Szent István University

Postgraduate School of Veterinary Science

Genetic analysis of adenoviruses from rodents and bats

Brief Summary of Doctoral Thesis

Márton Vidovszky

2015

Szent István University
Postgraduate School of Veterinary Science

Supervisor and members of the supervisory board:

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

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

Introduction and Objectives

Adenoviruses (AdVs) can cause serious and lethal diseases in context of veterinary (eg. canine AdV-1). This may be due to the host switch of the AdVs, during which the AdVs have not had enough time to adapt to the new host. Adenoviruses are usually host specific, facultative pathogens, and they can be detected in representatives of all major families of vertebrates. AdVs are non-enveloped, medium sized (70–90 nm) viruses, which have icosahedral shaped virions and double stranded DNA.

Prior to this research, only the genome sequence of murine adenovirus 1 (MAdV-1) was known among the rodent AdVs. This virus causes encephalomyelitis in newborn animals of some mouse strains. In contrast, the long term identified, but not sequenced, MAdV-2 causes milder diseases, such as diarrhea. Differences between the genomes of the two AdVs had been presupposed by the restriction patterns. One objective of the research detailed in this thesis was to carry out a full molecular analysis of the MAdV-2 genome, as the rodent AdVs can be good models for the potential application of AdVs in human medicine. While research began on the genome sequencing of MAdV-2, a new MAdV type (MAdV-3) had been isolated from striped field mouse (*Apodemus agrarius*) in Slovakia. The genome of MAdV-3 has been sequenced and annotated. Until now, only these three MAdV genomes have been characterised.

A new AdV from red squirrels (*Sciurus vulgaris*) with diarrheal symptoms has been described in collaboration with researchers from Hungary. The squirrels originated from north-east England. Subsequently, more cases were described in Cumbria, in the Island of Anglesey of Wales, in Scotland and in further provinces of Great Britain. Some AdV virions were found by electron microscope in the intestinal contents of the animals. Despite initial observations of signs of isolation on murine cell line, this strain could not be propagated further than one passage. A short part of the hexon gene of the “isolated” virus has been sequenced. Phylogenetic calculations showed that the squirrel AdV-1 (SqAdV-1) is a mastadenovirus, which is well isolated from known AdVs and represent a new species. In collaboration with local researchers in Germany, we were the first detecting AdV in squirrels outside Great-Britain. At the first attempts the isolation of the virus seemed to be successful, thus we started to sequence its genome. Furthermore, a screening study for AdVs in other species of rodents was initiated to explore novel potential AdV models.

Bats are popular subjects of the virus researches, because it is proved that they are carriers of several very interesting viruses. This research represents a comprehensive screening of AdVs in the Hungarian bat fauna, covering all the occurring species. Before this research with the bat AdVs, only one AdV isolation (BtAdV-1) was known from a Ryukyu flying fox (*Pteropus dasymallus*), isolated in Japan. As the screening study was initiated, our

laboratory was asked for collaboration in molecular characterization of a German bat AdV isolation. This virus (BtAdV-2) was isolated from a common pipistrelle (*Pipistrellus pipistrellus*). Since the genome analysis of BtAdV-2, two more bat AdVs were isolated in China (from *Myotis ricketti* – BtAdV-3) and India (from Leschenault's rousette, *Rousettus leschenaultii* – BtAdV-4). Currently only the BtAdV-3 genome has been determined beside BtAdV-2.

Materials and Methods

Origin of the samples

During the screening of both rodent and bat samples, organ homogenates and fecal samples were examined too. Most of the 64 rodent samples were collected by our colleagues or originated from the Budapest Zoo. These samples belong to 6 families of 3 suborders (Sciuromorpha, Myomorpha and Hystricomorpha) from the order Rodentia.

Almost all of the 165 Hungarian bat samples were collected fecal samples, while all the 194 bat samples from Germany were organ samples. The bat sample collection represents all 28 bat species present within Hungary and belonging to 3 families from both bat suborders of Chiroptera.

