Szent István University
Postgraduate School of Veterinary Science

Comparative Genome Analysis of Avian Orthoreoviruses

Brief Summary of the PhD Thesis

Eszter Bayer-Dandár

2016

Supervisor and members of the committee:

Dr. Krisztián Bányai

Hungarian Academy of Sciences - Institute for Veterinary Medical Research
Supervisor

Dr. Mária Benkő

Hungarian Academy of Sciences - Institute for Veterinary Medical Research
Committee member

Dr. Ádám Dán

National Food Chain Safety Office, Molecular Biology Laboratory
Committee member

Introduction

The Reoviridae family – according to current classification – includes 15 genera, namely*: Cardoreovirus, Mimoreovirus, Orbivirus, Phytoreovirus, Rotavirus, Seadomavirus, Aquareovirus, Coltivirus, Cypovirus, Dinovemavirus, Fijivirus, Idnoreovirus, Mycoreovirus, Orthoreovirus, Oryzavirus* (http://www.ictvonline.org). These viruses not only infect humans and other vertebrates (fishes, reptiles, birds, mammals), but are also found in invertebrates (clams, insects), fungi and plants (Medveczky et al., 1998, Mertens et al., 2000, Schiff et al., 2007). Some virus genera are important in public and animal health.

Avian orthoreoviruses (ARV) belong to the genus *Orthoreovirus*. These viruses are of economic importance, capable of causing significant damage to poultry farms. Clinical manifestation of orthoreovirus infections is very diverse, including: arthritis, gastrointestinal malabsorption syndrome, myocarditis, pericarditis, hepatitis, osteoporosis and respiratory symptoms (Chappell et al., 2000, Jones 2000). However, in approximately 80% of the cases the infection does not produce any symptoms (Benavente & Martinez-Costas, 2007).

The pathogen can spread through air or fecal-oral transmission; there is no known vector (Chappell et al., 2000). Most poultry infections are age-dependent; adult animals are more resistant against the virus than younger specimens (Benavente & Martinez-Costas, 2007).

Despite the fact that avian orthoreoviruses can cause significant losses in poultry farms, there is a lack of data regarding their characteristics. To this day, we still know very little - among other things – about their genome and their evolutionary background, the diversity of individual virus strain genomes, and the occurrence of different virus strains in same and foreign host species.

Aims of the study

1. Genome sequencing of avian orthoreovirus strains derived from various host species originating from Hungary and international sources using different genome amplification and sequencing methods.
2. Describing the genomic organization of avian orthoreovirus strains and investigating the phylogenetic connections using bioinformatics tools.

Materials and Methods

Reovirus positive samples, virus isolates

The 21 virus strains used in our experiments have been provided for us in the forms of tissue pieces, faeces, tissue supernatants and lyophilizates by CEVA Phylaxia Zrt., the National Food Chain Safety Office (NÉBIH) and members of staff at the Haartman Institute-University of Helsinki.

We analyzed orthoreovirus strains derived from the following host species: a crow (*Corvus corne cornix*), a domestic goose (*Anser anser domestica*), a pheasant (*Phasianus colchicus*) and a partridge (*Perdix perdix*); furthermore, two Muscovy ducks (*Cairina moschata*), three domesticated turkeys (*Meleagris gallopavo*) and twelve chickens (*Gallus gallus domesticus*).

Cell lines

The orthoreovirus strains isolated from chicken, domesticated turkey, Muscovy ducks, domestic goose, partridge, pheasant and crow have been amplified in CEK (Chicken Embryo Kidney) primary culture, LMH (chicken-derived leghorn male hepatoma) and/or BHK-21 cell lines and the cell-free supernatant has been used to extract viral RNA.

Molecular methods

Viral RNA extraction has been performed using the TRIzol method or the QIAmp Viral RNA Mini Kit (QIAgen), according to the manufacturers' instructions.

The gene amplification method has been selected depending on the sequencing method. We used one- or two-step RT-PCR upon determining the sequence of the amplified gene- or genome segments using the traditional (Sanger) sequencing. In this case, we designed specific and degenerate primers. In other cases, when using high-throughput next generation sequencing, we amplified the genome segments using random RT-PCR. We used traditional

(Sanger) sequencing for determining the genomic sequence of three strains. In all other cases we used next generation sequencing (Ion-Torrent RGM).

