**Brief Summary of Doctoral Thesis**

# Genetic analysis, biodiversity and phylogenetics of different viruses of ancient vertebrates

Zoltán László Tarján

Supervisor: Dr. Mária Benkő

UNIVERSITY OF VETERINARY MEDICINE
Doctoral School of Veterinary Science

Budapest, 2021

Supervisor and members of the supervisory board:

.....

Prof. Dr. Mária Benkő, DSc
supervisor
Veterinary Medical Research Institute,
Eötvös Loránd Research Network

Prof. Dr. Harrach Balázs, DSc
correspondent member of the HAS
member of the supervisory board
Veterinary Medical Research Institute,
Eötvös Loránd Research Network

Andor Doszpoly, PhD
member of the supervisory board
Veterinary Medical Research Institute,
Eötvös Loránd Research Network

.....

Zoltán László Tarján

# Introduction

Adenoviruses (AdVs) infecting ancient vertebrates have been studied with for decades using comparative genetic and phylogenetic methods in the research team of Molecular and Comparative Virology of the Veterinary Medical Research Institute. Based on their results, in addition to the original two genera of the *Adenoviridae* family (*Mastadenovirus* and *Aviadenovirus*), they have established three new genera named *Atadenovirus*, *Siadenovirus*, and *Ichtadenovirus*. Members of *Atadenovirus* were hypothesized to have coevolved with the reptiles, but later a growing number of data suggested that these viruses coevolved with the members of the order Squamata and may have switched hosts to birds and ruminants. Surprisingly, in 2007, at almost the same time, AdVs at a significant phylogenetic distance from those known previously were discovered in terrestrial and freshwater turtles in an American and two Hungarian laboratories respectively. Phylogenetic tree based on the PCR-amplified DNA polymerase gene fragment showed the newly discovered viruses as members of a lineage with genus-level distinction. It seemed interesting to screen for as many positive

samples as possible and to know and analyse the sequence of further fragments of the viral genome.

PCR screening of a large number of randomly collected samples from fishes and amphibians failed to detect either adenovirus or herpesvirus. However, I detected a hitherto unknown herpesvirus in a skin sample of a wels catfish showing skin lesion from a Hungarian fish farm. Within the recently established *Herpesvirales* order, members of the newly created *Alloherpesviridae* family are pathogens that infect fishes and amphibians and are important from a point of view of fisheries and animal health. Based on microscopic and electron microscopic findings, Hungarian researchers had attributed a skin lesion similar to the so-called carp pox to herpesvirus infection in catfish (*Silurus glanis*) native to natural waters of Hungary, but this has not been confirmed by molecular analysis for decades.

In my sample collection, I obtained a considerable number of positivity by PCR screening suitable for the general detection of circoviruses. The occurrence of viruses belonging to the family *Circoviridae* in different fish species was first described by Hungarian researchers, but until the beginning of my doctoral research, no circoviruses were detected in

amphibians and reptiles. Because of the jumping growth of sequence data obtained from environmental samples, the extraordinary number, diversity and prevalence of viruses called CRESS DNA viruses gradually became known. A common feature of these viruses is that they have a small, circular, single-stranded DNA genome that encodes a replication protein, too. The occurrence of CRESS DNA viruses has been detected in vertebrates and invertebrates, as well as in various plants and even in fungi. So far, little is known about their role in the virosphere and their potential pathogenicity.

# Objectives

- PCR screening of samples from ancient vertebrates to detect members of three virus families (*Adenoviridae*, *Alloherpesviridae* and *Circoviridae*).

- Determination of the nucleotide sequence of the whole genome or the largest possible genome fragments from the detected new viruses.

- Analysis of the newly obtained sequences, phylogenetic analysis of the viruses, determination of the taxonomic position of the new viruses.

- Study of hypothetical coevolution of viruses and host species.

# Materials and methods

## Origin of the examined samples

For the AdV screening of the turtles, I used samples from 148 individuals of 23 turtle species. I took cloacal swabs from the live animals and examined mixed organ samples from cadavers. Mixed organ samples from 40 fish species (407 individuals), 15 amphibian species (52 individuals) and 29 reptile species (107 individuals) were used to detect CRESS DNA viruses in ancient vertebrates. An alloherpesvirus was detected using a skin sample of a two-summer old catfish with smallpox-like proliferative skin lesion sent for testing to our laboratory from a Hungarian fish farm. DNA was extracted from the samples in each case.

