

WORKSHOP

Molecular tracing of viral diseases in aquaculture

S. S. Mikkelsen¹, L. Bigarré², B. Bang Jensen³,
A. B. Kristoffersen³, P. A. Jansen³, V. Panzarin⁴,
S. C. Bayliss⁵, J. C. Avarre⁶ and N. J. Olesen^{1*}

¹Section for Fish Diseases, European Union Reference Laboratory for Fish Diseases, National Veterinary Institute, Technical University of Denmark, Frederiksberg C, Denmark; ²Fish Viral Pathologies Unit, Laboratoire de Ploufragan/Plouzané, ANSES, Plouzané, France; ³Section for Epidemiology, Norwegian Veterinary Institute, Oslo, Norway; ⁴Division of Comparative Biomedical Sciences, Research & Innovation Department, OIE Reference Laboratory for Viral Encephalopathy and Retinopathy, Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro, Italy; ⁵Department of Biology & Biochemistry, University of Bath, Bath, United Kingdom; ⁶Institut des sciences de l'évolution de Montpellier, UMR 226 IRD-CNRS-UM2, Montpellier, France

Introduction

The workshop focused on new findings and tools for molecular tracing of viral diseases in aquaculture. The primary aim of the workshop was to bring colleagues involved in molecular biology, bioinformatics and epidemiology together in order to discuss improved methods for risk analysis and prevention of spread of serious diseases in aquaculture

One of the goals of molecular tracing of fish viruses is to suggest scenarios for the introduction and dissemination of a virus of unknown origin, by linking genomic sequences with epidemiological data. Starting from a virus in a sample, two steps are necessary for reaching this goal: first the identification of a molecular target (a gene, a non-coding region or similar) and, second, querying databases enriched with

large-scale data (Genbank, Fishpathogens, others). Both steps require continuous efforts from all of the scientific community to find the best molecular markers in a viral genome (particularly for new large DNA viruses) and to feed the databases with nucleic, geographic, time-related and biological data.

The MOLTRAQ (Molecular tracing of viral pathogens in aquaculture, EMIDA-ERA NET) project has contributed to both efforts on various models of economic importance, such as fish rhabdoviruses, alloherpesviruses and oyster malacoherpesviruses. To disseminate the results from this project a workshop was organized with the goal to provide participants with knowledge on the molecular and genetic tools available for providing molecular epidemiologi-

* Corresponding author's e-mail: njol@vet.dtu.dk

cal analysis and tracing of disease outbreaks in aquatic animals and to share results from the newly finished project. The workshop also provided an opportunity for the newly established WGS-AQUA.NET to share information on their project.

Molecular tracing of viral pathogens in aquaculture, a review presented by Dr. Laurent Bigarré, ANSES, France.

The aim of the MOLTRAQ project was to bring researchers working in different areas within fish health like epidemiologists, molecular biologists and bioinformaticians together to provide new information on disease control and prevention of aquatic diseases. This was done successfully during the project in a number of ways.

First of all, a large repository of genetic and epidemiologic data on fish and oyster viruses including Viral haemorrhagic Septicaemia Virus (VHSV), Cyprinid Herpesvirus 3 (CyHV-3), Eel Virus European X (EVEX) and Oyster Herpesvirus 1 (OsHV-1) were established and uploaded to Genbank and www.fishpathogens.eu, including new information on hundreds of isolates (Bellec et al., 2014, Cieslak et al., 2016, Mikkelsen et al., 2016).

A new method for capturing and sequencing large DNA viruses was implemented, leading to the full-genome sequencing of 19 CyHV-3 isolates and the first full-length sequencing of a CyHV-3 "variant" (Hamoumi et al., Submitted). Furthermore, 12 full-length OsHV-1 genomes were obtained and 10 microsatellites discovered (Renault et al., 2014). This genetic information provided a better vision of the variability of CyHV-3 and OsHV-1, by revealing conserved

and variable regions useful for diagnostics and genotyping respectively. They will be essential to set up molecular tools for the rapid genetic identification of strains in the future.

For rhabdoviruses, repositories were enriched with epidemiological information and fully sequenced G-genes of hundreds of new isolates of infectious hematopoietic necrosis virus (IHNV), VHSV and Eel rhabdovirus from several European countries. The extensive sequencing of viral isolates has provided new information that has been used for molecular tracing of the respective viruses, both on a national and a European scale, for creating new simulation models for risk assessment and intervention strategies and for finding new biomarkers (Bang Jensen et al., 2014).

Furthermore, a new database for betanodaviruses has been established at www.fishpathogens.eu with a repository of epidemiologic and sequence information of new isolates (Mikkelsen et al., 2015). Already established databases for VHSV and IHNV have been updated and work has been started on a new Salmonid alpha virus (SAV) database (Jonstrup et al., 2009, 2010).

