó et al., 2015). The GC content of the two SAdVs presently studied conformed to this rule (Table 1).

The size of SAdV genomes, sequenced to date, ranges between 31,045 (SAdV-7 (Roy et al., 2011)) and 35,726 bp (SAdV strain A1173 (Roy et al., 2012)). The two newly sequenced SAdV genomes fall within this range and have 37 putative genes typical of mastadenoviruses. Yet in both genomes, we found some unique differences. SAdV-2 is an exception from all the previously studied SAdVs as it does not contain E1B 19K gene but has an extra copy of the ITR and E4 ORF1 in this region instead. The ITR length of SAdV-2 (81 bp) falls within the broad ITR length range of the HAdV-G members (60–133 bp), whereas the ITRs of SAdV-17 are much shorter (59 bp) than those in SAdV-18 (180 bp), the only other known member of SAdV-F.

All OWM AdVs studied earlier, have been found to have only one VA-RNA gene (Kidd et al., 1995; Larsson et al., 1986), while human and chimpanzee AdVs (in species HAdV-B to HAdV-E) possess two of them (Kidd et al., 1995; Larsson et al., 1986). Next to the confirmed VA-RNA gene in SAdV-2, we also proved its presence in SAdV-17 (Fig. 1). Since there has been no information about the VA-RNA gene of SAdV-18, the other member of SAdV-F, we also identified it by in silico analysis, and confirmed its position (FJ025931.1; 10,597 to 10,766) and the same length as in SAdV-17 (169 nt). Notably, these findings confirm that all OWM AdVs studied to date, contain only one VA-RNA gene.

Part of the SAdV-2 genome (repetition of the ITR and E4 ORF1 in place of the E1B 19K gene; Fig. 1), which is different in this type from all the other types, was re-sequenced also by Sanger sequencing to confirm the result gained by the MPS. To exclude the possibility that the repetition was caused by a mutation during the propagation in Vero E6 cells in our laboratory, DNA samples from the virus both before and after the propagation were sequenced. While we excluded a possible sequencing mistake, we cannot exclude that this mutation happened during earlier propagation and was not present in the wild type virus. However, one of the vectors by Nabel et al. in 2011 (Patent application WO2011057254; accession no. JA453577), construct seq. 28, also contained the above-described duplicate ITR in the left part of the genome, upstream from the transgene cassette, in the position where the E1B 19K should be. Since several studies showed that artificial E1B 19K deletion mutant AdVs demonstrate enhanced tumour cell killing and therefore enhanced oncolytic potency (Liu et al., 2004; Sauthoff et al., 2000), there seems to be a promising potential for such a virus, which has three fiber genes and a natural deletion of E1B 19K gene, especially if the goal of the adenoviral vector application is targeted host cell destruction.

When analysing the fibers of e.g. SAdV-17 by BLASTp (Table 2), the SF1 and SF2 genes seemed to have different origins, most similar to SFs of HAdV-F and SAdV-C, respectively. The rest of the SAdV-17 genome, including LF, was highly similar to SAdV-18 of the same species: SAdV-F. As SAdV-18 contains only this fiber gene (LF), a possible evolutionary scenario is that while diverging from a common ancestor, SAdV-17 acquired two additional fibers in separate recombination events, and/or the earlier acquired short fibers were lost in SAdV-18. To test these hypotheses, we attempted to reconstruct the phylogenetic relatedness among fiber protein sequences from multiple-fibered HAdVs and SAdVs.

The structure-based approach for aligning the fiber protein sequences resulted in the best match, as seen in the conserved motifs of these highly diverse sequences, thus minimizing the possible distortion for the downstream applications. On the phylogenetic tree (Fig. 3), it is clearly outlined that all rightmost located fibers of each of these HAdV and SAdV strains cluster together in a monophyletic group. In this exclusive group of rightmost fibers, the sub-branching resembles that seen on the DNA-polymerase-based tree (Fig. 2), which implies that the common ancestor of these fibers was present in the progenitor virus of these primate AdVs and has coevolved/diverged with the rest of their genomes. Therefore, we designated these rightmost fibers as “ancient”.

The “non-ancient” fibers, however, do not form a single monophyletic cluster, but rather four distinct ones (“Clade1” to “Clade4” in Fig. 3a), suggesting multiple origins due to separate duplication and/or recombination events. “Clade1” delineates several fiber-1 s, including the first short fiber (SF1) of SAdV-17. The collinear arrangement is seen among these fiber-1 s and the “ancient” fibers of the same viruses, this