MAdV-2, BtAdV-2 and SqAdV-1, studied on the full genome level, were obtained from foreign collaborators. From these AdVs, only MAdV-2 and BtAdV-2 were isolated viruses.

Polymerase Chain Reaction (PCR), genome analysis

The general Wellehan PCR (2004) method was used to detect the AdV DNA. This method is perfect for a random screening, but due to the occasionally presence of nonspecific results, sequencing of the PCR product is unavoidable. Previous experience has shown that this PCR method is suitable to detect the members of all five accepted AdV genera.

Additional (nested or simple) PCRs were carried out to amplify more and longer gene fragments from MAdV-2 and SqAdV-1. In general, these PCRs could be used to amplify sequence regions only from mastadenoviruses. Nested PCRs were used to amplify conserved regions from the genes IVa2 and pVIII, while simple PCRs were used for the fragments of the hexon and the penton base. To design consensus primers, alignments were made from the amino acid (aa) sequences of homologue genes of closely related AdVs. For designing primers, strongly conserved aa motives were selected where the aa have as few as possible codon variations (1–2). In the ordered oligonucleotides, all the possible codon variations were present.

The genome fragments, amplified by the common primers, were connected with PCRs using specific primers. The longer genome fragments were sequenced by primer walking.

The sequencing of the MAdV-2 was made in Swiss collaboration, while sequencing of the SqAdV-1 was made exclusively in our laboratory. In case of the BtAdV-2, our work was the genome analysis and annotation.

Sequence analysis

The BioEdit program was used to visualize and inspect the quality of the sequence data. Incidental mistakes were corrected with this program as well. The Gap4 program of the Staden sequence analyzer software package was used to join the parts of the full or partial genome sequences.

The origin of the newly detected AdV sequences was identified mainly by the BLASTX program on the homepage of NCBI, comparing to the sequences of GenBank database. The sequences were also compared to our database, containing our unpublished AdV sequences, by using the BLASTX program of BioEdit software package.

Full genome sequences of MAdV-2 and BtAdV-2, and the partial genome sequence of SqAdV-1, were examined with the JavaScript DNA Translator 1.1 program. The ORFs, found by this program were presumed to be genes depending on their size and location on the genome and on their homology with other AdV genes. The potential splicing donor and acceptor sites were determined by individual examinations, looking for the consensus signals. Alignments were made with MultAlin and Genedoc programs.

Three different methods were used for phylogenetic calculations. Short gene sequences were analyzed using the protein distance matrix method as the more sophisticated likelihood methods proved to be unreliable due to the lack of sufficient data. For these calculations, the Protdist program of the PHYLIP software package and subsequently the Fitch program with global rearrangement were used. The Fitch program tests the position of each branch of the phylogenetic tree, which was obtained from the first calculation.

Long gene sequences or full genes were analyzed using the maximum likelihood and Bayesian methods. These calculations were made with the PhyML and MrBayes programs of the Topali 2.5 software package. To generate the best result, a preliminary model selection was run. For analyzing full protein sequences, bootstrap calculations were performed as well. This calculation repeats the analysis with samples randomly taken from the alignments, and makes a statistical comparison from the results at the end. The bootstrap analysis was made usually with 100 sampling. The distance analyses by the maximum likelihood calculation with bootstrap analysis were made on the Mobyle internet portal of the Pasteur Institute of Paris. For model selection, the Topali 2.5 program package and the ProtTest program were used. Phylogenetic trees were visualized and edited with FigTree v1.4 program.

