Final report to the postdoctoral research project 115519

Goose haemorrhagic polyomavirus (GHPV, A*nser anser polyomavirus 1* species, classified in the *Gammapolyomavirus* genus, *Polyomaviridae* family) may cause serious illness and sudden death of young geese. Although the disease, named haemorrhagic nephritis and enteritis of geese, is known for 50 years we hardly know anything about the virus. Isolation and culturing of GHPV is cumbersome, thus few data were published about the intracellular processes. The aim of the project was to gain information about genome variability and biological properties of GHPV, and to reveal potential hosts other than domestic goose and duck.

In the first year of the research (September 2015 – August 2016) we planned to investigate virus-host interactions using tissues of GHPV inoculated goslings and ducklings, and applying mRNA transcriptome sequencing. Usage of a virus isolate, free of co-infective microbes, is crucial for the experiments to avoid interactions in the host cells. In the absence of a virus isolate accepted protocol is testing of field samples with PCR specific for known microbes. As GHPV isolate was not available, field samples of geese and ducks (kidney and bursa of Fabricius) infected with GHPV were supplied by the Veterinary Diagnostic Directorate of the National Food Chain Safety Office, Hungary. First we performed circovirus detection in the specimens, but the circovirus positivity rate was very high, indicating the importance of pathogen testing and, if it is possible, usage of isolates.

In this period we received modified Muscovy duck origin cells permissive for GHPV infection (Mészáros et al., 2014), and we could make an attempt to isolate the virus for subsequent experiments. On the other hand, cell cultures could be used for transcriptome analysis avoiding animal experiments. Thus, instead of inoculation of geese and ducks with a probably ‘uncelar’ virus, the cell culture conditions were optimized and the cell culture was inoculated with homogenized field samples tested positive for GHPV. Virus replication was detected with PCR during passages, and cythopathic effects were visually checked.

Another plan of the first research year was collection of bird samples for screening with broad-spectrum polyomavirus specific PCR that may identify GHPV and other gammapolyomaviruses at the same time. Nucleic acid was extracted from 90 cloacal swab specimens of wild birds (including waterfowls) originally sent for influenza virus surveillance to the Veterinary Diagnostic Directorate of the National Food Chain Safety Office.

Polyomavirus could not been detected in any of those. Polyomavirus sequence was identified in the organ samples of a white-headed munia (*Lonchura maja*) suffering from liver failure, nephritis and myocarditis. The sample originated from the flock of a Hungarian aviculturist, where many birds of variable species died suddenly. The whole genome of the virus was amplified and the product was prepared for sequencing. The genome sequence (GenBank accession number KX756154) showed 91% identity with finch polyomaviruses. Based on the differences the International Committee on Taxonomy of Viruses named the virus Hungarian finch polyomavirus, and classified as a novel species, *Lonchura maja polyomavirus 1*, in the *Gammapolyomavirus* genus. The genome characterization was published in a peer-reviewed scientific journal (Marton et al., 2016).

Domestic waterfowl samples were collected as well. According to our plan, overlapping long PCR products were generated from the genome of the three novel GHPV strains, originating from samples of two geese and one duck, and DNA library was prepared from the mixed fragments of each strain. The genomes of the three strains were highly similar to the sequences determined previously in our laboratory and to the genomes deposited in the GenBank. The duck and goose origin genomes did not show significant differences.

Expression of the putative ORF-X gene was not previously proved in GHPV infected cell cultures or tissue samples of waterfowls. So far, the only evidence for ORF-X transcription was an experiment in which cloned GHPV genome was transfected into cultured cells (Johne and Müller, 2003). Although functional analysis of the putative ORF-X was planned to be executed in the second year of the research (September 2016 – August 2017), we performed preliminary examination in the first year. RNA was extracted from field samples and treated with DNase. Primers were designed for the amplification of the reverse transcribed, putative ORF-X mRNA. The PCR gave negative results. Of note, these samples were not collected and preserved to examine the RNAs. Therefore further studies were needed to elucidate whether ORF-X is transcribed from the GHPV genome.