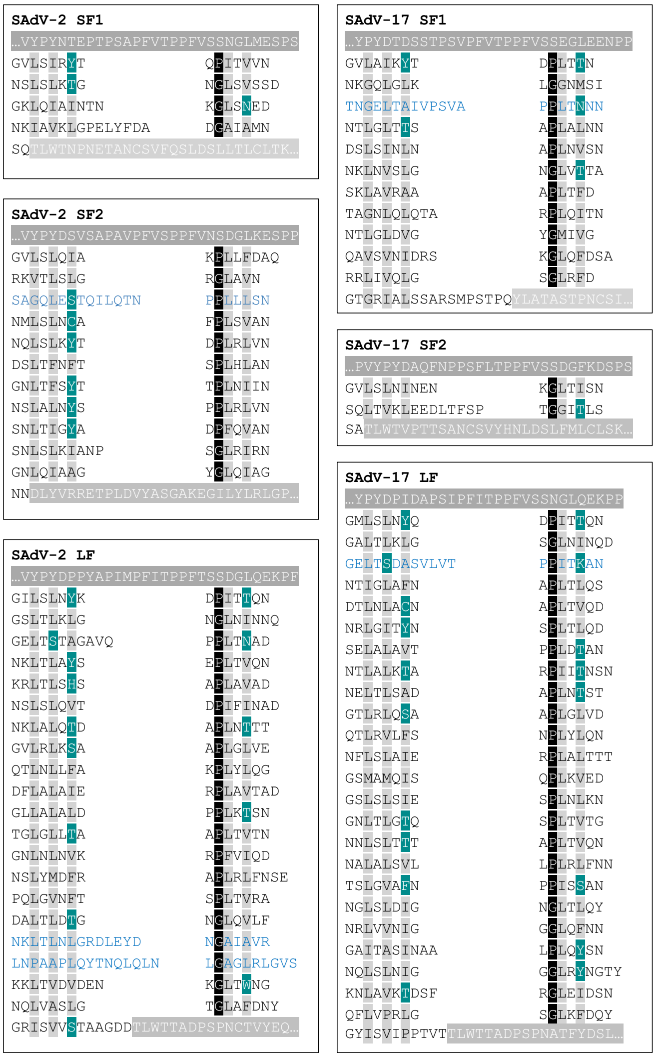

Fig. 5. Sequence analysis of fiber shafts of the newly described SAdVs. For simplicity, fiber tail (dark grey) and knob (light grey) domains are only partially showed at the beginning and at the end of the shaft domain, respectively, with white letters. Hydrophobic aa in the shaft are highlighted with light grey, polar aa with cyan, proline (P) and glycine (G) with black. The shaft repeats with the largest disruptions in the repeat pattern possibly causing kinks are printed in light blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

points to a common double-fibered ancestor for SAdV-B, -E, -F, and HAdV-F and two SAdV-C members. Among the currently known descendants of this probable ancestral virus, there is only one species where both single-fibered and triple-fibered types also occur: SAdV-F. In this species, the second short fiber (SF2) of SAdV-17, nonetheless clusters in “Clade2”, alongside the fiber-1 of the outlying SAdV-C member: SAdV-19, and with most of the HAdV-G short fibers. This second group of “non-ancient” fibers also points to the foretime existence of a progenitor with two fibers, which later diverged to the several strains found in today’s species HAdV-G, and its progeny could have been the donor for the SAdV-17 SF2. Regarding the two currently known SAdV-F members, the ancestor of SAdV-18 most probably has also had SF genes (as all other genes of SAdV-17 and SAdV-18 are very similar) but a large deletion affected almost the whole E3 region (leaving only part of 12.5 K, the leftmost gene in E3) deleting also the first two fiber genes from the three ones. Whether the progenitors of “Clade2” and “Clade1” originated from a double-fibered prime ancestor or were the results of two separate gene duplication events, could not be decided.

The fibers of SAdV-2 are almost identical (99%) with those of RhAdV-61 (HAdV-G). The SF1 shares the second highest identity (55%) with SAdV-19 (SAdV-F) SF as well as with other HAdV-G members. The SF2 and LF of SAdV-2 only show identity with members of their own species, HAdV-G. On the phylogenetic tree (Fig. 3a), the SAdV-2 LF (fiber-3) and SF1 clustered together with their expected homologs the HAdV-G “rightmost fibers” and the HAdV-G short fibers in “Clade2”, respectively. Nevertheless, in both cases, these HAdV-G fiber groups are paraphyletic, and to extend them monophyletic, both within the “rightmost fibers” and for the “Clade2” SFs, further SAdV-C LFs and the SAdV-19 SF plus SAdV-17 SF2 should also be included, respectively. This close kinship with the SAdV-19 is consonant with the BLAST results mentioned above. At the same time, SF2 of SAdV-2 is an outlier of both large SF-clades (“Clade1 and 2”, Fig. 3a) and together with the SF2s of five rhesus AdV isolates, it forms “Clade4”. The other outlier group of the tree, “Clade3” is formed by SF1s of five other rhesus AdV isolates. These ten isolates are the ones that encode three fiber genes, among the recently reported 16 rhesus isolates (Abbink et al., 2018), all belonging to HAdV-G based on the full pol aa phylogeny (Fig. 2). According to the fiber-phylogeny, we presume that the double-fibered HAdV-G progenitor picked up a third fiber gene at two different times during the evolution, once as the leftmost fiber gene (“Clade3” members), while the other time between the two extant fiber genes (in the middle position: “Clade4” members).