## PCR, DNA sequencing and molecular cloning

To detect the presence of viral nucleic acid, I applied widely used PCR methods reported in the international literature, which were designed to amplify conserved regions in specific genes of each virus family. In the case of turtle adenovirus and wels catfish herpesvirus, I attempted to amplify additional gene segments and connect them by normal- and nested

PCRs, respectively, using consensus and specific primers. From samples positive for the presence of CRESS DNA virus, I attempted to amplify the whole genome with self-designed specific primers using inverse, nested PCRs. Sequencing of the PCR products were performed by Sanger reaction using BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). For PCR products resulting in multiple signals, I separated the DNA fragments of the same size but not of viral origin by molecular cloning using the CloneJET™ PCR Cloning Kit (Thermo Scientific).

**Data analysis and phylogenetic studies**

The nucleic acid sequences determined from both sides were aligned, edited and corrected using the Gap v4.11.2 program of the Staden software package. Using the BLAST homology search application, I performed the verification of the obtained sequences online. Positional alignment of amino acid and nucleotide sequences was performed using T-Coffee. The derived amino acid sequences were generated using JavaScript DNA Translator 1.1. I used ProtTest v3.3 and jModelTest 2.1.10 to select the evolutionary model that best fits my amino acid and nucleic acid alignments. In the case of

red-eared slider AdV, I removed poorly aligned positions and divergent regions of protein alignments using the Gblocks v0.91 online server. For HVs and CRESS DNA viruses, I manually edited the alignments with BioEdit v7.0.9.0. For maximum likelihood analysis, I used PhyML 3.0. For the phylogenetic analysis based on Bayesian statistics, I applied MrBayes 3.2.7. Tanglegrams showing the hypothetical common evolutionary past of viruses and their natural hosts were generated using the Dendroscope v3.7.2 program. To create schematic genomic maps, I used the DNAPlotter r1.11 program of the Artemis r16.0.0 software package.

# Results and discussion

## Red-eared slider adenovirus

Based on the sequence of the PCR product obtained in the screening assays, a single variant of the newly discovered adenovirus line was present in a sample from the internal organs of a dead red-eared slider, making this sample a suitable starting material for sequencing the viral genome. Based on the sequences of the products previously amplified from the DNA polymerase as well as the hexon gene, I designed virus-specific primers to perform nested PCR. Using consensus primers designed for the IVa2 gene at the left side of the genome, I performed another connecting PCR to a known part of the DNA polymerase gene. After alignment of the sequences of the PCR products, a nucleotide sequence of 14,776 base pairs became known from the middle, conserved region of the viral genome. The estimated G+C content was 55.2%, which can be considered balanced. During sequence annotation, I identified two partial (IVa2 and hexon) and eight complete genes (DNA polymerase, pTP, 52K, pIIIa, penton base, pVII, pX and pVI). Their relative size,

orientation and sequence corresponded to the general characteristic organization of adenoviruses. A possible splicing donor and acceptor site was detected in the terminal protein precursor (pTP) gene.

I searched for the cleavage signal motifs of the viral protease in the amino acid sequence of five precursor proteins, as these are generally characteristic and conserved in members of each genus, so their number as well as their exact sequence may be suitable for taxonomic separation of the genera. For the pVI and pX precursor proteins, I found significant differences compared to the other genera. For pVII, the difference was less pronounced. Phylogenetic distance was examined by two complete (DNA polymerase and penton base) and partial hexon gene tree reconstructions. The virus, called red-eared slider adenovirus (RESAdV-1), appeared on all three trees on a branch well separated from the officially accepted genera, and its sister group was in each case the white sturgeon adenovirus 1 from the genus *Ichtadenovirus*. Also the demonstrated phylogenetic distance justified the classification of the virus as the member of a new species and the creation of a new genus for similar adenoviruses detected in turtles. Based on our results,

we proposed and the International Committee on Taxonomy of Viruses (ICTV) officially accepted the sixth genus with a name of *Testadenovirus*. The founder species was the new species *Pond slider testadenovirus 1*. The separation of the lineage of the new turtle adenoviruses was also confirmed by the results of coevolution studies. It can be hypothesized that the ancestor of the turtle adenoviruses already infected the ancient turtles, and during the evolution testadenoviruses evolved alongside in parallel with the speciation of the turtles.