Results

Novel adenoviruses

Seventeen (3 organ and 14 fecal samples) of the 64 (26 organ and 38 fecal) rodent samples proved to be PCR positive, resulting in a 26.56% positivity ratio. These 17 samples belonged to 3 species and 4 different AdVs (3 novel) were detected in them. Novel AdV sequences were detected in organ samples of the common vole (*Microtus arvalis*) and the capybara (*Hydrochoerus hydrochaeris*) and in the striped field mouse (*Apodemus agrarius*). In two (1 organ, 1 fecal) of the 15 positive striped field mouse samples, we found a novel murine adenovirus (MAdV-4). In the remaining 13 positive (fecal) samples, we identified MAdV-3, which had been isolated and described in Slovakia previously. The sequence of a truncated version of the E1B 19K gene (coding for 24 aa only) from the Slovakian isolate is available in the GenBank. The amplified E1B 19K gene sequence of the MAdV-3, detected in the Hungarian striped field mouse population showed 4 nt differences compared to the Slovakian isolate, but these caused 2 aa differences abolishing the stop codon. Thus, in the MAdV-3, detected by us, a “full” E1B 19K gene (coding for 174 aa) is present.

All together 359 bat samples were screened for AdVs from representatives from 28 bat species (194 samples from 17 species from Germany and 165 samples from 28 species from Hungary). The presence of AdVs was detected in samples of 18 species. From the 57 positive samples, 21 originated from Hungary and 36 from Germany. We have determined 34 distinct AdV sequences, from which 31 proved to be novel. All of them belong to the genus *Mastadenovirus*.

Genome analysis

The genome length of MAdV-2 K87 is proved to be 35,203 bp long with relatively high (63.35%) G+C content (GenBank acc. nr.: HM049560). The ITRs are 121 bp long and 28 genes were identified on the genome.

A 20,602 bp long genome fragment (between the genes of IVa2 and pVIII) was sequenced from SqAdV-1 detected in Germany in red squirrels that died after diarrhea. The sequenced genome fragment had a G+C content of 45.72% which can be considered as balanced. Seventeen genes corresponding to the genetic organization of mastadenoviruses were identified.

The genome sequence of BtAdV-2, isolated in Germany, was sequenced by next generation pyrosequencing. The assembled genome sequence (GenBank acc. nr.: JN252129)

was analyzed and the genes, ORFs and splicing donor and acceptor sites were determined by comparison to other mastadenoviruses. The genome size is 31,616 bp with a G+C content of 53.5%, ITRs of 146 bp and with 31 presumed genes.

Discussion

Novel adenoviruses

Within the class of mammals (Mammalia), rodents (Rodentia) and bats (Chiroptera) are the two orders richest in species, respectively. Their great importance in human health is due to their role as virus reservoirs.

This study describes the first detection of AdV in capybara (from the family Caviidae) and in common vole (which belongs to the family Cricetidae). Besides the confirmation of the presence of MAdV-3 in the Hungarian wild population of the striped field mouse, a novel AdV (MAdV-4) was also detected. Thirteen from the 19 samples were positive for MAdV-3. This shows that the virus occurs commonly in Hungary. Demonstration of SqAdV-1 in Germany is the first detection in continental European, as this virus had been detected on the islands of the United Kingdom only previously.

From 359 bat samples, screened by PCR, 57 (15.88%) proved to be positive. The positivity ratio in case of the German samples was a bit higher (36/194; 18.55%) compared to the Hungarian samples (21/165; 12.72%). This is mainly attributable to the origin of the samples. All the German samples were DNAs from organ samples of dead animals, while the significant part of the Hungarian samples comprised individual fecal samples or rectal swabs. At this latter kind of samples, the positivity ratio was extremely low (3/108; 2.7%). However, the positivity rate of guano samples, collected from under the Hungarian bat colonies, was also exceptionally high (17/54; 31.48%). Bats have been confirmed repeatedly as carriers of numerous pathogens causing zoonoses, including very dangerous viruses. An extremely high biodiversity is specific for bats. Compared to the average ~10% AdV positivity of all vertebrate samples, originating from diseased or dead individuals, found in our lab before, the number of positive samples among bats is outstanding. It is noteworthy, that within vertebrate samples, usually the number of AdV positive mammals is even lower (2–5%). In the individuals of a given bat species, multiple AdV types might be present just as in case of other animals and humans. There are AdVs which can infect representatives of more than one bat species. In some of the positive samples, the presence of several AdV types or genetic variants was verified. In three cases, the same AdV was revealed in geographically distant samples. In two cases the different regions were Germany and Spain, and in one case Germany and China. Interestingly, we didn't

find identical AdV sequences in the Hungarian or German samples even if they originated from the same host species. However, Hungary is situated geographically closer to Germany than Spain. This earlier mentioned phenomenon could be explained by the typical yearly Northeast-Southwest migration of European bats. Thus the opportunity for virus exchange is higher in individuals which migrate across the concerned countries.