Analyses of the primary and predicted secondary structures of the fiber proteins suggested the presence of SF tails (virus-anchoring domain) in SAdV-2 SF2 and SAdV-17 SF1 comparable in length to other previously described AdVs (Fig. 4). The tail of SAdV-17 SF1 and SF2 have a somewhat higher percentage of identities to AdV fibers of known sequences; however, the shafts and knobs have very little sequence identity, only up to 58% (Table 2). The predicted central shaft domain of SAdV-17 SF2 was the shortest one observed so far, with two putative triple β-spiral repeats. Twice as long shaft was observed in SAdV-2 SF1 (four repeats), and the longest ones in SAdV-17 SF1 (11 repeats), comparable in length to other members of SAdV-F, HAdV-F and -G. The putative C-terminal knob (head) domains also varied in length, with the longest one observed in SAdV-2 SF2.

The beginning of the fiber knob region can be recognised by the characteristic sequence-motif: TLWT immediately after the last β-repeat of the shaft, which was described well conserved in HAdVs (Xia et al., 1994). This typical knob sequence was absent in the examined SAdV-2 SF2 and SAdV-17 SF1. Analysis of SAdV-19 also showed that it does not contain this sequence, and neither HAdV-F members do. Earlier studies showed that TLWT is not conserved either in other genera, e.g. in bovine AdV-7 (Li and Tikoo, 2002), an atadenovirus or in aviadenoviruses: fowl adenovirus 4 (FAdV-4) (Mase et al., 2010), FAdV-8 and -11 (Grgić et al., 2014), etc. We indicated the knob-starting motifs of all multiple-fibered HAdVs and SAdVs on our fiber-based phylogenetic tree (Fig. 3a) and concluded that the classical sequence motif is conserved in all rightmost fibers and exits in several non-ancient fibers (SAdV-17 SF2 and numerous HAdV-G short fibers in “Clade2”). Yet, most of the non-ancient fibers contain a slightly or largely different motif, namely in all “Clade1, 3, and 4” members, as well as in the sole SAdV-C member of the “Clade2” alongside the fiber-1s of five double-fibered HAdV-G members. Presumably, the double-fibered progenitor of the HAdV-G species contained the classical knob-motif on both of its fibers and the acquisition of the third fibers, either to the left or to the middle positions, might have used the same or similar donor AdVs, as the motif is the same (DLYV) in both the “Clade3 and 4” short fibers. To define the possible donor AdV-lineage of the third fibers for the HAdV-G members, we performed a comprehensive phylogenetic comparison of all published HAdV and SAdV fiber sequences, including also the numerous single-fibered types. Unfortunately, this analysis could not find close relatedness of the above-mentioned “Clade3 and 4” members with any other known HAdV or SAdV fiber (data not shown), thus leaving the question of the donor lineage unanswered. But “Clade1” and “Clade2”, as well as the “Clade3” and “Clade4” are clearly connected to each other, supported by high bootstrap values. Thus, we may surmise a single occasion of fiber-gene duplication and a single double-fibered ancestor for all multiple-fibered HAdVs and SAdVs (“Clade1 and 2” connected) as well as suppose that the third fiber in HAdV-G members originated from a single source, but integrated twice, either to the left or to the middle position, and in these positions, they diverged with the rest of the genome.

The fiber mediates the primary interaction with host cells, i.e., cell tropism. The entry into the cells requires two sequential steps, first the interaction of the fiber protein with its cellular receptor (Bergelson et al., 1997; Tomko et al., 1997), and the second the interaction between the RGD sequence of the AdV penton base and integrins on the cell surface (Bai et al., 1993; Wickham et al., 1996), after which the internalization of the virus is possible. Therefore, by incorporating ligands into the viral capsid or by swapping the fibers, it is possible to retarget the AdV vectors. As shown previously in many studies, fibers can vary a lot in flexibility, depending on the shaft length, i.e. on the number of β spiral repeats in the shaft (Van Raaij et al., 1999). Short fibers are thought to be more rigid and thus stricter in virus-cell receptor interactions and specificity, therefore the fiber length can have a role in cellular tropism. It is also known that disruptions of the sequence repeat pattern within the shaft result in shaft kinks (Chroboczek et al., 1995). The observed repeat pattern sequence disruption in the third repeat of SAdV-17 SF1 could be correlated to such kink. Some previously studied AdVs proved to have both long and short-shafted fibers, for example HAdV-52, which gives them the ability to bind more than one type of cellular receptors, i.e. CAR and Sia-containing glycans in the case of HAdV-52 (Lenman et al., 2015). The observed three fiber genes in the studied SAdVs suggest a possible even greater receptor diversity and therefore greater infectivity properties of these AdVs.

