

Szent István University
Postgraduate School of Veterinary Science

Molecular characterisation of simian adenoviruses

Brief Summary of Doctoral Thesis

Iva Podgorski

2016

Szent István University

Postgraduate School of Veterinary Science

Supervisors and members of the supervisory board:

Prof. Dr. Mária Benkő
supervisor
Institute for Veterinary Medical Research
Centre for Agricultural Research
Hungarian Academy of Sciences

Prof. Dr. Balázs Harrach
supervisor
Institute for Veterinary Medical Research
Centre for Agricultural Research
Hungarian Academy of Sciences

Dr. Győző Kaján
consultant
Institute for Veterinary Medical Research
Centre for Agricultural Research
Hungarian Academy of Sciences

Introduction and objectives

Adenoviruses (AdVs) are medium-sized (70-90 nm), non-enveloped viruses which were isolated more than 60 years ago from human adenoid tissue, hence the name. Most of the AdVs are apathogenic in healthy individuals, but some can cause disease and in rare cases even death. Simian AdVs (SAdVs) have been found to be associated with several diseases in primates, including diarrhoea, pneumoenteritis, conjunctivitis, and hepatitis, and some of them have been reported to induce tumours when injected into neonate rodents.

Based on the molecular phylogeny, primates can be grouped into two big suborders, Strepsirrhini (containing the prosimians) and Haplorrhini (containing New World monkeys (NWMs), Old World monkeys (OWMs), gibbons, great apes, and humans). In the Haplorrhini suborder, more than 150 AdV types are known, albeit with an overwhelming majority of ape and human AdVs (HAdV) compared to the more ancient primates (NWMs and prosimians) with very limited or no information about the prevalence, evolution, and genome characteristics of their AdVs. Since the human population is widely infected by AdVs and the existing specific antibodies significantly limit the medical applicability of HAdVs, there is an increased interest in the possible use of non human, especially SAdVs.

Ape AdVs, isolated from chimpanzees, bonobos and gorillas, have been proposed to be members of species *Human mastadenovirus B* (HAdV-B), *C* and *E*, respectively. Additional recently described ape AdVs showed to be the closest to HAdV-A, -D, or -F. Species *Simian mastadenovirus A* was, until this study, the only species officially approved for OWM AdVs exclusively. Our knowledge concerning the NWM and prosimian AdVs is even more limited. The only fully sequenced NWM AdV is the titi monkey AdV (TMAdV), whilst there have not been any AdVs reported in prosimians as yet.

In this study the aims were to find and characterize novel AdVs in prosimians, NWMs, OWMs, orangutans and gibbons, and to get more data about the OWM AdVs isolated previously, with the main focus on the host specificity, virus-host evolution, AdV variability, and differences between them on molecular level (e.g., novel genes, spliced genes, receptors of the fibre genes). We were interested also in the amelioration of the taxonomy of the family *Adenoviridae* by establishing correct novel species for the newly characterized primate AdVs.

Materials and methods

Viruses and DNA extraction

OWM SAdV strains (SAdV-1 to 20) deposited in the American Type Culture Collection (ATCC) were studied by PCR, DNA sequencing and/or bioinformatics. For the purpose of the next generation sequencing (NGS), Vero cells E6 were infected with all the non-sequenced OWM SAdV strains. The DNA of SAdV-2, -8, -11, and -17 was extracted from the concentrated virions by phenol-chloroform extraction method. For the receptor binding studies, A549 cells were infected with SAdV-1, -2, -11 and HAdV-5, as a positive control. The virions were subjected to purification with CsCl gradient. To find novel AdVs, a total of 138 fecal or organ samples (originating from three Hungarian zoos, a Croatian zoo, a French zoo, a French university, and Madagascar) were screened: 10 from apes, 11 from OWMs, 19 from NWMs, and 98 from prosimians. DNA was extracted from fecal samples with the E.Z.N.A.® Stool DNA Kit (OMEGA bio-tek) according to the manufacturers' instructions, and from the organ samples with a house protocol.