We adapted the method described by Lambden et al. (1992) to determine sequences at the 5' and 3' ends of the genome segments.

Sequence analysis and phylogenetics

To analyze the sequencing data, we did sequence alignments (MULTALIN, GeneDoc). Genes were identified using BLASTN and BLASTX (Altschul et al., 1990) alignment search tools.

Data generated by the Ion Torrent sequencing machine was analyzed by the CLC Genomics Workbench (www.clcbio.com) software.

To generate phylogenetic trees, we used the MEGA 5 software (Tamura et al., 2011). We selected the best fit substitution model with the help of maximum likelihood (ML) tree-reconstruction algorithm and Bayesian information criterion. The reliability of the phylogenetic tree has been validated by bootstrap analysis repeated 500 times.

Results

Genomic composition of avian orthoreoviruses

The full-length size of the orthoreovirus genomes analyzed by us, depending on the strain, were between 22969-23579 base pairs (bps). Each genome segments started and ended with a short, highly conserved section. In the case of the crow orthoreovirus, they differed from the others on the 5' end, whereas the segment at the 3' end was typical in all of the analyzed strains.

The sizes of segments sequenced in full-length was the following: L1 3958-3993 bps, L2 3796-3829 bps, L3 3902-3907 bps, M1 2282-2288 bps, M2 2154-2158 bps, M3 1996-2026 bp, S2 1322-1327 bp, S3 1169-1203 bps, S4 1192-1201 bps. The length of untranslated regions was between 12-62 bps on the 5' ends and 30-98 on the 3' ends. The segments – as generally seen in orthoreoviruses – are usually monocistronic. The sizes of translated proteins were the following: λA 1293-1295 amino acid (aa), λB 1259 aa, λC 1283-1285 aa, μA 732-737 aa, μB 676-675 aa, μNS 630-639 aa, σA 416 aa, σB 367 aa, σNS 367-368 aa. The σC protein coding S1 and S4 segments were exceptions. In the case of chicken, turkey, partridge and crow orthoreoviruses, the S1 consisted of three partially overlapping ORFs; they coded products with the following lengths: p10: 96-99 aa, p17: 146-172 aa and σC: 326 aa. In the Muscovy duck and domestic goose S1-equivalent S4 segment was found to be polycistronic. These strains carried two partially overlapping ORFs on this gene segment, encoding the σC (269 aa) and the p10 (95 aa) proteins.

Phylogenetic analyses - Tvärminne avian virus (TVAV)

The comparison of main members of the genus *Orthoreovirus* revealed that the topology of the phylogenetic trees, even if slightly, varies by genes. At the same time, in general, the TVAV belonged to the same clade as avian orthoreoviruses and appeared to be closely related to some bat orthoreoviruses and in an especially close relationship with the unique sea lion-isolate SSRV strain. Sequencing data that are more detailed than S-class segments suggest that the SSRV and parrot reoviruses might belong to the same species and therefore in the case of these genes the TVAV also appears to be in close relationship with the psittacine orthoreovirus strains.

The analysis of sequence similarity numbers revealed that six out of ten genes shows less than 60% (nt) value and only 3 core proteins (λA, λB, σA), and the outer capsid forming μB gave >60%, but <75% nt identity with the reference sequences.

Phylogenetic analyses – Gallinaceous bird and waterfowl orthoreoviruses

Most of the analyzed orthoreovirus strains originated from poultry and species that are bred for hunting purposes. In my dissertation, I give a detailed comparative analysis of these strains by genes and host species. After summarizing the results in short, in general, the phylogenetic analyses put the orthoreoviruses of waterfowl (ducks, geese) and gallinaceous birds in two separate evolutionary lineages.