Interestingly, in the SAdV-17 SF2 knob (SF2K), we could identify the polySia-binding RGN motif that has previously been found to be present only in the SFs of members of species HAdV-G (SAdV-1, -2, -7 and -11) and SAdV-C (SAdV-19). The motif was not found in the SAdV-2 SF2K and SAdV-17 SF1K either. In our earlier study, the binding properties of RGN-containing AdV knobs from members of HAdV-G, including SAdV-2 SF1K, to polySia-expressing cells were analysed (Lenman et al., 2018). Despite possession of the RGN motif, SAdV-2 SF1K was not able to bind polySia. We presumed this to happen because the knob sequence was only distantly related to the other HAdV-G knobs, which might result in a different overall arrangement of the residues (Lenman et al., 2018). We have shown earlier with the phylogenetic analysis of SFs that the SAdV-19 fiber is distant from all the known AdVs (Podgorski et al., 2016), but more closely related to HAdV-G members than to those in its own species, SAdV-C. Therefore, we should not exclude the possibility that SAdV-2 SF1 went through recombination processes with yet-unknown members of SAdV-C.

Our findings give potential for novel virus-host binding studies, as well as implications for the design of optimised AdV-based vectors with improved and triple receptor binding functions, based on the properties of different fibers they might harbour. Still, we lack fundamental knowledge regarding the AdVs with two or more fiber genes, the practical advantage of these genes for the viruses and their properties that may enhance the mechanisms of infection. Future studies should reveal if all of the described fiber genes express functional fiber proteins that are efficiently incorporated into the adenoviral vertex and are able to bind the receptors on the host cells.

Funding

This work was supported by Hungarian Scientific Research Fund (grant numbers K100163 and NN140356); the EU grants ADVance (FP7-290002) and AD-VEC (FP7-324325); and the Tempus Public Foundation travel grant for IIP by the Hungarian Balassi Institute.

CRediT authorship contribution statement

Iva I. Podgorski: Methodology, Investigation, Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **Balázs Harrach:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Mária Benkő:** Conceptualization, Funding acquisition, Resources, Writing – review & editing. **Tibor Papp:** Methodology, Formal analysis, Visualization, Investigation, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors have no relevant financial or non-financial interests to disclose.

Data availability

The novel sequences of this study were deposited to GenBank under accession numbers MG969551 and KP329566. All other sequences analysed in the manuscript are available in GenBank and listed on page: https://sites.google.com/site/adenoseq/

Acknowledgements

The authors would like to gratefully thank the help of Dr. Jerome Custers at Janssen Vaccines, Janssen R&D, Johnson & Johnson, Leiden, The Netherlands. His scientific support as well as that of his research group colleagues, was very valuable during the researcher exchange programmes of the EU grants ADVance and AD-VEC.

References

Abbink, P., Maxfield, L.F., Nganga, D., Borducchi, E.N., Iampietro, M.J., Bricault, C.A., Teigler, J.E., Blackmore, S., Parenteau, L., Wagh, K., Handley, S.A., Zhao, G., Virgin, H.W., Korber, B., Barouch, D.H., 2015. Construction and evaluation of novel rhesus monkey adenovirus vaccine vectors. J. Virol. 89, 1512–1522. https://doi.org/10.1128/JVI.02950-14.

Abbink, P., Kirilova, M., Boyd, M., Mercado, N., Li, Z., Nityanandam, R., Nanayakkara, O., Peterson, R., Larocca, R.A., Aid, M., Tartaglia, L., Mutetwa, T., Blass, E., Jetton, D., Maxfield, L.F., Borducchi, E.N., Badamchi-Zadeh, A., Handley, S., Zhao, G., Virgin, H.W., Havenga, M.J., Barouch, D.H., 2018. Rapid cloning of novel rhesus adenoviral vaccine vectors. J. Virol. 92, https://doi.org/10.1128/JVI.01924-17 e01924–17.

Albinsson, B., Kidd, A.H., 1999. Adenovirus type 41 lacks an RGD α(v)-integrin binding motif on the penton base and undergoes delayed uptake in A549 cells. Virus Res. 64, 125–136. https://doi.org/10.1016/S0168-1702(99)00087-8.

Alonso-Padilla, J., Papp, T., Kaján, G.L., Benkő, M., Havenga, M., Lemckert, A., Harrach, B., Baker, A.H., 2015. Development of novel adenoviral vectors to overcome challenges observed with HAdV-5 based constructs. Mol. Ther. 24, 6–16. https://doi.org/10.1038/mt.2015.194.

Bai, M., Harfe, B., Freimuth, P., 1993. Mutations that alter an Arg-Gly-asp (RGD) sequence in the adenovirus type 2 penton base protein abolish its cell-rounding activity and delay virus reproduction in flat cells. J. Virol. 67, 5198–5205. https://doi.org/10.1128/JVI.67.9.5198-5205.1993.

Baker, A.T., Greenshields-Watson, A., Coughlan, L., Davies, J.A., Uusi-Kerttula, H., Cole, D.K., Rizkallah, P.J., Parker, A.L., 2019. Diversity within the adenovirus fiber knob hypervariable loops influences primary receptor interactions. Nat. Commun. 10, 741. https://doi.org/10.1038/s41467-019-08599-y.

Benkő, M., Aoki, K., Arnberg, N., Davison, A.J., Echavarría, M., Hess, M., Jones, M.S., Kaján, G.L., Kajon, A.E., Mittal, S.K., Podgorski, I.I., San Martín, C., Wadell, G., Watanabe, H., Harrach, B., 2022. ICTV virus taxonomy profile: *Adenoviridae* 2022. J. Gen. Virol. 103, 1721. https://doi.org/10.1099/jgv.0.001721.