PCR and DNA sequencing

Shorter or longer fragments of the genes of four well-conserved adenoviral proteins (namely IVa2, DNA-dependent DNA polymerase (*pol*), penton base and hexon) were obtained by PCR from 14 SAdVs. Later on, some of these AdVs were fully sequenced by NGS (SAdV-2, -8, -11 and -17; Illumina HiSeq2500 system) or traditional methods (SAdV-13 (mostly sequenced by a colleague), -16 and -19; capillary electrophoresis). To detect novel AdVs, nested PCR targeting the IVa2 gene of mastadenoviruses was used. One of the NWM AdVs was subjected to sequencing of longer fragments of the genome by traditional methods.

Cell culture methods and end-point dilution assay

To isolate newly detected AdVs, positive samples were prepared to infect the cmt93, A549, HEK293, 3T6, Vero E6, CHO-K1 and MFC cells. Blind passages were repeated for up to seven times, after which the suspension was subjected to PCR with IVa2 primers. Three OWM AdVs (SAdV-5, -12 and -15) were not clean types, therefore were subjected to end-point dilution assay on Vero E6 cells, to isolate at least one type from each mixture. PCR with penton base targeting primers determined if the virus type was clean or not. If clean, the virus was produced in larger amounts and was sent to a commercial

NGS service (BGI, China). If not clean, it was sent to a partner company in Netherlands (Batavia) where it was subjected to plaque purification.

Molecular cloning of fibre-1 knobs, fluorescence activated cell sorting (FACS)

Knobs of the fibre-1 genes of SAdV-1, -7, -11 and -19 were cloned into the pQE-30 Xa plasmid to express the knobs and study the cellular receptors they can attach to. Niklas Amberg's group (Umea University, Sweden) expressed the knobs and produced them in larger amounts. To investigate the ability of these fibre knobs to use the sialic acid-containing glycans as receptors, A549 cells treated with neuraminidase (cleaves polysialic acids) from *V. cholera* were used. Primary RGS-His, mouse IgG and secondary Polyclonal Rabbit Anti-mouse Immunoglobulin/FITC antibody were used, and the binding was measured on FACS LSR II machine (Becton Dickinson).

Bioinformatics

The identities of the newly gained sequences were examined using the BLASTX program (NCBI). Primate AdV sequence alignments were prepared with the ClustalW program of the MEGA6 package. Phylogenetic calculations were performed using the ProtDist and PhyML algorithms (Mobyli portal), and the ProtTest program. The ProtDist analyses were run with the JTT substitution model, followed by Fitch-Margoliash analysis, applying the global rearrangements option. PhyML calculations were based on a user tree obtained by ProtDist and a substitution model determined by ProtTest. Phylogenetic analysis, based on nt alignment of partial penton base gene sequences, were performed using PhyML in TOPALi v2 platform. Bootstrap analysis with 100 sampling replicates was applied for every tree. The trees were visualized using the MEGA6 program. Similarity plots and bootscanning analyses were performed with Simplot 3.5.1 with window size 1000 bp, step size 50 bp.