In the period of September 2016 – April 2017 (in a shortened period because of medical reasons, childbearing), as part of the second research year, we continued the cell culture experiments started in the first research year. Unfortunately, GHPV could not be isolated with this method; cytopathic effects were lacking and the amount of PCR products, targeting the gene of the main capsid protein gene, decreased during the passages which means reduced or lacking viral genome replication in these cells. As we later saw, very high amount of GHPV (or

other polyomaviruses, described in the literature) may be necessary for successful inoculation of cultured cells, that could be the problem in case of field samples.

In the period May 2019 – August 2019 (in a shortened period), as part of the second research year, we continued analyses of the GHPV sequences. 23 complete genome sequences, characterized in our lab, were subjected to phylogenetic, evolutionary and selection constraint analysis along with the available GenBank data. The results indicated a low mean evolutionary rate in the order of 10^{-5} – 10^{-6} for the complete genome and extracted individual GHPV genes, and a dN/dS ratios <1. Recently, GHPV VP1 sequences were identified in tissue and fecal matter of wild birds, i.e. white stork, rock dove, European herring gull, common buzzard, grey heron and Eurasian jackdaw. These and further goose and duck origin VP1 sequences were retracted from the GeneBank and analysed. Accumulation of nt and aa mutations was detected in the sequences amplified from samples of wild birds. The sequence of a GHPV genome that originated from the sample of a diseased goose infected during the first outbreaks in 1969 was analysed as well but did not differ significantly from others. The results were published a peer-reviewed scientific journal (Kaszab et al., 2020).

In the last year of the project (September 2019 – August 2020) evolutionary and selection constraint calculations were also executed for other gammapolyomaviruses. Some of the covered sequences were determined in our laboratory. The results indicated that GHPV has one order lower mean evolutionary rate and more conservative genome compared to budgerigar fledgling disease virus, and finch and canary polyomaviruses. The highest dN/dS ratios (0.963 and 1.43, respectively) were calculated for the ORF-X of GHPV and VP4d of budgerigar fledgling disease virus. Based on sequence comparisons the ORF-X may have functions similar to the VP4/VP4d. The VP4 is a capsid component, while its splicing variant, the VP4d protein induces apoptosis in budgerigar fledgling disease infected cells. The significance of the mutations in these regions is unknown. A manuscript, that contains these results, was submitted to a peer-reviewed scientific journal (Kaszab et al., 2021a).

Unfortunately, the tested Muscovy duck cell culture did not to permit GHPV propagation. Furthermore, the field samples did not prove to be free of other pathogens/microbes. Thus the previously planned transcriptome and ORF-X analyses could not be properly executed in previous research periods. In the third year of the project a primary goose embryo fibroblast cell culture and a GHPV isolate (free of other pathogens; kindly provided by a co-operative lab) was successfully used for the investigation of viral RNA expression and replication, avoiding animal experiments. The GHPV infected cell culture was

sampled at 0, 3, 6, 9, 12, 24, 48, 72, 120 and 168 hours post infection, and the extracted DNA and RNA was prepared for viral replication and transcription analysis, respectively. There are no data in the literature about these courses for GHPV. In the study of Johne and Müller (2003) viral DNA was transfected into cultured cells and transcription of the small and large T antigen, the three capsid protein and the ORF-X genes was detected. However, they did not investigate temporal changes of the viral DNA and RNA copy number with quantitative methods. Although we originally planned to use sequencing for the transcription analyses, mRNA levels were measured for each predicted viral gene with hydrolysis probe based (TaqMan-type) quantitative real-time RT-PCR. With this method the measurements could be easily reproduced, with a lot of repeats and a cost-effective way. Furthermore, the low copy number mRNAs could be more reliably detected. We estimated the temporal appearance of viral GHPV mRNAs and the level/rate of those compared to each other. High-level transcription of the predicted ORF-X gene was proved along with expression of the early (LTA, STA) and late (VP1-VP3) genes known for all polyomaviruses. The splicing site of the ORF-X and LTA was identified with PCR and Sanger sequencing. The results were comparable with that of described for mammalian polyomaviruses. Previously, goose embryo fibroblast cells were unsuccessfully used for GHPV propagation, but here we demonstrated that GHPV may infect this cell types. However, similarly to mammalian polyomaviruses, high copy number of GHPV was necessary in the starting material. Accumulation of the LTA and STA mRNAs started before 9 hours post infection that was followed by the accumulation of the VP1-VP3 capsid protein encoding mRNAs and the ORF-X mRNA before 24 hours post infection. Commencement of the DNA replication was detected at 48 hours post infection implying that transcription is not strictly connected to DNA replication. Time-scaled phylogenetic analysis revealed introduction of novel GHPV strains, along with reintroduction and co-circulation of variable strains, at the same time at different towns in Hungary. A manuscript, that contains these results, was submitted to a peer-reviewed scientific journal (Kaszab et al., 2021b).