Bergelson, J.M., Cunningham, J.A., Droguett, G., Kurt-Jones, E., Krithivas, A., Hong, J.S., Horwitz, M.S., Crowell, R.L., Finberg, R.W., 1997. Isolation of a common receptor for coxsackie B viruses and adenoviruses 2 and 5. Science. 275, 1320–1323. https://doi.org/10.1126/science.275.5304.1320.

Borkenhagen, L.K., Fieldhouse, J.K., Seto, D., Gray, G.C., 2019. Are adenoviruses zoonotic? A systematic review of the evidence. Emerg. Microbes Infect. 8, 1679–1687. https://doi.org/10.1080/22221751.2019.1690953.

Bots, S.T.F., Kemp, V., Dautzenberg, I.J.C., Hoeben, R.C., 2022. Genome analyses of ten new ape adenoviruses with similarity to human mastadenovirus C. Int. J. Mol. Sci. 23, 9832. https://doi.org/10.3390/ijms23179832.

Butynski, T., De Jong, Y.A., 2022. *Chlorocebus aethiops*. The IUCN red list of threatened species 2022: e.T4233A214886892. [WWW document].. https://doi.org/10.2305/IUCN.UK.2022-1.RLTS.T4233A214886892.en.

Chiu, C.Y., Yagi, S., Lu, X., Yu, G., Chen, E.C., Liu, M., Dick, E.J., Carey, K.D., Erdman, D.D., Leland, M.M., Patterson, J.L., 2013. A novel adenovirus species associated with an acute respiratory outbreak in a baboon colony and evidence of coincident human infection. MBio 4, e00084. https://doi.org/10.1128/mBio.00084-13.

Chroboczek, J., Ruigrok, R.W.H., Cusack, S., 1995. Adenovirus fiber. In: Doerfler, W., Böhm, P. (Eds.), The Molecular Repertoire of Adenoviruses. Springer, pp. 163–200. https://doi.org/10.1007/978-3-642-79496-4.

Descamps, D., Benihoud, K., 2009. Two key challenges for effective adenovirus-mediated liver gene therapy: innate immune responses and hepatocyte-specific transduction. Curr. Gene Ther. 9, 115–127. https://doi.org/10.2174/156652309787909544.

Doszpoly, A., Harrach, B., LaPatra, S., Benkő, M., 2019. Unconventional gene arrangement and content revealed by full genome analysis of the white sturgeon adenovirus, the single member of the genus *Ichtadenovirus*. Infect. Genet. Evol. 75, 103976 https://doi.org/10.1016/j.meegid.2019.103976.

Felsenstein, J., 1993. PHYLIP (Phylogenetic Inference Package, version 3.5c). Dep. Genet. SK, Univ, Washington, Seattle, WA. https://evolution.genetics.washington.edu/phylip.html.

Grgić, H., Krell, P.J., Nagy, É., 2014. Comparison of fiber gene sequences of inclusion body hepatitis (IBH) and non-IBH strains of serotype 8 and 11 fowl adenoviruses. Virus Genes 48, 74–80. https://doi.org/10.1007/s11262-013-0995-y.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704. https://doi.org/10.1080/10635150390235520.

Harrach, B., Tarján, Z.L., Benkő, M., 2019. Adenoviruses across the animal kingdom: a walk in the zoo. FEBS Lett. 593, 3660–3673. https://doi.org/10.1002/1873-3468.13687.

Hull, R.N., Minner, J.R., Mascoli, C.C., 1958. New viral agents recovered from tissue cultures of monkey kidney cells. III. Recovery of additional agents both from cultures of monkey tissues and directly from tissues and excreta. Am. J. Hyg. 68, 31–44. https://doi.org/10.1093/oxfordjournals.aje.a119947.

Jones, M.S., Harrach, B., Ganac, R.D., Gozum, M.M., Dela Cruz, W.P., Riedel, B., Pan, C., Delwart, E.L., Schnurr, D.P., 2007. New adenovirus species found in a patient presenting with gastroenteritis. J. Virol. 81, 5978–5984. https://doi.org/10.1128/JVI.02650-06.

Kaján, G.L., Stefancsik, R., Ursu, K., Palya, V., Benkő, M., 2010. The first complete genome sequence of a non-chicken aviadenovirus, proposed to be Turkey adenovirus 1. Virus Res. 153, 226–233. https://doi.org/10.1016/j.virusres.2010.08.006.

Kidd, A.H., Chroboczek, J., Cusack, S., Ruigrok, R.W., 1993. Adenovirus type 40 virions contain two distinct fibers. Virology 192, 73–84. https://doi.org/10.1006/viro.1993.1009.

Kidd, A.H., Garwicz, D., Oberg, M., 1995. Human and simian adenoviruses: phylogenetic inferences from analysis of VA RNA genes. Virology 207, 32–45. https://doi.org/10.1006/viro.1995.1049.