We identified two distinct clusters of waterfowl orthoreoviruses using phylogenetic analyses. One of the clusters consists of the classical disease causing Asian and European Muscovy Duck Reoviruses (MDRV), and, in the case of at least eight genes, also the strain that has been isolated from the Hungarian domestic goose. The other cluster contains the recently identified novel duck and goose strains (Novel MDRV, N-MDRV) described in China. We did not observe the segregation of strains by host species in the latter cluster. In the case of the μNS and σA coding genes, the goose strain from Hungary was a member of a separate cluster. To make things more complex, in case of some genes, the MDRV and N-MDRV strains were categorized in the opposite clusters. Looking at the μB coding gene, the N-MDRV strains showed closer genetic relationship with the gallinaceous bird orthoreoviruses than the classical disease causing MDRV strains' homologous gene segments.

The segregation by host species – partly due to the increased number of sequences and the more diverse geographical source – was more explicit in the case of gallinaceous bird orthoreoviruses. The chicken orthoreoviruses, depending on the genes, can be organized in two (or in some cases three) major clusters. One of the major clusters contains the strains that have spread through large geographical locations, whereas the other major cluster consisted almost exclusively of strains from Hungary. The sequence identity between clusters in most comparisons was between 75% nt and 80% aa.

We observed even lower numbers in the case of σC coding genes. Another interesting thing came to our attention while analyzing μB. We identified a formerly unknown heterogenic form of this gene in three strains derived from Hungarian chickens. This gene variant showed ~65% nt and ~69% aa identity with other chicken reoviruses. This similarity value is close to the sequence similarity threshold that is used when orthoreoviruses are categorized as separate species.

The analyzed turkey orthoreoviruses originated from two countries: USA and Hungary. Surprisingly, in most cases, the genetic identity of the strains collected from two separate geographical regions showed a high level (typically >90% for both nt and aa). The μB gene was an exception, in which case the turkey derived strains belonged to two independent evolutionary lineages. In general, we also observed that turkey orthoreoviruses were in

closer relationship to the dominant chicken orthoreovirus cluster, than with the genetic cluster containing most chicken orthoreoviruses found in Hungary.

The phylogenetic state of the genetically formerly not defined pheasant and partridge orthoreoviruses showed further interest. Albeit to this day we could not determine the complete genomic sequence of the pheasant strain, we gathered information about most of its genome. According to the phylogenetic calculations, the domestically isolated pheasant and partridge orthoreovirus strains shared the same cluster in the case of at least eight genes. These shared clusters, in certain cases, showed closer relationship with the domesticated turkey orthoreoviruses, whereas in other cases with the chicken orthoreoviruses. In the matter of the λA and μNS coding genes, either each or both strains represented an independent lineage, separate from the clades containing waterfowl derived and the chicken and turkey derived strains.

Discussion

Although in the past two-three years, the amount of data regarding the genomic organization of orthoreoviruses, therefore avian orthoreoviruses has increased rapidly, by the time I started my PhD work only two complete or almost complete ARV genome sequences were known. The data later generated by us and other research groups not only aided in developing a better understanding of the genomic organization of ARV strains, but also helped in understanding the evolutionary connections between the strains.

Genomic organization of orthoreoviruses

The structural organization of avian orthoreovirus genomes and genome segments did not differ significantly from the previously published results. The size of the genome averaged between 23-23.5 bps. We identified untranslated regions at the 5' and 3' ends of the 10 genome segments in the 21 analyzed ARV strains that had evolutionarily conserved sequences by genes and host species. Regarding the protein coding regions that are localized between the untranslated regions, we found that L1-L3, M1-M3 and S2-S4 segments were monocistronic in all cases, whereas the S1 segment is bi- or tricistronic. The number of coded proteins is 11 or 12; their size is proportional to the length of the genome segment (Benavente & Martinez-Costas, 2007). The polycistronic S1 genome segment was an exception, as it is the longest among the S-class segments, but the encoded σC protein is the shortest among structural proteins.

In the case of waterfowl, one important structural difference was between the S1 and its equivalent S4 segments. In the matter of classical MDRVs, the S4 segment is bicistronic with two partially overlapping ORFs, while the S1 segment of N-MDRVs – similarly to gallinaceous bird reovirus strains – is tricistronic (Costas et al., 2005; Schmulevitz & Duncan, 2000).