Results

ATCC strains

We were able to grow 12 of the 14 non-sequenced SAdVs on the Vero E6 cells, and eight of them (SAdV-2, -4, -5, -8, -11, -12, -15 and -17) were produced in large amounts and their DNA was extracted for the NGS. However, in case of SAdV-4, a parvovirus was in majority, and three of the viruses (SAdV-5, -12 and -15) were mixtures of at least two types, therefore the full genome sequence of these SAdVs was not obtained in this study. Partial sequencing of SAdV-12 and -15 indicated the presence of two different types in each case, belonging to the species HAdV-G, whereas the two types of SAdV-5 were found to be members of different species, HAdV-G and SAdV-B. The end-point dilution assay on the Vero E6 cells was unsuccessful in obtaining clean types. For SAdV-5 we got partial sequence (about half of the genome) of a clean type, but it was exactly the same as SAdV strain A1335 (GenBank acc. number JN880456) from species SAdV-B. The three mixtures were sent to a partner company in Netherlands (Batavia) where they were subjected to plaque purification. The phylogeny reconstructions supported existence of previously established HAdV-G and SAdV-A, and previously proposed SAdV-B and SAdV-C, and also indicated existence of five additional OWM AdV species, SAdV-D to SAdV-H. Four SAdVs (SAdV-2, -8, -11 and -17) were successfully sequenced by NGS, and two (SAdV-16 and -19) with the traditional Sanger sequencing. Genomic assemblies revealed that studied genomes contain 36 (SAdV-8, -11, -16, -19) or 37 (SAdV-2, -17) putative coding regions characteristic for mastadenoviruses, including two or three genes predicted to code for the fibre of AdVs, respectively. SAdV-2 is an exception as it does not contain E1A 19K gene, but has an extra copy of ITR and E4 ORF1 gene in this region instead. A single copy of VA-RNA gene and three exons of the U exon protein (UXP) were predicted for all the fully sequenced SAdVs. Phylogeny inference based on the full hexon aa sequences pointed out divergent relationships among several AdVs, which were confirmed by the SimPlot and BootScan analyses, indicating recombination event(s) in the hexon gene of SAdV-19. In the receptor-binding studies, pre-treatment of the A549 cells with neuraminidase reduced the binding of the fibre-1 knobs of positive controls and studied SAdVs to background levels.

Novel adenoviruses

The PCRs resulted in an approximately 300 bp amplicons from 23 samples (16.6%). Nucleotide and aa sequences of the amplified IVa2 region showed that some prosimian AdVs are identical with each other, as well as some NWM AdVs. There were eight different AdVs detected in prosimians, seven in NWMs (10 in prosimians and 15 in NWMs

if we include those detected by other colleagues), one in OWMs, and two in apes. Unfortunately, none of the viruses propagated on any of the cells used. Maximum likelihood analysis of partial IVa2 aa sequences (84 aa) separated novel AdVs from the earlier established species of primate AdVs. Two new primate AdV lineages appeared: prosimian AdVs as the most ancient ones of primate AdVs, and NWM AdVs, grouped together with previously described TMAV. Additionally, we sequenced 25,885 bp of the red-handed tamarin (RHT) AdV-1 genome, ranging from the E1B 55K to the UXP gene. The sequenced region shared genetic composition characteristic for mastadenoviruses, with difference seen in the E3 region: two ORFs in this region represent novel genes, not seen in any other AdV before. Phylogeny inference placed the RHTAdV-1 among other NWM AdVs, as expected.

Discussion

ATCC strains

Although an OWM cell line (Vero E6) was used for the production of OWM AdVs for NGS, some of the strains (SAdV-10 and -13) did not grow, which may have been the consequence of having not living strains. Adeno-associated virus (AAV) DNA detected in some cases is not surprising, especially as most of the non-human primate AAVs have been isolated as contaminants of AdV preparations. The receptor binding studies indicated that the fibre-1 knob of the studied SAdVs uses sialic acid-containing glycans as the receptor on A549 cells. Since data about the binding of fibre-2 of SAdVs is still missing, an interesting study might be investigation of binding properties of both fibre-1 and -2 from all available SAdVs, comparison to HAdVs and defining the possibilities of using them in gene therapy approaches.