An additional manuscript was prepared for submission in which we collected and summarized the GHPV positivity data of the years 2005-2020 in Hungary and we described the results of polyomavirus testing of wild birds. The results showed fluctuation of the GHPV positive case number in domestic waterfowls with peaks in 2008, 2013 and 2017 (Kaszab et al., 2021c).

Although the unforeseen professional difficulties and the interrupted research periods made the executions harder and we had to change the research plan, the project provided valuable data about GHPV:

- GHPV was propagated in goose embryo fibroblast cells and transcription of the six main genes were first identified in infected cells. Both viral replication and transcription were pursued in time. To best of our knowledge, autovaccines are the only tools used for prevention of diseases caused by GHPV. These vaccines are produced in the same cell culture system in Hungary. Our results may help vaccine development.
- Complete sequences of novel GHPV strains were determined and sequence variability was investigated. Monitoring of circulating strains and potential hosts is of high importance for vaccine production as well.
- A novel species of the gammapolyomaviruses has been established, named *Lonchura maja polyomavirus 1* or Hungarian finch polyomavirus.

The result were summarized in 2 accepted manuscripts and 3 further manuscripts submitted to peer-reviewed scientific journals (2) or prepared for submission (1). A PhD student, Eszter Kaszab, joined the research project, and the result will be incorporated into the PhD thesis. Some results were included in the MSc thesis of Krisztina Bali.

References:

Johne R, Muller H. The genome of goose hemorrhagic polyomavirus, a new member of the proposed subgenus Avipolyomavirus. Virology 308:291–302, 2003. https://doi.org/10.1016/s0042-6822(02)00103-4

Kaszab E, Marton S, Dán Á, Farsang A, Bálint Á, Bányai K, **Fehér E**. Molecular epidemiology and phylodynamics of goose haemorrhagic polyomavirus. Transbound Emerg Dis 10(1):10301, 2020. https://doi.org/10.1111/tbed.13608

Kaszab E, Marton M, Erdélyi K, Bányai K, **Fehér E**. Genomic evolution of avian polyomaviruses with a focus on budgerigar fledgling disease virus. 2021a. (közlésre beküldött)

Kaszab E, Szabadi L, Kepner A, Bajnóczi P, Lengyel G, Bányai K, **Fehér E**. Viral gene expression profile of goose haemorrhagic polyomavirus in susceptible primary cell. 2021b. (közlésre beküldött)

Kaszab E, Bali K, Bálint Á, Bányai K, **Fehér E**. A lúd haemorrhagiás polyomavírus fertőzés jellemzése és hazai előfordulása. 2021c. (közlésre előkészítve)

Marton S, Erdélyi K, Dán Á, Bányai K, **Fehér E.** Complete Genome Sequence of a Variant Pyrrhula pyrrhula polyomavirus 1 Strain Isolated from White-Headed Munia (Lonchura maja). Genome Announc 4(6):e01172-16, 2016. https://doi.org/10.1128/genomeA.01172-16

Mészáros I, Tóth R, Bálint A, Dán A, Jordan I, Zádori Z. Propagation of viruses infecting waterfowl on continuous cell lines of Muscovy duck (Cairina moschata) origin. Avian Pathol 43(4):379-86, 2014 https://doi.org/10.1080/03079457.2014.939941