Tvärminne avian virus (TVAV) – a potentially new orthoreovirus species

The TVAV showed mild sequence-identity with the previously known strains that are classified as separate orthoreovirus species; upon individual gene comparison, in many cases these values were lower than the species demarcation threshold levels accepted by the ICTV. In conclusion, the sequence similarity data and the phylogenetic calculations suggests that TVAV can be considered as a new orthoreovirus species (Dandár et al., 2014).

The evolutionary history of gallinaceous bird and waterfowl orthoreoviruses

Our molecular studies revealed significant genetic similarities between the Classical MDRV and the novel N-MDRV variants. One exception is the μB coding gene, responsible for virus penetration, in which case the N-MDRVs are in close relationship with the chicken derived ARVs. The waterfowl derived orthoreovirus strains investigated by us – according to their molecular structure – can be classified as Classical MDRVs. The analyses showed that in case of the strains derived from two Muscovy ducks and a domestic goose, similarly to a number of MDRV and N-MDRV isolates from China, belong to the ARV virus species. However, on the phylogenetic trees, these strains – with the exception of the previously mentioned μB – were in all cases separated from the chicken and turkey derived orthoreovirus species.

Phylogenetic analysis of the gallinaceous bird orthoreoviruses revealed that although the turkey derived strains form separate branches inside the evolutionary trees, these branches often belong to the same cluster along with chicken derived orthoreovirus strains. One explanation could be that chicken and turkey are taxonomically related to each other and these species were kept together in poultry yards in the past; therefore, their orthoreoviruses probably had a common ancestor and adapted to their currently known host species through various evolutionary mechanisms.

Opposite of the relatively small genetic variability observed in turkey derived orthoreovirus strains, the genes of chicken orthoreovirus strains could be organized to many large clusters. In some of the major clusters, we can only find strains originating from Hungary. The long branches in the topology of some phylogenetic trees suggested the amplification of point mutations, whereas the heterogenic distribution of some strains suggested reassortment.

Analysis of the M2 segment made our observations regarding the importance of reassortment less clear. Three chicken derived orthoreovirus strains clustered together as a separate branch of the phylogenetic tree. Due to the not so close genetic relationship, we assume that the M2 segment was a component of a sole orthoreovirus species derived from a foreign host species, which was later inserted into the homologous chicken orthoreovirus genome through reassortment. If our assumption is correct, that could mean that in certain cases it is possible to exchange genes between different orthoreovirus species.

Phylogenetic analysis of the homologous pheasant and partridge derived orthoreovirus strains currently only leaves us with limited options in exploring the evolutionary background of these strains. At the same time, genomes of the analyzed strains showed a mixed, part chicken and part turkey derived, mosaic genetic composition that was further complicated by the acquisition of genes with standalone phylogenetic ancestry, during the evolution of these strains.

New scientific achievements and conclusions

1. We were the first to determine the genomic sequence of the hooded crow *(Corvus corne cornix)* derived orthoreovirus. After phylogenetic analyses, we suggested the creation of the new reovirus species *Corvid reovirus* (CRV).
2. We were the first to determine the genomic sequence of the Muscovy duck *(Cairina moschata)* and domestic goose *(Anser anser domestica)* derived virus strain. We found that all three strains belong to the MDRVs with classical genome organization, however we pointed out the possibility of reassortment between Classical MDRV and N-MDRV strains; disproving the geographical separation hypothesis of the two pathotypes.
3. We were the first to analyze the pheasant *(Phasianus colchicus)* derived ARV strain with molecular methods. Furthermore, we were the first to determine the genomic sequence of a partridge (*Perdix perdix)* derived orthoreovirus isolate. Due to their roles as reservoirs, we suspect that these host species are important in the construction/maintenance of the genetic composition of turkey and chicken reoviruses.
4. We were the first to determine the protein coding sequences of the domesticated turkey *(Meleagris gallopavo)* derived orthoreovirus originating from Europe, pointing out the high genetic similarity between strains isolated from different geographical locations.
5. We were the first to determine the genomic sequences of orthoreovirus isolates that were derived from chicken *(Gallus gallus domesticus)* without tenosynovitis.
6. After the analysis of a couple of dozen orthoreovirus strains that we investigated phylogenetically, we concluded that there are conserved patterns in the genetic composition of orthoreoviruses in different host species, however the reassortment with strains derived from a foreign host species allows the insertion of new gene variants.

