Hungarian University of Agriculture and Life Sciences

Molecular characterisation of reptilian RNA viruses

Summary of Phd thesis

DOI: 10.54598/003290

Katalin Tóth-Ihász

Keszthely

2022

The doctoral school

Name: Doctoral School of Environmental Sciences

Discipline: Environmental Sciences

Head of school: Dr. Angéla Anda DSc

University professor
Hungarian University of Agriculture and Life Sciences
Institute of Environmental Sciences

Supervisors: Dr. Szilvia Farkas

Clinical veterinarian, PhD
University of Veterinary Medicine

Dr. Szabolcs Tamás Nagy

University professor, Doctor of Science
Hungarian University of Agriculture and Life Sciences

……. ……

Approval of the Head of School Supervisor approval

…….
Supervisor approval

1. Background of the work and objectives

Nowadays, keeping exotic animal becoming increasingly popular worldwide, but our knowledge about viruses of reptiles is limited. In the absence of sufficient data, diagnosis of these viruses is difficult; the pathogenic background of infectious diseases of reptiles remains unclear. Various next-generation sequencing (NGS) methods can provide an opportunity to collect complete genomic sequence data of the circulating virus strains, broadening our knowledge about genetic diversity and phylogenetic relationships of the viruses causing infections. We have worked with RNA viruses infecting reptiles, focusing on arenaviruses, orthoreoviruses and picornaviruses.

The family *Arenaviridae* (order *Bunyavirales*) comprises two out of four genera infecting reptiles: *Hartmanivirus* and *Reptarenavirus*. Members of *Reptarenavirus* genus – which is relevant to our work – have a bisegmented single-stranded RNA genome. Currently, five species are classified by the International Committee on Taxonomy of Viruses (ICTV) (*California-*, *Golden-*, *Giessen-*, *Ordinary-*, and *Rotterdam reptarenavirus*). The reptarenaviruses are the causative agents of boa inclusion body disease (BIBD). The name of the disease originated from the large eosinophilic inclusions described in the cytoplasm of almost all cell types of the infected animals. The clinical and histological manifestations and progression of BIBD differ in boid snakes and pythons, but emaciation and central nervous symptoms are common. While there are known asymptomatic carriers in boas, the disease is more severe in pythons. BIBD and reptarenaviruses has been described and detected in several snake species, but the majority of cases are known from red-tailed boas (*Boa constrictor*).

The genus *Orthoreovirus* is one of the numerous members of the family *Reoviridae*. The double layered icosahedral virions contain the ten double-stranded RNA genome segments. Although reoviruses are commonly identified

in reptiles and associated with different symptoms of the respiratory tract and central nervous system, their pathogenic role is not clear. There are currently ten orthoreovirus species in the genus. Reoviruses are isolated from squamatas were considered as members of the species *Reptilian orthoreovirus* (RRV).

The family *Picornaviridae* includes small, vertebrate-infecting, single-stranded RNA viruses, currently comprising 65 genera. Picornaviruses have been detected in several species of tortoises, usually as a co-infecting agent with other pathogens. Infections have been associated with various clinical signs and pathological lesions including softening of the carapax of young animals. On the other hand, asymptomatic carriers are also known, so the exact pathogenetic role of tortoise picornaviruses remains to be clarified. Picornaviruses of tortoise species are divided into two genera: *Rafivirus* and *Torchivirus*.

The aim of our study was:

- to determine the whole genome sequence and taxonomic classification of a reptile-derived arenavirus
- to determinate and analyze the whole genome sequence of a reptile-derived orthoreovirus
- to survey the prevalence of RRV infection in Hungary, to determine the complete or partial genome sequences of the detected viruses, to gain a better understanding of the genetic divergence of RRV
- to isolate picornaviruses from terrestrial tortoise species and to determine the whole genome sequences of the isolated viruses
- to develop an effective diagnostic system specific to tortoise picornaviruses.

2. Materials and methods

Samples

Tissue samples (liver, stomach, intestine, kidney, ovary, heart, trachea, tongue, esophagus) were collected from an adult, female, captive red-tailed boa (*Boa constrictor*) that had succumbed at a private owner in Hungary.

The determination of a reptilian orthoreovirus whole genome sequence was performed on the bush viper (*Atheris squamigera*) reovirus strain (47/02) isolated in Germany. To survey the prevalence of RRV infection in Hungary, we collected organ samples from dead animals (n=111) from pet shops and hobby keepers.