The phylogenetic reconstructions confirmed the theory of coevolution of AdVs and their hosts corresponding also to the bat taxonomy. This is best demonstrated by the perfect separation of AdVs of bats belonging to the former suborders of Microchiroptera and Megachiroptera. The AdVs of bat families Vespertilionidae, Hipposideridae, Rhinolophidae and Pteropodidae also formed well-separated groups, however unfortunately these were not fully monophyletic.

Genome analysis

Analysis of the genome sequence of MAdV-2 confirmed its place in the genus *Mastadenovirus*. The few data that had been available in the GenBank previously from the MAdV-2 was identical with our sequences except 12 nt that caused 3 reading frame shifts in the submitted sequence. Our genome analysis corrects the mistakes, and determines the correct length of the concerned genes (of the protease and the DNA-binding protein). During the comparative genome analysis, substantial size difference between MAdV-2 (35,203 bp) and the other MAdVs (MAdV-1 – 30,944 bp, MAdV-3 – 30,570 bp) was noted. We concluded that it is attributable to the length rather than to the number of the genes, which are of the same number or just slightly more. Despite of the significant difference in size, the genome organization of the three MAdVs are very similar, only a few small differences can be observed. The absence of 12.5K and U exon (the first exon of a probable gene) located in the E3 region in MAdV-1 and 3 is an exceptional thing. In contrast, the counterparts of these genes are identifiable in MAdV-2. While MAdV-1 and MAdV-3 proved to be monophyletic in phylogeny reconstructions based on almost any genes, MAdV-2 appeared consistently on a separate, though monophyletic or neighboring branch. Based on the phylogenetic calculations, it is ascertainable that MAdVs seem to be the most ancient among the known mastadenoviruses. The phylogenetic distances between all the three MAdVs are greater than 5-10% which is a prerequisite of the classification AdVs to separate species. Consequently, all three MAdV types can be classified to a separate virus species.

The presence of squirrel AdV was detected always in dead or diseased red squirrels, but in one case in healthy grey squirrel (*Sciurus caroliensis*). Consequently it can be supposed that this AdV was introduced with its original host, the invasive grey squirrel, which increasingly occupies the red squirrel's territory in Europe. The genome organization of the studied part of

SqAdV-1 was typical of mastadenoviruses. This region contains 16 conserved genes present in all AdVs and also the gene of protein V which can be found in mastadenoviruses only. The genes of the IVa2, DNA polymerase, pTP and 33K have splicing sites in this virus as well, so these genes consist from two exons.

Just as all bat AdVs known to date, BtAdV-2 also belongs to the genus *Mastadenovirus*. The ITRs of mastadenoviruses are usually longer (93-371 bp) than in AdV of other genera. Comparing ITRs of BtAdV-2 (146 bp) with the close relatives CAdV-1 (199 bp) and CAdV-2 (198 bp), it is a bit shorter. In addition to BtAdV-2, the only other bat AdV of which almost the full genome is known is BtAdV-3. Unfortunately the end sequences of BtAdV-3 are not determined, thus no information is available about the length of its ITRs. This leads to the fact that BtAdV-2 is the first bat AdV, where the entire genome is known. The first 40 bp of BtAdV-2 ITR is identical with the corresponding sequence part of CAdV-1 and CAdV-2, which is additional evidence to the genetic relationship between bat and canine adenoviruses. The splicing pattern typical of the genus *Mastadenovirus* is strengthened by bat AdVs (BtAdV-2, BtAdV-3) as well.