Scientific publications

Published articles in peer-reviewed scientific journals with impact factor

Kugler R., Dandár E., Fehér E., Jakab F., Mató T., Palya V., Bányai K., Farkas S. L.: **Phylogenetic analysis of a novel reassortant orthoreovirus strain detected in partridge (Perdix perdix),** Virus Res., 215. 99-103, 2016.
IF: 2,324

Dandár E., Fehér E., Bálint Á., Kisfali P., Melegh B., Mató T., Kecskeméti S., Palya V., Bányai K., Farkas S.L.: **Genome sequences of three turkey orthoreovirus strains isolated in Hungary,** Genome Announc., 3. e01333-15, 2014.
IF: -

Farkas S.L., Dandár E., Marton S., Fehér E., Oldal M., Jakab F., Mató T., Palya V., Bányai K.: **Detection of shared genes among Asian and European waterfowl reoviruses in the whole genome constellations,** Infect. Genet. Evol., 28. 55-57, 2014.
IF: 3,264

Dandár E., Farkas S.L., Marton S., Oldal M., Jakab F., Mató T., Palya V., Bányai K.: **The complete genome sequence of a European goose reovirus strain,** Arch. Virol., 8. 2165-2169, 2014.
IF: 2,030

Dandár E., Huhtamo E., Farkas S.L., Oldal M., Jakab F., Vapalahti O., Bányai K.: **Complete genome analysis identifies Tvärminne avian virus as a candidate new species within the genus Orthoreovirus,** J. Gen. Virol., 95. 898-904, 2014.
IF: 3,127

Dandár E., Bálint Á., Kecskeméti S., Szentpáli-Gavallér K., Kisfali P., Melegh B., Farkas S.L., Bányai K.: **Detection and characterization of a divergent avian reovirus strain from a broiler chicken with central nervous system disease**, Arch. Virol., 158. 2583-2588, 2013.
IF: 2,030

Dandár E., Borzák R., Bányai K., Farkas S. L.: **Hüllők, madarak és emlősök orthoreovírus okozta megbetegdései – Irodalmi áttekintés,** Magy. Állatorvosok., 134. 564-573, 2012.
IF:0,146

Bányai K., Dandár E., Dorsey, K.M., Mató T., Palya V.: **The genomic constellation of a novel avian orthoreovirus strain associated with runting-stunting syndrome in broilers**, Virus Genes, 42. 82-89, 2011.
IF: 1,170

Abstracts or proceedings published in international conference handbooks related to the doctoral research subject

Dandár E., Bálint Á., Kecskeméti S., Szentpáli-Gavallér K., Farkas S.L., Bányai K.: **Detection and characterization of an avian reovirus strain with divergent genes from central nervous system disease of broiler chicken**, Federation of European Microbiological Societies (FEMS), Lipcse, 2013.

Bányai K., Dandár E., Dorsey, K. M., Mató T., Palya V.: **Novel avian orthoreovirus strain associated with runting-stunting syndrome in broilers,** International Meeting on Emerging Diseases (IMED), Bécs, 2011.

Abstracts or proceedings published in international conference handbooks not related to the doctoral research subject

Dandár E., Doszpoly A., Jánoska M., Heltai M., Szabó L., Benkő M.: **PCR Screening of Mammalian predators *(Carnivora)* for adeno- and herpesviruses**, 8th International Congress of Veterinary Virology, In: *Proceedings of the ESVV 8th Int Congr Vet Virol.* Szerk.: Benkő M., Harrach B. Budapest, p. 226., 2009.

László B., Papp H., Dandár E., Deák J., Gray J., Iturriza-Gomara, M., Jakab F., Juhász Á., Kovács J., Kónya J., Lengyel Gy., Martella, V., Mészáros J., Mészner J., Mihály I., Molnár P., Nyúl Z., Pátri L., Puskás E., Schneider F., Tóth A., Tóth E., Szűcs Gy., Bányai K.: **Emerging rotavirus strains, 2007-2010, Hungary**, 4th European Rotavirus Biology Meeting, Altafiumara, 2011.