Kovács, G.M., Davison, A.J., Zakhartchouk, A.N., Harrach, B., 2004. Analysis of the first complete genome sequence of an Old World monkey adenovirus reveals a lineage distinct from the six human adenovirus species. J. Gen. Virol. 85, 2799–2807. https://doi.org/10.1099/vir.0.80225-0.

Kovács, 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. https://doi.org/10.1099/vir.0.80757-0.

Larsson, S., Svensson, C., Akusjärvi, G., 1986. Characterization of a low-molecular-weight virus-associated (VA) RNA encoded by simian adenovirus type 7 which functionally can substitute for adenovirus type 5 VA RNAI. J. Virol. 60, 635–644. https://doi.org/10.1128/JVI.60.2.635-644.1986.

Lasswitz, L., Chandra, N., Arnberg, N., Gerold, G., 2018. Glycomics and proteomics approaches to investigate early adenovirus–host cell interactions. J. Mol. Biol. 430, 1863–1882. https://doi.org/10.1016/j.jmb.2018.04.039.

Lenman, A., Liaci, A.M., Liu, Y., Årdahl, C., Rajan, A., Nilsson, E., Bradford, W., Kaeshammer, L., Jones, M.S., Frängsmyr, L., Feizi, T., Stehle, T., Arnberg, N., 2015. Human adenovirus 52 uses sialic acid-containing glycoproteins and the coxsackie and adenovirus receptor for binding to target cells. PLoS Pathog. 11, e1004657 https://doi.org/10.1371/journal.ppat.1004657.

Lenman, A., Manuel Liaci, A., Liu, Y., Frängsmyr, L., Frank, M., Blaum, B.S., Chai, W., Podgorski, I.I., Harrach, B., Benko, M., Feizi, T., Stehle, T., Arnberg, N., 2018. Polysialic acid is a cellular receptor for human adenovirus 52. Proc. Natl. Acad. Sci. U. S. A. 115, E4264–E4273. https://doi.org/10.1073/pnas.1716900115.

Li, X., Tikoo, S.K., 2002. Genetic organization and sequence analysis of pVIII, fiber and early region 4 of bovine adenovirus type 7. Virus Genes 25, 59–65. https://doi.org/10.1023/A:1020122208784.

Liu, T.C., Hallden, G., Wang, Y., Brooks, G., Francis, J., Lemoine, N., Kirn, D., 2004. An E1B-19 kDa gene deletion mutant adenovirus demonstrates tumor necrosis factor-enhanced cancer selectivity and enhanced oncolytic potency. Mol. Ther. 9, 786–803. https://doi.org/10.1016/j.ymthe.2004.03.017.

Lopez-Gordo, E., Podgorski, I.I., Downes, N., Alemany, R., 2014. Circumventing antivector immunity: potential use of nonhuman adenoviral vectors. Hum. Gene Ther. 25, 285–300. https://doi.org/10.1089/hum.2013.228.

Marek, A., Ballmann, M.Z., Kosiol, C., Harrach, B., Schlötterer, C., Hess, M., 2014a. Whole-genome sequences of two Turkey adenovirus types reveal the existence of two unknown lineages that merit the establishment of novel species within the genus *Aviadenovirus*. J. Gen. Virol. 95, 156–170. https://doi.org/10.1099/vir.0.057711-0.

Marek, A., Kaján, G.L., Kosiol, C., Harrach, B., Schlötterer, C., Hess, M., 2014b. Complete genome sequences of pigeon adenovirus 1 and duck adenovirus 2 extend the number of species within the genus *Aviadenovirus*. Virology 462–463, 107–114. https://doi.org/10.1016/j.virol.2014.04.033.

Marttila, M., Persson, D., Gustafsson, D., Liszewski, M.K., Atkinson, J.P., Wadell, G., Arnberg, N., 2005. CD46 is a cellular receptor for all species B adenoviruses except types 3 and 7. J. Virol. 79, 14429–14436. https://doi.org/10.1128/jvi.79.22.14429-14436.2005.

Mase, M., Nakamura, K., Imada, T., 2010. Characterization of fowl adenovirus serotype 4 isolated from chickens with hydropericardium syndrome based on analysis of the short fiber protein gene. J. Vet. Diagn. Investig. 22, 218–223. https://doi.org/10.1177/104063871002200207.

Medkour, H., Amona, I., Akiana, J., Davoust, B., Bitam, I., Levasseur, A., Tall, M.L., Diatta, G., Sokhna, C., Hernandez-Aguilar, R.A., Barciela, A., Gorsane, S., La Scola, B., Raoult, D., Fenollar, F., Mediannikov, O., 2020. Adenovirus infections in African humans and wild non-human primates: great diversity and cross-species transmission. Viruses 12, 657. https://doi.org/10.3390/v12060657.

Nilsson, E.C., Storm, R.J., Bauer, J., Johansson, S.M.C., Lookene, A., Ångström, J., Hedenström, M., Eriksson, T.L., Frängsmyr, L., Rinaldi, S., Willison, H.J., Domellöf, F.P., Stehle, T., Arnberg, N., 2011. The GD1a glycan is a cellular receptor for adenoviruses causing epidemic keratoconjunctivitis. Nat. Med. 17, 105–109. https://doi.org/10.1038/nm.2267.