All in all, the MOLTRAQ project has succeeded in developing new tools for the genetic characterization of large DNA viruses, establishing new models for tracing the spread of aquatic diseases and creating new and large repositories of information about a large variety of important aquatic viruses for further development of new, more specific assays and simulation models.

Scenario simulation models for control options by Dr. Peder A. Jansen, NVI, Norway

In Norway, the fish farmers experience large problems with SAV which causes outbreaks in about 120-140 fish farms every year. A stochastic model was created within the MOLTRAQ project that integrates spatio-temporal epidemiological data and which could include phylo-geographic information to help predict the outcome of intervention strategies (Aldrin et al., 2015). The simulation model showed that the seaway distance between fish farms was one of the most important factors in the transmission pathways whereas other factors like previous infections at the farm were less related to the consequence. Simulation models like this can be used to investigate the outcome of various strategies like mandatory slaughtering of fish after an outbreak or re-location of fish farms. Subsequently, we have coupled the model with a cost-benefit analysis, to help determine the most efficient and economically beneficial strategies to implement (Manuscript in preparation).

Use of sequence data in epidemiological analysis by Dr Anja B. Kristoffersen, NVI, Norway

While the use of molecular data is becoming more common in epidemiological studies, it is important to remember to be aware of the statistical models being used. When calculating genetic distances, a distance matrix is being obtained. The distances in this matrix are dependent which means that normal statistical methods that rely on independence cannot be used. Instead, it is possible to use permutations as developed by Mantel in 1967. The Mantel framework was originally developed

for correlation which is a univariate approach, but the framework can be extended and used within many different statistical methods. In this talk, the methodology was presented, and an example given where the Mantel framework is used together with logistic regression, to trace transmission patterns of Infectious Salmon Anaemia virus (ISAV).

Molecular epidemiology and evolution of VHSV and IHNV in Italy by Dr. Valentina Panzarin, ISZVe, Italy

Despite the application of eradication programs, Italy still struggles with outbreaks of VHSV and IHNV every year at several trout farms, leading to great economic losses. A study was undertaken to investigate the epidemiology and evolution of VHSV and IHNV in Italy during the past years. Representative isolates from outbreaks on different fish farms were sequenced on the G-gene. The sequences were then aligned and phylogenetic and evolutionary analyses were carried out.

The phylogenetic analysis of the Italian VHSV strains showed that they belong to the freshwater genotype Ia, while IHNV strains are all situated in the European (E) genogroup.

Our data also suggest that Italian VHSV and IHNV possess different evolutionary dynamics. Sequence information related to each viral isolate was combined with epidemiological data, and different scenarios describing the distribution and the spread of VHSV and IHNV in Italy were identified.

WGS-Aqua.NET: capacity building for the widespread adoption of whole genome sequencing (wgs) for the molecular epidemiology of aquaculture pathogens “Investigating the diversity of *Flavobacterium psychrophilum* using whole genome sequencing.” Presented by Dr Sion Bayliss, UoB, UK

Whole genome sequencing of pathogens is beginning to be routinely applied for identification and tracing of pathogens in a manner that was previously the province of individual genes and other molecular markers. This necessitates the identification of tools and standardized methods to analyse WGS data and the establishment of centralized databases to facilitate investigation.

WGS-AQUA.NET is NERC/BBSRC funded, capacity-building consortium of academics and stakeholders, led by the University of Bath, the University of Stirling and Cefas (UK), that aims to implement whole genome sequencing for the management of aquaculture pathogens. The project will establish databases for a number of bacterial and viral aquaculture pathogens that will allow for epidemiological analysis from raw WGS data. Databases for Koi Herpes Virus (KHV), *Flavobacterium psychrophilum* and *Vibrio anguillarum* are currently being established. Reference collections, that will be made publicly available, are being generated for these and other aquaculture pathogens. The database will allow for users to upload their own sequences, publicly or privately, for comparison to these reference collections.

Discussion

As one of the goals of MOLTRAQ as well as WGS-AQUA.NET was and is to establish da-

tabases for storing genetic and epidemiological information on isolates, it was discussed how to keep the databases running. It is often seen that databases are established *via* project funding, but as soon as the funding runs out, the databases become inactive or are even removed. For Fishpathogens.eu, the solution has been to link the database to the European Reference Laboratory for Fish Diseases as a task to keep it running. Specific funding assures server space and that development on it is continued.

An additional deliverable of WGS-AQUA.NET is to provide guidelines for the analyses of bacterial whole genome data for aquaculture disease management.