All phylogenetic calculations, except the one based on full hexon aa sequence, indicate that nine of the studied SAdVs belong either to the species SAdV-A (SAdV-4, -9, -10 and -14) or HAdV-G (SAdV-2, -5 (one of the types in the mixture), -11, -12 (both types) and -15 (both types)), established earlier. The third most numerous cluster is that of SAdV-5, -8, -49, -50, nine other rhesus macaque isolates and BaAdV-1, recently officially accepted by the Executive Committee of ICTV as SAdV-B. SAdV-19 proved to be a member of recently officially accepted species SAdV-C. SAdV-13, -16 and -20 turned out to be the only members of a separate lineages suggested as new species, SAdV-D, -E and -G, respectively. SAdV-17 and -18 were proposed to form a new species, SAdV-F. SAdV-54 strain 23336, described recently, was proposed to be a member of a new species SAdV-H. The G+C content, hemagglutination-inhibition test results, host origin, phylogenetic distance, genome organization differences, and number of VA-RNA genes (one in OWM SAdVs vs. two in human and ape AdVs) supported the proposed species classification. All newly sequenced OWM SAdVs have 36 or 37 putative genes characteristic for the genus *Mastadenovirus*. For the first time we detected three fibre genes in AdVs (in SAdV-2 and -17), which gives great potential to these AdVs for vectorizing purpose, since the fibres can be switched to use certain receptors on the host cells. Sequence alignments (aa) of the UXPs predicted in SAdVs and described in HAdVs from HAdV-C showed that the first and the second exons are relatively well conserved, even in different species of AdVs, whilst the third exon is more variable both in sequence content and length. Genomic localisation of all the three exons was comparable (the third exon starts at the same upstream position and is always in the same reading frame relative to the DBP-encoding region), which might be of help in defining the UXP sequences in other AdVs in the future. SAdV-2 has deletion/repetition never seen in other

AdVs before: repetition of the ITR and E4 ORF1 in place of the E1B 19K gene. Since 19K deletion mutant AdVs demonstrated enhanced tumour cell killing and enhanced oncolytic potency, there seems to be a promising potential of the virus with three fibre genes and a natural deletion of 19K gene, especially if the goal of the vector application is targeted host cell destruction. The hexon-based tree shows different relationship among several simian and HAdVs which might be the result of recombination, supported by the SimPlot analysis of SAdV-19, indicating that its hexon gene resulted from a recombination event between SAdV-19 and a yet unknown member of HAdV-F. It is clear now that SAdVs cluster as well as HAdVs do, and that the future may bring more primate AdV species for OWMs exclusively, as well as for more ancient AdVs, i.e. for those of NWMs and prosimians.

Novel adenoviruses

In this study we detected novel AdVs in apes, OWMs, NWMs and prosimians. IVa2 nested PCR was found to detect mastadenoviruses very effectively, and for the primate AdVs it proved to be even more sensitive than the *pol* PCR. Although the samples were collected from captive and wild animals from different parts of Europe and Madagascar, we found the same AdV (sometimes on aa level, sometimes even on nt level) in several hosts from different areas, and even in different species. Anyway, all the cases where different species harbor the same AdV are not unexpected since these species are evolutionarily very close.

NWM and prosimians AdVs clustered into two new lineages of primate AdVs. In both lineages we can see several clusters, which we compared to the host clusters, and in most cases they correspond. The phylogenetic place of TMAAdV together with tufted capuchin and common squirrel monkey AdVs is unexpected since its host is evolutionarily quite far from them. Anyway, the case of high fatality rate from the TMAAdV outbreak (83%) indicates that titi monkey is not the native host for TMAAdV. Furthermore, it was already seen that cross-species transmission can result in a pathogenicity usually not seen in the original host. This indicates that AdVs can indeed cross the species barriers among different nonhuman and human primate hosts, but sometimes even between primate and non-primate hosts. Ape AdVs appeared on the phylogenetic tree at the expected place: Sumatran orangutan AdV among the HAdV species, closest to the HAdV-C, to which many ape AdVs were proposed to belong to (chimpanzee, bonobo and gorilla AdVs), whereas siamang AdV, the first representative of gibbon AdVs, found its place among the OWM AdV species (nevertheless, it stands alone, as expected, although the position closer to the ape and HAdVs would be more expected). All the positive samples were used in the AdV isolation trials on several different cell lines originating from human