The canine AdVs (CAdV-1 and -2) are very similar to some bat AdVs both phylogenetically and in genome organization. We supposed that they might have originated from some bat AdVs. The full genome sequence and analysis of the equine AdV-1 (EAdV-1) have been described recently and it became clear that the EAdV-1 probably also has bat AdV origin. Examination of the E3 region of AdVs is very important due to its high variability. The majority of the proteins encoded by the E3 genes play crucial role in the interaction of the AdV and the host's immune system. The genes in this region usually are not homologues. If they are then they can be found in very closely related AdVs only. There isn't any single E3 gene, which would occur in every mastadenovirus. In the E3 region of BtAdV-2, there is only one ORF besides the 12.5K. Apart from the E3 region of CAdV-1, -2 and EAdV-1, no homologue of this ORF has been found in any other mastadenovirus. In BtAdV-2, four additional ORFs (ORF A-D) are present right from the gene of the 34K in the E4 region, which is a similarly variable region. The homologues of these ORFs can be found in CAdVs, BtAdV-3 and EAdV-1 as well, however the role of the hypothetical proteins coded by these putative genes is still not known. On the phylogeny reconstructions, the CAdVs and EAdV-1 are monophyletic and placed always next to bat AdVs. The close relation of these AdVs, based on the similarities in their E3 and E4 regions is unambiguous.

New scientific results

1. I performed a PCR screening for the presence of adenoviruses in samples representing every bat species known to occur in Hungary, and the majority of bat species registered in Germany. We detected 31 novel bat adenoviruses and demonstrated that the same specimen can be infected by several different adenoviruses.

2. I analyzed and annotated the entire genome of bat AdV-2 and verified that certain mastadenoviruses of veterinary importance have close common ancestry with bat adenoviruses.

3. I demonstrated the presence of novel AdVs in rodents, namely in the striped field mouse, the common vole and the capybara. We published the first occurrence of squirrel AdV-1 outside of Great-Britain.

4. I participated in the full genome sequencing of murine AdV-2, analyzing the virus genome and its phylogeny.

5. I sequenced and analyzed two-thirds of the genome of squirrel AdV-1.

6. I have repeatedly detected MAdV-3 in the domestic striped field mouse population. I identified the full-length version of the E1B 19K gene, which is present in a truncated form in the original isolate of MAdV-3. I verified that the MAdV-3 circulating in the wild population of striped field mouse contains an intact E1B 19K gene.

Publications and conference abstracts, this work was based on

Publications

Hemmi, S., Vidovszky, M.Z., Ruminska, J., Ramelli, S., Decurtins, W., Greber, U.F., Harrach, B.: **Genomic and phylogenetic analyses of murine adenovirus 2**, Virus Res., 160. 128-135, 2011. *IF: 2,941*

Jánoska, M., Vidovszky, M., Molnár, V., Liptovszky, M., Harrach, B., Benkő, M.: **Novel adenoviruses and herpesviruses detected in bats,** Vet. J., 189. 118-21, 2011. *IF: 2,239*

Peters, M., Vidovszky, M.Z., Harrach, B., Fischer, S., Wohlsein, P., Kilwinski, J.: **Squirrel adenovirus type 1 in red squirrels (*Sciurus vulgaris*) in Germany**, Vet. Rec., 169. 182, 2011. *IF: 1,248*

Vidovszky, M.Z., Boldogh, S.: **Detection of adenoviruses in the Northern Hungarian bat fauna [in Hungarian]**, Magy. Allatorvosok, 133. 747-753, 2011. *IF: 0,201*

Kohl, C., Vidovszky, M.Z., Mühldorfer, K., Dabrowski, P.W., Radonić, A., Nitsche, A., Wibbelt, G., Kurth, A., Harrach, B.: **Genome analysis of bat adenovirus 2: indications of interspecies transmission**, J. Virol., 86. 1888-1892, 2012. *IF: 5,076*

Vidovszky, M.Z., Kohl, C., Boldogh, S., Görföl, T., Wibbelt, G., Kurth, A., Harrach, B.: **Random sampling of the European bat fauna reveals the existence of numerous hitherto unknown adenovirus**, Acta Vet. Hung., 63. 508-525, 2015. *IF: 0,646*

Conference abstracts

Vidovszky, M.Z., Harrach, B.: **Novel adenoviruses detected in bats in Hungary**, XV. International Congress of Virology, Sapporo, Japan, 2011.