Pantó, L., Podgorski, I.I., Jánoska, M., Márkó, O., Harrach, B., 2015. Taxonomy proposal for Old World monkey adenoviruses: characterisation of several non-human, non-ape primate adenovirus lineages. Arch. Virol. 160, 3165–3177. https://doi.org/10.1007/s00705-015-2575-z.

Pei, J., Kim, B.-H., Grishin, N.V., 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 36, 2295–2300. https://doi.org/10.1093/nar/gkn072.

Pénzes, J.J., Menéndez-Conejero, R., Condezo, G.N., Ball, I., Papp, T., Doszpoly, A., Paradela, A., Pérez-Berná, A.J., López-Sanz, M., Nguyen, T.H., van Raaij, M.J., Marschang, R.E., Harrach, B., Benkő, M., San Martín, C., 2014. Molecular characterization of a lizard adenovirus reveals the first atadenovirus with two fiber genes and the first adenovirus with either one short or three long fibers per penton. J. Virol. 88, 11304–11314. https://doi.org/10.1128/JVI.00306-14.

Perelman, P., Johnson, W.E., Roos, C., Seuánez, H.N., Horvath, J.E., Moreira, M.A.M., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., Schneider, M.P.C., Silva, A., O'Brien, S.J., Pecon-Slattery, J., 2011. A molecular phylogeny of living Primates. PLoS Genet. 7, e1001342 https://doi.org/10.1371/journal.pgen.1001342.

Persson, B.D., John, L., Rafie, K., Strebl, M., Frängsmyr, L., Ballmann, M.Z., Mindler, K., Havenga, M., Lemckert, A., Stehle, T., Carlson, L.-A., Arnberg, N., 2021. Human species D adenovirus hexon capsid protein mediates cell entry through a direct interaction with CD46. Proc. Natl. Acad. Sci. U. S. A. 118, e2020732118 https://doi.org/10.1073/pnas.2020732118.

Podgorski, I.I., Pantó, L., Papp, T., Harrach, B., Benkö, M., 2016. Genome analysis of four Old World monkey adenoviruses supports the proposed species classification of primate adenoviruses and reveals signs of possible homologous recombination. J. Gen. Virol. 97, 1604–1614. https://doi.org/10.1099/jgv.0.000465.

Podgorski, I.I., Pantó, L., Földes, K., Winter, I., Jánoska, M., Sós, E., Chenet, B., Harrach, B., Benkő, M., 2018. Adenoviruses of the most ancient primate lineages support the theory on virus−host co-evolution. Acta Vet. Hung. 66, 474–487. https://doi.org/10.1556/004.2018.042.

Rajan, A., Persson, B.D., Frängsmyr, L., Olofsson, A., Sandblad, L., Heino, J., Takada, Y., Mould, A.P., Schnapp, L.M., Gall, J., Arnberg, N., 2018. Enteric species F human adenoviruses use laminin-binding integrins as co-receptors for infection of Ht-29 cells. Sci. Rep. 8, 10019. https://doi.org/10.1038/s41598-018-28255-7.

Rapoza, N.P., 1967. A classification of simian adenoviruses based on hemagglutination. Am. J. Epidemiol. 86, 736–745. https://doi.org/10.1093/oxfordjournals.aje.a120781.

Roelvink, P.W., Lizonova, A., Lee, J.G., Li, Y., Bergelson, J.M., Finberg, R.W., Brough, D.E., Kovesdi, I., Wickham, T.J., 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. https://doi.org/10.1128/JVI.72.10.7909-7915.1998.

Roy, S., Vandenberghe, L.H., Kryazhimskiy, S., Grant, R., Calcedo, R., Yuan, X., Keough, M., Sandhu, A., Wang, Q., Medina-Jaszek, C.A., Plotkin, J.B., Wilson, J.M., 2009. Isolation and characterization of adenoviruses persistently shed from the gastrointestinal tract of non-human primates. PLoS Pathog. 5, e1000503 https://doi.org/10.1371/journal.ppat.1000503.

Roy, S., Clawson, D.S., Adam, V.S., Medina, A., Wilson, J.M., 2011. Construction of gene transfer vectors based on simian adenovirus 7. J. Gen. Virol. 92, 1749–1753. https://doi.org/10.1099/vir.0.032300-0.

Roy, S., Sandhu, A., Medina, A., Clawson, D.S., Wilson, J.M., 2012. Adenoviruses in fecal samples from asymptomatic rhesus macaques, United States. Emerg. Infect. Dis. 18, 1081–1088. https://doi.org/10.3201/eid1807.111665.

Sauthoff, H., Heitner, S., Rom, W.N., Hay, J.G., 2000. Deletion of the adenoviral E1b-19kD gene enhances tumor cell killing of a replicating adenoviral vector. Hum. Gene Ther. 10, 379–388. https://doi.org/10.1089/10430340050015851.