(A549, HEK293), monkey (Vero E6, MFC), hamster (CHO-K1) or mouse (3T6, cmt93) tissues. However, despite multiple serial blind passages, isolation was unsuccessful, most probably because of the incompatibility of these cell lines with the viruses we are trying to isolate. However, the OWM and ape samples were expected to propagate on Vero and/or human cell lines, but the problem here could be the very low amount of the virions in the fecal samples. Although the isolation of the RHTAdV-1 was unsuccessful, we sequenced the majority of its genome with the traditional (Sanger) method. Most of the genes are shorter than those of the TMAAdV, a few are similar, and only one is considerably longer (RID- α in the E3 region). The identities of the genes it shares with the TMAAdV range from 30 to 82%. There are two novel genes in the E3 region which do not represent additional genes, but rather a “replacement” of some genes seen in other SAdVs. In fact, it seems that one gene is missing, since this virus has five genes in the E3 region, instead of the usual six. Majority of the novel AdVs detected confirms the theory on the virus-host co-evolution and gives us a broad picture about primate AdVs, representing other lineages beside the previously seen human, ape and OWM AdVs. Additionally, we proposed a novel species for the TMAAdV to the ICTV, putatively named *Platyrrhini mastadenovirus A*.

New scientific results

1. Partial genome sequencing of the 14 non-sequenced, serotyped ATCC Old World monkey AdVs enabled their assignment to the previously established (*Human mastadenovirus G* and *Simian mastadenovirus A*) and recently proposed (*Simian mastadenovirus B* and *C*) or newly proposed (*Simian mastadenovirus D* to *G*) AdV species.
2. Six Old World monkey AdVs (SAdV-2, -8, -11, -16, -17 and -19) fully sequenced and molecularly characterized. For the first time, three fibre genes detected in AdVs.
3. For the first time in SAdVs, all the three exons of the so called U exon protein characterized.
4. An almost fully sequenced New World monkey AdV confirms the difference of the E3 region of Old World and New World monkey AdVs.
5. Achievement of the official establishment of two SAdV species, *Simian mastadenovirus B* and *Simian mastadenovirus C* in the ICTV. Proposal to ICTV for six SAdV species, *Simian mastadenovirus D* to *H*, and *Platyrrhini mastadenovirus A*.
6. First receptor binding studies with SAdVs show that fibre-1 knob of members of species *Human mastadenovirus G*, as well as that of SAdV-19, binds the sialic acid-containing glycans.
7. Novel AdVs detected in apes, Old World monkeys, New World monkeys, and for the first time in prosimians support the theory on virus–host co-evolution.

Publications and conference abstracts

Publications

1. 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. IF: 3.755

2. 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. IF: 2.390

*L. Panto and I. I. Podgorski contributed equally to this work.

3. Podgorski, I.I., Pantó, L., Papp, T., Harrach, B. & Benkő, 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*. Published online: 24/03/2016 (in press). DOI: 10.1099/jgv.0.000465. IF: 3.183

Conference abstracts

1. Podgorski, I. I., Pantó, L., Benkő, M. & Harrach, B. (2012). Genomic and bioinformatics analysis of simian adenovirus 19 confirms the need to establish a new adenovirus species. Annual meeting of the Hungarian Society for Microbiology, Keszthely, Hungary.

2. Podgorski, I. I., Pantó, L., Benkő, M. & Harrach, B. (2014). Genomic and bioinformatics analysis of simian adenovirus 19 confirms the need to establish a new adenovirus species. XVI. International Congress of Virology, Montreal, Canada.

3. Podgorski, I. I., Benkő, M. & Harrach, B. (2014). Adenoviruses – good or bad fellows? 1st Croatian Virus Workshop, Basic and Translational Virus Research, Rijeka, Croatia.