Six tortoise picornavirus strains were provided from our german collaborating partner originated from greek tortoise (*T. graeca*, 124/10, 144/10, 5/04), African spurred tortoise (*Geochelone sulcata*, 9/05), Hermann's tortoise (*Testudo hermanni*, 14/04), spider tortoise *(Pyxis arachnoides,* 5/03). One strain originated from a greek tortoise (*T. graeca*) was isolated in our laboratory (2013/T4).

Molecular methods

Viral nucleic acid was purified from organ samples using TRI Reagent and tested for arena-, orthoreo- and picornaviruses by broad-range reverse transcription PCR (RT-PCR) with universal primer sets. PCR products were cut and purified from agarose gel. The nucleotide sequences of the amplified PCR products were determined by Sanger sequencing.

Viral nucleic acid was purified using TRI Reagent. Reverse transcription was performed with the viral RNA using the oligonucleotide FR26RV-N consisting of a 3' random hexamer tag. cDNA was amplified in PCR using the FR20RV oligonucleotide. The PCR product was run in agarose gel, products

between the size of 200-2000 base pairs were excised and extracted. Libraries were prepared and whole genome sequencing was performed on IonTorrent PGM platform. To obtain the terminal sequences a modified 5' and 3' RACE method was applied.

We developed primers for specific detection of tortoise picornaviruses based on nearly complete genome sequences determined by NGS. To test the specificity, reptilian origin orthoreo-, adeno-, irido-, rana-, herpes- and paramyxovirus strains were used as templates in the assay. Analytical sensitivity was tested by serial dilutions of tortoise picornavirus isolates.

Virus isolation

We try to isolate arena-, reptilian orthoreo- and tortoise picornavirus strains from the positive RT-PCR tissue samples on the appropriate Russell's viper (VH 2, ATCC CCL-140), iguana heart (IgH-2, ATCC CCL-108) and tortoise heart cell lines (TH-1, ATCC CCL50). Virus isolates originated from our collaborating partners were also maintained on these cell lines.

Bioinformatics

CLC Genomics Workbench v7 was used to clean and assemble NGS data. Sanger sequencing reads were edited by BioEdit and Geneious Prime softwares. Geneious Prime and AliView softwares were used to assemble NGS contigs and Sanger reads. BLASTn and BLASTp algorithms were used to identify homologous genes among sequences deposited in GenBank. Codon-based multiple alignments were made using Geneious Prime and TranslatorX softwares. Phylogenetic analysis was performed, and sequence identity values were calculated using the MEGA6 and MEGA10 package. Gene-specific substitution models were evaluated, and the best-fit models were selected based on the Bayesian information criterion. Maximum-likelihood phylogenetic trees were generated, and tree topologies were validated by bootstrap analysis (500

or 1000). Nucleotide (nt) and amino acid (aa) sequence distances were calculated using the p-distance method. Hypothetical proteolytic cleavage sites were predicted using NetPicoRNA online software.

3. Results and discussions

Investigation of a reptile-derived arenavirus

The RT-PCR targeting the *GPC* gene of arenaviruses gave positive results and the sequencing confirmed the presence of an arenavirus in all the examined organs of the red-tailed boa (liver, stomach, intestine, kidney, ovary, heart, trachea, tongue, oesophagus). The virus was tentatively named Coldvalley virus. Attempts for the isolation in cell culture remained unsuccessful after several passages. Isolation and propagation of reptarenaviruses using VH 2 and other mammalian and reptilian cell lines have been reported with various degrees of success.

A large fragment of the viral genome sequence (8755nt) could be determined by viral metagenomics approach, directly from the liver sample. After combining this with the supplementary information obtained from the Sanger sequencing, two contigs were assembled and based on bioinformatic analyses one L and one S segment was identified. The complete sequence of the L segment was determined (6860 nt), and the S segment was partially determined (1985 nt). The 5' untranslated region (UTR) of the L segment is 87 nt, the 3' UTR is 47 nt long. The segment contains two ORFs: the Z protein (115 aa) in the virus genome-sense sequence and the viral RNA-dependent RNA polymerase (RdRp) protein (2068) in the genome-complementary sequence with the intergenic region in between (172 nt). The 5' end of the S segment is unknown, the 3’ end encode the nucleoprotein (NP) (584 aa) in the genome-complementary sequence. The 3' UTR (29 nt) and the intergenic region (111 nt) has been determined also. Considering that the L segment contains the *Z* gene in addition to the *RdRp*, our results suggest that Coldvalley virus was similar to that of reptarenaviruses.