Seiradake, E., Cusack, S., 2005. Crystal structure of enteric adenovirus serotype 41 short fiber head. J. Virol. 79, 14088–14094. https://doi.org/10.1128/jvi.79.22.14088-14094.2005.

Seiradake, E., Lortat-Jacob, H., Billet, O., Kremer, E.J., Cusack, S., 2006. Structural and mutational analysis of human Ad37 and canine adenovirus 2 fiber heads in complex with the D1 domain of coxsackie and adenovirus receptor. J. Biol. Chem. 281, 33704–33716. https://doi.org/10.1074/jbc.M605316200.

Singh, M., Kumar, A., Kumara, H.N., 2020. *Macaca mulatta*. The IUCN red list of threatened species 2020: e.T12554A17950825. [WWW document]. https://doi.org/10.2305/IUCN.UK.2020-2.RLTS.T12554A17950825.en.

Stasiak, A.C., Stehle, T., 2020. Human adenovirus binding to host cell receptors: a structural view. Med. Microbiol. Immunol. 209, 325–333. https://doi.org/10.1007/s00430-019-00645-2.

Tarján, Z.L., Pénzes, J.J., Tóth, R.P., Benkő, M., 2014. First detection of circovirus-like sequences in amphibians and novel putative circoviruses in fishes. Acta Vet. Hung. 62, 134–144. https://doi.org/10.1556/AVet.2013.061.

To, K.K., Tse, H., Chan, W.M., Choi, G.K., Zhang, A.J., Sridhar, S., Wong, S.C., Chan, J.F., Chan, A.S., Woo, P.C., Lau, S.K., Lo, J.Y., Chan, K.H., Cheng, V.C., Yuen, K.Y., 2014. A novel psittacine adenovirus identified during an outbreak of avian chlamydiosis and human psittacosis: zoonosis associated with virus-bacterium coinfection in birds. PLoS Negl. Trop. Dis. 8, e3318 https://doi.org/10.1371/journal.pntd.0003318.

Tollefson, A.E., Ying, B., Doronin, K., Sidor, P.D., Wold, W.S., 2007. Identification of a new human adenovirus protein encoded by a novel late l-strand transcription unit. J. Virol. 81, 12918–12926. https://doi.org/10.1128/JVI.01531-07.

Tomko, R.P., Xu, R., Philipson, L., 1997. HCAR and MCAR: the human and mouse cellular receptors for subgroup C adenoviruses and group B coxsackieviruses. Proc. Natl. Acad. Sci. U. S. A. 94, 3352–3356. https://doi.org/10.1073/pnas.94.7.3352.

Van Raaij, M.J., Mitrakl, A., Lavigne, G., Cusack, S., 1999. A triple β-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. Nature 401, 935–938. https://doi.org/10.1038/44880.

Wang, H., Li, Z.Y., Liu, Y., Persson, J., Beyer, I., Möller, T., Koyuncu, D., Drescher, M.R., Strauss, R., Zhang, X.B., Wahl, J.K., Urban, N., Drescher, C., Hemminki, A., Fender, P., Lieber, A., 2011. Desmoglein 2 is a receptor for adenovirus serotypes 3, 7, 11 and 14. Nat. Med. 17, 96–104. https://doi.org/10.1038/nm.2270.

Wickham, T.J., Segal, D.M., Roelvink, P.W., Carrion, M.E., Lizonova, A., Lee, G.M., Kovesdi, I., 1996. Targeted adenovirus gene transfer to endothelial and smooth muscle cells by using bispecific antibodies. J. Virol. 70, 6831–6838. https://doi.org/10.1128/jvi.70.10.6831-6838.1996.

Wu, E., Pache, L., Von Seggern, D.J., Mullen, T.M., Mikyas, Y., Stewart, P.L., Nemerow, G.R., 2003. Flexibility of the adenovirus fiber is required for efficient receptor interaction. J. Virol. 77, 7225–7235. https://doi.org/10.1128/JVI.77.13.7225-7235.2003.

Xia, D., Henry, L.J., Gerard, R.D., Deisenhofer, J., 1994. Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. Structure 2, 1259–1270. https://doi.org/10.1016/S0969-2126(94)00126-X.

Zeng, Z., Zhang, J., Jing, S., Cheng, Z., Bofill-Mas, S., de Motes, C., Hundesa, A., Girones, R., Seto, D., Zhang, Q., 2016. Genome sequence of a cynomolgus macaque adenovirus (CynAdV-1) isolate from a primate colony in the United Kingdom. Genome Announc. 4 https://doi.org/10.1128/genomeA.01193-16 e01193–16.

Zhao, J., Zhong, Q., Zhao, Y., Hu, Y.X., Zhang, G.Z., 2015. Pathogenicity and complete genome characterization of fowl adenoviruses isolated from chickens associated with inclusion body hepatitis and hydropericardium syndrome in China. PLoS One 10, e0133073. https://doi.org/10.1371/journal.pone.0133073.

Web references

https://sites.google.com/site/adenoseq, 14.07.2022.
https://ictv.global/report/adenoviridae, 14.07.2022.