Another subject brought up in the discussion was the need for accurate and large datasets to create models for disease control. Models have been used with success for terrestrial viruses including avian influenza, but models are also only as good as the data they are based on. Information on aquatic viruses is plentiful when it comes to optimal infection conditions and virus dispersal, but specific fish farm data like stocking times, herd sizes and transport and sale of fish are often lacking.

These data can be difficult to obtain as there can be legislative restrictions which prevent data from being recorded or released. It can also be a problem for the industry if they are on the stock market as information about outbreaks might be considered confidential information as it can affect the stock price. Furthermore, farmers are worried about the consequences of releasing these data to the public. It was agreed that it was imperative to get access to more data in order to establish more accurate models for

disease control and one way of doing this was by working with the aquaculture industry and the farmers to show them what can be gained by these models so they can see the value of sharing data.

Another topic during the discussion was the use of tools for modelling and data analysis. As of now there is a lack of standardization of tools and methods and some are only useful for experts and others can be hard to compare. Therefore, it would be valuable to try and standardize the use of tools and to create guidelines for the more inexperienced users, which hopefully also could lead to more information and datasets being shared. WGS-AQUA.NET is working on some guidelines which will be published on their platform but these will most likely only concern bacteria.

It is also important to develop new tools for data analysis, particularly for larger viruses. It is important to look at all kinds of variations to distinguish viruses, including variable number tandem repeats (VNTR) and not to throw away too much information when large dataset must be trimmed in order to reduce size of data for storage. Another subject to consider is how to handle quasi-species and how to trace these.

Conclusion

The MOLTRAQ project has secured large amounts of data on several of the economically important fish viruses like the rhabdoviruses and the alloherpesviruses as well as the malacoherpesviruses from oysters, which will be available to the public and which have given a fundament for molecular epidemiological studies.

In conclusion molecular tracing will need more investments from all the stakeholders to face the huge sanitary issues provoked by the spread of viruses *via* international trade and the domestication of new fish species. Furthermore there is a need for up-to-date databases containing both genetic and epidemiological data as well as harmonized and standardized tools for data analysis and for these databases and tools to be available in perpetuity.

Acknowledgements

We want to thank all who attended the workshop and participated in the discussion. We would also like to thank EMIDA-ERA for funding.

References

- Aldrin M, Huseby RB and Jansen PA (2015). Space-time modelling of the spread of pancreas disease (PD) within and between Norwegian marine salmonid farms. *Preventive Veterinary Medicine* **121**, 132-141.
- Bang Jensen, B, Ersbøll AK, Korsholm H, Skall HF and Olesen NJ (2014). Spatio-temporal risk factors for viral haemorrhagic septicaemia (VHS) in Danish aquaculture. *Diseases of Aquatic Organisms* **109**, 87-97.
- Bellec L, Cabon J, Bergmann S, de Boissésion C, Engelsma M, Haenen, O, Morin T, Olesen NJ, Schuetze H, Toffan A, Way K and Bigarre L (2014). Evolutionary dynamics and genetic diversity from three genes of Anguillid rhabdovirus. *Journal of General Virology* **95**, 2390-2401.
- Cieslak M, Mikkelsen SS, Skall HF, Baud M, Diserens N, Engelsma MY, Haenen OLM, Panzarin V, Mousakhani S, Wahli T, Olesen NJ and Schütze H (2016). Epidemiology of Viral Hemorrhagic Septicemia in Farmed Rainbow Trout by Viral Phylodynamics. *PLoS Pathogens* (Submitted).
- Jonstrup SP, Gray T, Kahns S, Skall HF, Snow

- M and Olesen N J (2009). FishPathogens.eu/vhsv: a user-friendly viral haemorrhagic septicaemia virus isolate and sequence database. *Journal of Fish Diseases* **32**, 925-929.
- Jonstrup SP, Schuetze H, Kurath G, Gray T, Jensen BB and Olesen NJ (2010). An isolate and sequence database of infectious haematopoietic necrosis virus (IHNV). *Journal of Fish Diseases* **33**, 469-471.
- Mikkelsen SS, Panzarin V, Jonstrup SP, Bigarre L, Baud M, Gray T, Agapow PM and Olesen NJ (2015). Fishpathogens.eu/noda: a free and handy online platform for Betanodavirus targeted research and data sharing. *Journal of Fish Diseases* **38**, 755-760.
- Mikkelsen SS, Panzarin V, Fusaro A, Schuetze H and Olesen NJ (2016). Molecular tracing of VHS outbreaks in Denmark between 1982 and 2009. *Frontiers in Microbiology* (In preparation).
- Renault T, Tchaleu G, Faury N, Moreau P, Segarra A, Barbosa-Solomieu V and Lapegue S (2014). Genotyping of a microsatellite locus to differentiate clinical Ostreid herpesvirus 1 specimens. *Veterinary Research* **45**, 3.