4. Podgorski, I. I., Pantó, L., Harrach, B. & Benkő, M. (2015). Characterization of novel non-human primate adenoviruses with promising future in human medicine. BSGCT Annual Conference, Glasgow, UK.

5. Podgorski, I. I., Pantó, L., Jánoska, M., Márkó, O. & Harrach, B. (2015). Taxonomy proposal for Old World monkey adenoviruses existence of several non-human, non-ape

primate adenovirus lineages. 17th International Congress of the Hungarian Society for Microbiology, Budapest, Hungary.

6. **Harrach B. & Podgorski I.I. (2014).** Jó vírus -- rossz vírus? (Gyógyító vírusok), In: Szerk.: Janda T III. ATK Tudományos Nap: A jövőt alakító Európai Unió agrárkutatói programok az MTA ATK-ban. Martonvásár: MTA ATK, 30-35.

Acknowledgements

The very first people I would mention here are dr. Balázs Harrach and dr. Mária Benkő, because without their guidance and support I wouldn't be in this life period. I've met many supervisors in the last few years, and learned indubitably that Balázs and Mari are the most encouraging and supporting people whose help was priceless in my early scientific career. I appreciate all the advices I was getting from them for both the private and professional life, making the living and studying in a foreign country much easier.

Next to them stand the people who are under their guidance in the lab, and who helped me with both the work and the fun times in Budapest and abroad. All these people taught me most of the things I know today about the lab work and I am very grateful for that. Special thanks to dr. Tibor Papp who provided great help vitally for the start of the living in Hungary, and also for the help in the projects I've been working on. Thanks to Mónika Ballmann who was my teacher in the first months and helped me to start with my projects. Thanks to Laura Pantó who worked with the simian AdVs prior to my arrival to the lab, and for giving me great basis to continue working on. Thanks to dr. Győző Kaján for his expertise in any problem I had during my stay in Budapest and Sweden. Thanks to Zoltán László Tarján, dr. Andor Doszpoly and dr. Judit Péntzes, my temporary "lab roommates" who were always there to help with my Hungarian language issues. Special thanks to Judit for the fun times we had listening to her concerts. I also thank to dr. Márton Vidovszky and Kinga Böszörményi for making my second part of the stay in Budapest and abroad in very friendly environment, with a lot of jokes and fun times! It was a pleasure to supervise someone as adroit and friendly as Kinga! And thanks to Marci for all the help he provided in preparation of this thesis. I would also like to thank to dr. Edit Eszterbauer for the great help she provided for my private life in Hungary, being an interpreter at any time I needed. And finally, thanks to all the other people of the Institute who made my working days in Hungary easier with just a few nice words or smiles while meeting in the halls.

I also thank to all the people who provided the samples essential for my work, namely dr. Viktor Molnár and dr. Endre Sós from Hungarian zoo, dr. Baptiste Chenet from French zoo, dr. Nadine Mestre-Francés from French University, Maša Ljuština from Croatian zoo, and Katalin Földes for help with the collection of samples. Special thanks to Iris de Winter from Wageningen University for providing the samples from Madagascar. Thanks also to Niklas Arnberg and his group for hosting me in their lab for three weeks, as well as to Eric Kremer and his group for the help during my secondment in Montpellier. And I would also like to thank to the whole ADVance team, who made it possible to work on the project together.

Enormous thanks and gratitude to my whole family, especially to my mom and dad who enabled me to become a person I am. Without their guidance and support I would not be here today. Thanks for always being there for anything I need, for selfless support of every idea and wish I have, and for making my life as fulfilled and happy. The person who I am thankful the most in another way is Luka, who generously and indisputably followed me everywhere, supporting my first steps in the scientific career and being a person I can count on in every situation.

Financial supports were granted by the EU grant ADVance FP7-290002, by the Hungarian Scientific Research Funds NN107632 and K100163, and the Balassi Institute (Tempus Public Foundation).