Vidovszky, M., Ruminska, J., Ramelli, S., Decurtins, W., Doszpoly, A., Skoda, G., Hemmi, S., Harrach, B., Greber, U.: **Characterisation of the murine adenovirus 2 genome and partial sequences from similar rodent adenoviruses**, Acta Microbiol. Immunol. Hung., 58. s112-113. 2011.

Vidovszky, M.Z., Kohl, C., Boldogh, S., Görföl, T., Wibbelt, G., Kurth, A., Harrach, B.: **Novel adenoviruses detected in bats in Hungary and Germany**, 10th International Adenoviral Meeting, Umeå, Sweden, 2012.

Vidovszky, M.Z., Kohl, C., Boldogh, S., Görföl, T., Wibbelt, G., Kurth, A., Harrach, B.: **PCR screening of the German and Hungarian bat fauna reveals the existence of**

numerous hitherto unknown adenoviruses, XVIth International Congress of Virology, Montreal, Canada, 2014.

Further publications

Ballmann, M.S., Vidovszky, M.Z.: **Detection of broad-host-range psittacine adenovirus (PsAdV-2) in representatives of different parrot species**, Magy. Allatorvosok, 135. 78-84, 2013. *IF: 0,185*

Nguyen, T.H., Vidovszky, M.Z., Ballmann, M.Z., Sanz-Gaitero, M., Singh, A.K., Harrach, B., Benkő, M., van Raaij, M.J.: **Crystal structure of the fibre head domain of bovine adenovirus 4, a ruminant atadenovirus**, Virol. J., 12. 81, 2015. *IF: 2,181*

Acknowledgements

I would like to say thank you first of all to my supervisor Dr Balázs Harrach and to Dr. Mária Benkő, for the unbelievable familial atmosphere, which they make available in our laboratory to any student who has at least a small interest in research. A lot of thanks for their patience, motivation, professional advices and supports; as they keep guiding me even these days as “scientific parents”.

Thanks for the continuous help to all our colleagues with whom we have worked together on scientific manuscripts for months; to dr. Viktor Molnár for the many exotic animal samples from the Budapest ZOO, to dr. László Egyed, Tamás Görföl, dr. Péter Estók and especially to dr. Sándor Boldogh, for collecting samples without sparing effort, and thanks for dr. Susan Compton for the isolated MAdV-2 sample. Additional thanks to dr. Thomas Wilton for the help using correct English in the English parts of my thesis.

Many thanks are due to my students, Mónika Ballmann, Tamás Görföl and Katalin Földes, who while working together, helped a lot sometimes even overtaking me in some fields. Thanks for the help to all of my colleagues working next to me in the laboratory. To the walkers before me who introduced me into the secrets of techniques and methods. Including but not limited to dr. Győző Kaján, dr. Tibor Papp, dr. Andor Doszpoly, dr. Judit Pénzes and especially with friendly love to Máté Jánoska†. We have started the discovery of the wonderful world of bat AdVs together, and who unfortunately cannot be with us today.

Gratitude is to my family: to my father, mother, brother and my sister. They believed and still believe in me, supported me from the first class of primary school, and their example has always confirmed that there is no impossibility just inability. Thank you, last but not least, my lovely wife Sára Vidovszky and my children Hanga and Marci jr., to grant me the family background, which gave undoubtedly enormous strength and security to a researcher, balancing on a knife-edge between the great discoveries and great falls of science. I am also grateful to them for their patience exhibited during the preparation phase of my thesis.

The financial support of my work was provided by the Hungarian Scientific Research Fund grant NN107632.