The currently used classification of arenaviruses is based pairwise sequence comparisons (PASC) of coding-complete genomes: S and L segment nt sequence identities for viruses within the same genus need to be higher than 40% and 35%, respectively. Two arenaviruses should be classified in a common species if their nt sequence identity values are higher than 80% and 76% for the S and L segments, respectively, and the aa sequence similarity values are higher than 88% in case of the NP.

Comparing the Coldvalley virus sequences with homologous segments of representative members of the family *Arenaviridae*, the highest identity values were seen with members of the genus *Reptarenavirus* (L segment: 56.3-98.9.0%, S segment: 65.4-98.7%).

Genome segment sequences of the Coldvalley virus as well as the nt and aa sequences of *RdRp*, *NP* and *Z* genes were compared with other members of the genus *Reptarenavirus* availabe at the GenBank. In all cases, the highest identity values were obtained with the University of Helsinki virus strains (L segment 98.8-99.3% nt, S segment 98.9% nt; RdRp 99.3% nt, 99.5% aa; NP 98.9% nt, 99.6% aa; Z 99.8% nt, 100% aa), including those closely related to UHV-3.

Besides the above-mentioned sequences, the *RdRp* of UHV-1 (KF297881) and ROUTV (KC508670) strains, type sequences of the *Rotterdam reptarenavirus* species, showed relatively high nt and aa identities with the Coldvalley arenavirus sequences (82.1–87.2% nt and 84.9–87.8% aa identity). Regarding the *NP*, the ROUTV (KC508669) strain grouped together with the Coldvalley arenavirus (98.5% nt and 98.9% aa identity), while the UHV-1 (KF297880) represented a slightly more distant cluster (77.6% nt and 85.0% aa identity). This finding suggested potential recombination events among variable reptarenaviruses that have been described in other studies while it was not detected for Coldvalley virus and UHV-3.

The phylogenetic analyses and pairwise identity calculations revealed that the sequence of the Coldvalley arenavirus was most similar to, and grouped together on the phylogenetic trees with the sequences named UVH-3 within the *Rotterdam reptarenavirus* species.

Investigation of a reptile-derived orthoreoviruses

The cytopathogenic effect of bush viper reovirus on VH 2 cell line showed syncytium formation characteristic of orthoreoviruses. The successful isolation provided an opportunity for further investigations: the PAGE system showed genome segmentation and the genome sequence (24043 nt in total) was determined from the purified nucleic acid. The structure of the bush viper reovirus genome segments was typical of orthoreoviruses: the ten double stranded RNA genome segments are divided into three class in size, the nucleotide sequences at the genome segment termini were found to be highly conserved (5' GUUA/CUU, 3' UCAUC), all segments encode a single protein except the bicistronic S1 genome segment.

According to the current taxonomy in the genus Orthoreovirus, specific species demarcation criteria have been defined based on sequence identity values. >75% nt sequence identity between homologous genes is the cut-off value for most genomic segments to classify virus strains into the same species, and <60% between species. The cut-off values for most conserved core proteins >85% aa identity within the species and <65% between the species. For the more divergent outer capsid proteins: >55% identity within the species and <35% between the species.

The genes of RRV 47/02 had low to moderate nt and aa sequence identities to homologous genes of other members of the genus *Orthoreovirus*, whereas the most similar virus was *Testudine orthoreovirus* (TRV). The close relationship with TRV is supported by the topologies of the phylogenetic trees, in addition to the bicistronic genomic segment structure: their S1 segment

encodes the σC, as well as the cell fusion protein (p14). However the identity values (49,4-71,6% nt, 47,05-82,32% aa) are below the cut-off values for identical species. Compared to the partial RRV sequences in the GenBank, the bush viper reovirus shows high identity values (65-95% nt, 75-90% aa).

Based on the genome organization, identity values and the topology of phylogenetic trees the bush viper reovirus is a member of the RRV species, so we successfully determined the first complete RRV genome sequence, which is designated as a type strain by ICTV.

Another aim of our studies was to survey the prevalence and genetic diversity of orthoreoviruses. Applying RT-PCR amplification of a short, 245 nt long sequence of the orthoreovirus *RdRp* gene, 111 organ samples from 20 exotic reptile species were tested and 5 (4.5%) positive samples were found, from ball python (*Python regius*, n=2), green lizard (*Iguana iguana*, n=1), Schneider's skink (*Eumeces schneideri*, n=1) and an unidentified snake species (n=1). Compared with representative members of *Orthoreovirus* genus, the highest identity values were obtained with the bush viper reovirus (78-96% nt, 86-100% aa) and the TRV (68-78% nt, 85-90% aa). The partial nt and aa sequences of the RRV strains show high identity values (71-100% nt, 88-100% as) compared to the homologous sequences of RRV strains deposited in the GenBank. These aa values are below the cut-off values for the same species classification (>85% for core proteins), while not all strains reach a nt identity value higher than 75%. Based on the phylogenetic tree of the *RdRp* gene, all reovirus strains detected (2013/12, 2013/67, 643/47, 2013/KP1, 2013/KP3, 2013/54) are member of the RRV species. Reoviruses were isolated on cell culture from all the RT-PCR positive and 2013/KP3 samples. Electron microscopic examination of strain 2013/KP1 revealed virus particles with icosahedral capsid morphology characteristic of orthoreoviruses.