**Wels catfish herpesvirus**

Smallpox-like skin lesions observed on a wels catfish sent to our institute for examination were similar to the skin lesions attributed to herpesvirus infection based on light and electron microscopic examinations described by Békési et al. (1981). The sequence of a successful PCR product targeting the DNA polymerase gene of alloherpesviruses suggested the presence of a new alloherpesvirus. The PCR product of the terminase gene yielded an unreadable sequence, suggesting that the amplicon is a mixture. After molecular cloning, the fragment from the terminase was isolated. To our

surprise, however, it was found that among the various non-specific products, there was even a fragment of viral origin, which showed homology to a gene of a protein with a previously unknown function in some members of the genus *Cyprinivirus*. Based on the sequences of this PCR product and the DNA polymerase gene fragment, specific primers were designed to connect these genes. By sequencing the obtained product, I had the opportunity to analyse a 15,222 bp genome stretch. Nine ORFs (two partial and seven complete) found in the sequence of the genome segment were accepted as genes based on the homologous genome parts of anguillid herpesvirus 1 and cyprinid herpesvirus 1. Four of the seven complete genes were found to be specific for wels catfish herpesvirus. Based on the results of comparative genome studies, the amplified partial genome segment of wels catfish HV is clearly more similar in organization to the corresponding genome part of the CyHV-1. Because the full genome sequence of an alloherpesvirus found in glass catfish has recently been reported as silurid herpesvirus 1, we propose the silurid herpesvirus 2 (SiHV-2) name for the new virus detected in wels catfish and for a new species incuding it the name *Silurid herpesvirus 2*. According to phylogenetic

tree reconstructions based on partial or complete gene sequences, this new virus was shown to belong to the genus *Cyprinivirus*, which contains eel herpesvirus in addition to carp alloherpesviruses. Phylogeny of viruses belonging to the *Cyprinivirus* and that of their host species showed a similar tree topology.

**CRESS DNA viruses of lower vertebrates**

With nested PCR designed for the conserved region of the replication protein, I obtained positive results in 8 of 566 samples during screening, so that the entire viral genome was subsequently amplified. The nested PCR method (Halami et al., 2008), originally designed to detect circoviruses, has also been shown to detect cyclovirus (CyV) and CRESS DNA viruses related to the *Circoviridae* family. Positive PCRs were obtained by examining samples from a bream, stone morokoes, an eel, a gudgeon, an Amazonian milk frog and a red-eared slider. Each of the viruses has proven to be new to science. I detected porcine circovirus type 2 (PCV-2) from a mixture of organ samples from a monkey goby. Due to the heterogeneity of the PCR product, in some cases (stone moroko CV-1 and 2, gudgeon-associated CyV, Amazon milk frog-associated CRESS DNA virus),

molecular cloning was required to seprate the viral-derived amplicon. The entire genome of PCV-2 found in the monkey goby was 1,766 nucleotides, and from the sequences in the GenBank database, BLASTn showed the largest match (99.2%) with a sample from Brazil (KT819161). Its presence in fish is thought to be of food origin. Based on the results of whole genome studies, 4 of the 7 new viruses (detected in common bream, eel and two viruses from stone morokos) belong to the circoviruses. The virus described from the gudgeon proved to be a cyclovirus, while the Amazon milk frog virus can be classified only as belonging to the CRESS DNA viruses. The virus from the red-eared slider is also a CRESS DNA virus based on the sequences, but its genome size is smaller. The sequence identity calculated from the pairwise alignment of the total genomic DNA of the two CVs detected from stone morokos is 97.6%, so they can be considered as two variants of one virus. Phylogenetic tree reconstructions supported the results of genome studies. Circoviruses detected in fish are close to each other but do not form a monophyletic lineage and their sister group is the clade of fish cycloviruses. Based on this observation, we hypothesized that the genus *Circovirus* may be derived

from fish cycloviruses and may have been induced by genome inversion. Based on the results, 4 of the 7 new viruses (circovirus of bream and stone morokos, cyclovirus of the gudgeon, and CRESS DNA virus of the Amazon milk frog) are presumably independent, recent viruses. Due to the small size of the capsid protein gene, the circovirus-like sequence detected in eel, probably originated from a DNA segment integrated into the host genome. The sequence detected from the red-eared slider does not have a capsid protein gene homologue, so it is unlikely to function as an independent virus. Several scenarios are conceivable for its origin. Options may include genomic integration, free replicon, or replication and capsid genes packaged in separate capsids, similar to some plant viruses. Regarding the host origins of the two CRESS DNA viruses, their position on the phylogenetic tree does not provide reliable information. The virus detected in the milk frog sample is presumed to be of insect origin and entered the frog via food.