László B., Papp H., Dandár E., Deák J., Gray, J., Iturriza-Gomara, M., Jakab F., Juhász Á., Kovács J., Kónya J., Lengyel Gy., Martella, V., Mészáros J., Mészner J., Mihály I., Molnár P., Nyúl Z., Pátri L., Puskás E., Schneider F., Tóth A., Tóth E., Szűcs Gy., Bányai K.: **Emerging rotavirus strains, 2007-2010, Hungary**, International Meeting on Emerging Diseases (IMED), Bécs, 2011

Scientific publications not related to the doctoral research subject

Bayer-Dandár E, Kassa Cs.: **EBV fertőzések őssejt-transzplantáltakban: klinikum, diagnosztika és terápiás lehetőségek,** Focus Medicinae, 17. 16-21, 2015.
IF:-

László B., Kónya J., Dandár E., Deák J., Farkas Á., Gray, J., Grósz G., Iturriza-Gomara, M., Jakab F., Juhász Á., Kisfali P., Kovács J., Lengyel G., Martella, V., Melegh B., Mészáros J., Molnár P., Nyúl Z., Papp H., Pátri L., Puskás E., Sántha I., Schneider F., Szomor K., Tóth A., Tóth E., Szűcs G., Bányai K.: **Surveillance of human rotaviruses in 2007-2011, Hungary: exploring the genetic relatedness between vaccine and field strains**, J. Clin. Virol., 55. 140-146, 2012.
IF: 3,287

Sós E., Molnár V., Dandár E., Bálint Á., Bakonyi T.: **Szerológiai vizsgálatok hazai túzok-(Otis tarda) állományokban,** Magy. Állatorvosok., 134. 361-365, 2012.
IF: 0,146

Dandár E., Szabó L., Heltai M., Doszpoly A.: **Adeno- és herpeszvírusok előfordulásának felmérése ragadozók (Carnivora) mintáinak PCR-es vizsgálatával,** Magy. Állatorvosok., 132. 302-308, 2010.
IF: 0,300

Bányai K., Papp H., Dandár E., Molnár P., Mihály I., van Ranst, M., Martella, V., Matthijnssens, J.: **Whole genome sequencing and phylogenetic analysis of a zoonotic human G8P[14] rotavirus strain,** Infect. Genet. Evol., 10. 1140-1444, 2010.
IF: 2,792

Acknowledgment

I wish to thank my supervisor, Dr. Krisztián Bányai, for supporting my ideas, my research and for introducing me to the secrets of virology. His theoretical and practical guidance, his critical remarks have greatly contributed to this work.

I also wish to thank Dr. Szilvia Farkas Dr. Enikő Fehér, who worked hard on getting rid of a number of errors from the manuscript during the writing.

I thank the current and past members of the *Discovery of new diseases* group: Anett Bartókné Horváth, Réka Borzák, Hajnalka Papp, Dr. Eszter Kovács, Dr. Szilvia Marton, Dr. Boglárka Sellyei, Katalin Ihász and Renáta Dóró for supporting my everyday work and with their help I could overcome the everyday challanges.

I am also grateful to Dr. Ádám Bálint, Dr. Ádám Dán, Dr. Katalin Szentpáli-Gavallér, Dr. Sándor Kecskeméti, Judit Michna, Dr. Ákos Thuma, Dr. Ferenc Jakab, Dr. Tamás Mató, Dr. Vilmos Palya, and Dr. Eili Huhtamo and Dr. Olli Vapalahti for all the help, that made this work possible.

I thank my family and friends, who always believed in me and supported and encouraged me over the past years, which helped greatly overcome the tasks before me. I am especially grateful to my mother, who has been restlessly reading over and over again the manuscript in preparation.

Finally, I wish to thank my husband, László Bayer, for enduring the ordeals of the final months with infinite patience and understanding and for helping me get rid of my anxiety and concerns of failure many times.

This dissertation was supported by the OTKA grant no. K108727.