Investigation of a tortoise-derived picornaviruses

Applying the combination of next-generation and Sanger sequencing we obtained the nearly complete (7065-7079 nt, excluding the 5' UTR) genome sequences of seven picornavirus strains from different terrestrial tortoise species. Six of those were from our collaborators in Germany and one was isolated in our laboratory. All strains possess a single ORF – flanked by 5' and 3' UTRs – encoding a single polyprotein which is co-and post-translationally processed into mature viral polypeptides. The predicted ORFs were 6651-6657 nt, the encoded polyproteins were 2218 aa (except the 5/03 strain 2216 aa) long. The complete genome sequence of the 3' UTRs was determined (223-231 nt), the lengths of the 5' UTRs remain to be determined.

Low to moderate similarity values of the tortoise picornavirus strain were observed comparing the genome sequences of the picornavirus strains with the representative members of the family: the greatest similarity values were seen with members of the genera *Mosavirus* (61.5-63.1%). In the comparison of the polyprotein sequences Mischivirus D (KY512802, 68.1-71.7% nt) and Mosavirus A (JF973687, 51.2-52.2% aa) were most closely related. When comparing the whole genome sequence of tortoise derived Rafivirus A (KJ415177) we obtained moderate identity values (49.0-50.2%), with the polyprotein sequence showing only 30.3-31.1% nt and 44.7-45.7% aa identities.

Comparison of the polyprotein and sequences of the seven tortoise picornavirus strains gave high nt (81,2-99,0%) and aa (90,1-97,6%) identity values. Also high values were seen (71,9-98,9%) in the case of the nearly complete genome segments, while comparing individual genomic regions gave high identity values (P1 nt 81.3-99.0%, aa 92.5 99.1%; P2 nt 78.6-98.7%, aa 88.1-98.9%; P3 nt 81.3-99.0, aa 91.3-99.2%; 2C nt 81.4-99.7%, aa 92.8 100%; 3C nt 82.3-98.9%, aa 92.5-93.5%; 3D nt 80.0-98.6%, aa 87.3-99.4%).

Members of the *Picornavirus* genus cluster on the same branch in phylogenetic trees. The members of the known picornavirus genera differ by significant divergence of the ortholog proteins of the strains analyzed in the present study (P1 region >66%, 2C, 3C and 3D region >64%). Picornavirus species share a significant degree of aa identity of P1, 2C, 3C and 3D proteins, monophyly in phylogenetic trees, identical genome maps and significant degree of compatibility in proteolytic processing.

We also compared the sequence of the P1, 2C, 3C and 3D regions and obtained a divergence of 1-9.5% for the P1 region, 0.3-9.6% for the 2C region, 0.3-16.4% for 3C, and 0.9-20.0% for the 3D region. These values indicate the calssification of the strains into the same genus.

In the P1 and 3D phylogenies, the tortoise picornavirus strains formed a distinct lineage and most closely related to the Mosavirus A, which indicate their close relationship.

Based on the identity values and phylogenetic topologies described above, our tortoise picornavirus strains are distinct from the members of the genus *Rafivirus* and form a separate genus within the family. In the light of these results, we submitted a proposal to ICTV to form a separate genus for our strains, which was accepted and named *Torchivirus*. There is only a single species in the genus, *Torchivirus A*.

We developed a diagnostic RT-PCR system for rapid and specific detection of the *Torchivirus* genus based on the identified seven tortoise picornavirus strains. To test the specificity of the newly developed RT-PCR assay, a total of 25 virus isolates were tested for the presence of picornavirus RNA using two different RT-PCR assays. Similar cytopathogenic effect was detected in all samples. The tortoise picornavirus–specific primer sets amplified two regions of the genome, containing a portion of the 5′-untranslated region, the Leader peptide, and the VP4 capsid protein (Pic-gen-FOR1 and Pic-

gen-Rev1), and a portion of the picornavirus RNA-dependent RNA polymerase (3D) (Pic-gen-FOR2 and Pic-gen-REV2), respectively. All picornavirus isolates proved to be positive with both assays from all isolates, cloacal swabs and organ samples, in contrast, no specific bands could be observed on the isolates of known reptile-pathogen viruses. Sensitivity testing yielded positive results and can be successfully used to detect small amounts of picornavirus in the samples.