## New scientific results

1. I determined the nucleotide sequence of a nearly 15 kb middle genome part of an adenovirus detected from a red-eared slider representing a new virus lineage significantly different from the known genera of the *Adenoviridae* family. Analysis and phylogenetic study of the encoded proteins supported the creation of a new genus that was officially approved by the ICTV as *Testadenovirus* based on our proposal. The founding member of the genus is the *Pond slider testadenovirus A* species.

2. I detected a new alloherpesvirus from a wels catfish with smallpox-like skin lesions. I determined more than 16 kb sequences from the viral genome. Based on the analysis of the encoded proteins, I found the virus to belong to the genus *Cyprinivirus* of the family *Alloherpesviridae*. I named the new virus silurid herpesvirus 2, which is a close relative of the herpesviruses of eel and carp, and I propose the establishment of a new species for it.

3. Using PCR, I obtained previously unknown *rep* gene fragments from samples of various ancient vertebrates. The complete circular genome was also amplified from seven samples. Based on the sequence analysis, the presence of viruses belonging to the *Circoviridae* family and the CRESS DNA virus group can be assumed in an eel, a common bream, stone morokos, a gudgeon, an Amazon milk frog and a red-eared slider.

4. In most cases, I demonstrated a phylogenetic relationship between the detected viruses and their putative host species, suggesting a long coevolutionary history.

# Scientific publications

**Publications on the topic of the thesis in peer-reviewed journals**

Harrach B., Tarján Z.L., Benkő M.: **Adenoviruses across the animal kingdom: a walk in the zoo**, FEBS Lett., 593. 3660–3673, 2019.

Tarján Z.L., Doszpoly A.: **Diversity of adenoviruses occurring in turtles (Autoreferatum and literature review)**, Magy. Állatorvosok (in Hungarian), 141. 747–757, 2019.

Doszpoly A., Tarján Z.L., Glávits R., Müller T., Benkő M.: **Full genome sequence of a novel circo-like virus detected in an adult European eel *Anguilla anguilla* showing signs of cauliflower disease**, Dis. Aquat. Organ., 109. 107–115, 2014.

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

Doszpoly A., Wellehan J.F. Jr., Childress A.L., Tarján Z.L., Kovács E.R., Harrach B., Benkő M.: **Partial characterization of a new adenovirus lineage**

**discovered in testudinoid turtles**, Infect. Genet. Evol., 17. 106–112, 2013.

**Publications on other topics in peer-reviewed journals**

Kaján G.L., Doszpoly A., Tarján Z.L., Vidovszky M.Z., Papp T**.: Virus–host coevolution with a focus on animal and human DNA viruses**, J. Mol. Evol. 88. 41-56, 2019.

Csépányi-Kömi R., Sáfár D., Grósz V., Tarján Z.L., Ligeti E.: **In silico tissue-distribution of human Rho family GTPase activating proteins**, Small GTPases, 4. 90–101, 2013.

# Acknowledgements

It is my honour to thank to Dr. Mária Benkő, who, as my supervisor and a private person, has always stood by me in everything, helped me with her theoretical and practical knowledge, experience, advices and led the way on becoming a researcher.

I am also grateful to academician Dr. Balázs Harrach, who has often inspired and steered my ever-wandering thoughts in the right direction. In addition to the careful improvement of my dissertation, I could always count on him both professionally and humanly.

I would like to say thank to the current and former members of the Molecular and Comparative Virology and Fish Parasitology research teams for creating the research environment and for being able to consult anyone, anytime, about methods, results, ideas, or even just the neat and bad things in life.

I am grateful to Dr. Edit Eszterbauer for her help in testing catfish HV for providing me with an organ sample for DNA extraction and for using the laboratory she led if necessary.

I would like to thank to my mother, my godmother, my younger brother Szabolcs, my wife Nóri

and her parents, and Geri, Vili and Misa for creating the peaceful family atmosphere without which thoughts cannot be put together. Szabolcs also provided fish samples, and together with Nóri also helped me with the field samplings, for which I am especially grateful.

The financial background was provided by the OTKA K100163, NN107632 and NN128309 grants and the PhD framework of the Doctoral School of Veterinary Medicine of the Szent István University.