The nt and aa sequence identity values of these strains were found to be high for all three genes compared to each other (leader peptide nt: 67.3-100%, aa: 67.9-100%; VP4 nt: 72.9-100%, aa: 94.2-100%; partial 3D nt: 77.5-100%, aa: 86.9-100%), while moderate identity values were obtained with RaV-A1 (VP4 nt: 39.8-44.4%, aa 22.8-26.3%; leader peptide nt: 23.1-27.1%, aa: 9.3-12.0%; 3D nt: 49.4-52.7%, aa: 39.1-41.4%). Examination of the topologies reveals a similar structure for all three trees: two main lineages are separated, but no association could be observed between the geographic distribution and the genetic relatedness of the strains.

4. New scientific results

1. We were determined by NGS and analyzed the nearly complete genome sequence of a hungarian red-tailed boa-derived reptarenavirus. Phylogenetic analysis of the sequence and taxonomic classification of the virus strain was performed. This is the first Hungarian reptarenavirus genome sequence.

2. We first determined and analyzed the whole genome sequence of a *Reptilian orthoreovirus* strain 47/02 from a green bush viper.

3. We investigated the prevalence of RRV infection of different reptile species from private collectors and pet shops in Hungary. Using partial sequence data, phylogenetic analyses were performed and six RRVs were successfully isolated in VH 2 cell line.

4. Tortoise-derived picornavirus was isolated and their complete genome sequence was determined together with six other tortoise-derived picornavirus strains. The sequences were analyzed, which revealed that these strains show only moderate similarity to other picornaviruses. In the light of our results, we proposed to ICTV the formation of a separate genus, named *Torchivirus*. The type species of the genus became strain 14/04 as *Torchivirus A* species.

5. Using the sequences of the determined tortoise picornavirus strains, we developed a diagnostic RT-PCR system that can specifically detect the nucleic acid of the members of the *Torchivirus* from both virus isolates and tissue samples, and can therefore be used for diagnostic purposes.

5. Publications in peer-reviewed journals

Farkas, S. L., Ihasz, K., Erdelyi, K., Marschang, R. E., Papp, T., Pilis, T. & Banyai, K. (2013). Kígyók sejtzárványos betegsége. Irodalmi áttekintés. *Magyar Állatorvosok Lapja* 135, 751-757.

Banyai, K., Borzak, R., Ihasz, K., Feher, E., Dan, A., Jakab, F., Papp, T., Hetzel, U., Marschang, R. E. & Farkas, S. L. (2014). Whole-genome sequencing of a green bush viper reovirus reveals a shared evolutionary history between reptilian and unusual mammalian orthoreoviruses. *Archives of Virology* 159, 153-8. https://doi.org/10.1007/s00705-013-1796-2

Ihasz, K., Farkas, S. L., Lengyel, G., Bányai, K. & Gál, J. (2014). Orthoreovírusok előfordulásának és genetikai diverzitásának vizsgálata egzotikus hüllőfajokban Magyarországon. *Magyar Állatorvosok Lapja* 135, 247-252.

Farkas, S. L., Ihasz, K., Feher, E., Bartha, D., Jakab, F., Gal, J., Banyai, K. & Marschang, R. E. (2015). Sequencing and phylogenetic analysis identifies candidate members of a new picornavirus genus in terrestrial tortoise species. *Archives of Virology* 160, 811-6. https://doi.org/10.1007/s00705-014-2292-z

Marschang, R. E., Ihasz, K., Kugler, R., Lengyel, G., Feher, E., Marton, S., Banyai, K., Aqrawi, T. & Farkas, S. L. (2016). Development of a consensus reverse transcription PCR assay for the specific detection of tortoise picornaviruses. Journal of Veterinary Diagnostic Investigation 28, 309-14. https://doi.org/10.1177/1040638716628584

Ihasz, K., Marton, Sz., Feher E., Banyai, K. & Farkas, S. L. (2022). Genetic characterisation of a novel reptarenavirus detected in a dead pet red-

tailed boa (*Boa constrictor*). *Acta Veterinaria Hungarica* 70(1), 77-82. https://doi.org/10.1556/004.2022.00001