

Technical University of Denmark



FishPathogens.eu/vhsv: a user-friendly viral haemorrhagic septicaemia virus isolate and sequence database

Jonstrup, Søren Peter; Gray, Tanya; Kahns, Søren; Skall, Helle Frank; Snow, Micheal; Olesen, Niels Jørgen

Published in:
Journal of Fish Diseases

Link to article, DOI:
[10.1111/j.1365-2761.2009.01073.x](https://doi.org/10.1111/j.1365-2761.2009.01073.x)

Publication date:
2009

Document Version
Early version, also known as pre-print

[Link back to DTU Orbit](#)

Citation (APA):
Jonstrup, S. P., Gray, T., Kahns, S., Skall, H. F., Snow, M., & Olesen, N. J. (2009). FishPathogens.eu/vhsv: a user-friendly viral haemorrhagic septicaemia virus isolate and sequence database. Journal of Fish Diseases. DOI: 10.1111/j.1365-2761.2009.01073.x

DTU Library
Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FishPathogens.eu/vhsv: A user-friendly Viral Haemorrhagic Septicaemia Virus (VHSV) isolate and sequence database.

Søren Peter Jonstrup¹, Tanya Gray², Søren Kahns¹, Helle Frank Skall¹, Mike Snow³ and Niels Jørgen Olesen¹

1 Community Reference Laboratory for Fish Diseases, Section for Fish Diseases, Division of Poultry, Fish and Fur Animals, National Veterinary Institute, Technical University of Denmark, Hangøvej 2, DK-8200 Århus N, Denmark

2 Symantix Ltd, Wiltshire SN4 9BN, UK

3 Fisheries Research Services (FRS) Marine Laboratory, PO Box 101, 375 Victoria Road, Aberdeen AB11 9DB, UK

Correspondence to Søren Peter Jonstrup, Community Reference Laboratory for Fish Diseases, Section for Fish Diseases, Division of Poultry, Fish and Fur Animals, National Veterinary Institute, Technical University of Denmark, Hangøvej 2, DK-8200 Århus N, Denmark, spjo@vet.dtu.dk

Runing title: FishPathogens.eu/vhsv

Abstract

A database has been created, www.FishPathogens.eu, with the aim of providing a single repository for collating important information on significant pathogens of aquaculture, relevant to their control and management. This database will be developed, maintained and managed as part of the European Community Reference Laboratory for Fish Diseases function. This concept has been initially developed for VHSV and will be extended in future to include information on other significant aquaculture pathogens. Information included for each isolate comprises sequence, geographic origin, host origin and useful key literature. Various search mechanisms make it easy to find specific groups of isolates. Search results can be presented in several different ways including table based, map based, and graph based outputs. When retrieving sequences, the user is given freedom to obtain data from any selected part of the genome of interest. The output of the sequence search can be readily retrieved as a FASTA file ready to be imported into a sequence alignment tool of choice, facilitating further molecular epidemiological study.

Keywords: Database, Viral Haemorrhagic Septicaemia Virus, Fish pathogens.

Introduction

We live in a world where the amount of information available is enormous. The internet has provided an excellent platform to exchange this information. In order to keep track of the available knowledge, databases are needed to collect, store, and sort it. The field of biology is no exception. An overwhelming amount of species are known, millions of sequences are available, and the number of scientific publications is abundant just to mention a few examples. Broad databases are provided from big institutions such as National Institutes of Health to keep track of this information. The information available in broad databases can be overwhelming (No Author 2007). Therefore smaller niche databases are suitable for providing valuable and easily accessible data on a very specific subject. Several small databases have over time become a vital tool for people interested in specific topics (Galperin 2008).

Viral Haemorrhagic Septicaemia Virus (VHSV) is an important fish virus that has caused several large scale fish kills in both fresh and saltwater fish (Skall, Olesen & Møllergaard 2005). The virus is an enveloped negative-stranded RNA virus belonging to the family *Rhabdoviridae* and the genus *Novirhabdovirus*. Its genome consists of around 11000 bases and codes for six proteins: The nucleoprotein N, the phosphoprotein P, the matrix protein M, the glycoprotein G, the non-structural viral protein Nv, and the polymerase L arranged in the order 3'-N-P-M-G-Nv-L-5' (SchützeMundt & Mettenleiter 1999). Several isolates of VHSV have been sequenced and four genotypes of the virus are recognized (Genotypes I-IV) (Benmansour, Basurco, Monnier, Vende, Winton & de Kinkelin 1997, Einer-Jensen, Ahrens, Forsberg & Lorenzen 2004, Einer-Jensen, Ahrens & Lorenzen 2005, Nishizawa, Iida, Takano, Isshiki, Nakajima & Muroga 2002, Snow, Bain, Black, Taupin, Cunningham, King, Skall & Raynard 2004, Snow, Cunningham, Melvin & Kurath 1999, Stone, Way & Dixon 1997, Thiery, de Boisseson, Jeffroy, Castric, de Kinkelin & Benmansour 2002). Genotype I consists of European freshwater VHSV isolates and a group of marine isolates from the Baltic Sea (Skall *et al.* 2005). This genotype has further been divided into five subtypes called Ia-Ie (Einer-Jensen *et al.* 2004). Genotype II and III are marine European isolates while genotype IV is found in North America and East Asia (Skall *et al.* 2005). Genotype IV has further been divided into two subtypes named IVa and IVb (Elsayed, Faisal, Thomas, Whelan, Batts & Winton 2006).

The growing number of VHSV isolates and their sequences is quite overwhelming. Therefore we have made a database to keep track of isolates and their sequences. Being the European Community Reference Laboratory for Fish Diseases and the OIE reference laboratory for VHSV, as well as having a long tradition working with VHSV we have access to a huge amount of data. Our goal is to include as many VHSV isolates as possible. This will be achieved by entering all the data we have to the database as well as relying on partners in other laboratories to enter their data. All publicly available data in the database has been reviewed by a pathogen expert ensuring a high quality of the data. The VHSV database is the first fish pathogen database available at <http://www.fishpathogens.eu>, but we are already in the process of establishing databases of other important fish pathogens.

Overview of the VHSV database

The aim of the VHSV database is to provide a comprehensive and searchable database of VHSV isolates and their characteristics. In order to achieve this we will encourage laboratories from all around the world to submit data on VHSV isolated in their laboratory. New isolates and sequences can easily be added by filling out forms on the website. Available information on sequences, geography, host, date of isolation, useful literature, and much more is added for each isolate. Some information can be given as free text, however in several cases information is given by choosing from a list of defined values in order to make the database easy to search. Manuals for uploading are available at the website. It is also possible to contact us to get help and if there is an interest we can arrange database training courses.

We expect that many will also use the website to search the accessible data. The website includes browsable lists of all isolates and sequences stored in the database. However, in order to optimize the possibility for locating specific isolates two alternative search functions are available on the website. It is possible to search the isolates visually using Google MapsTM or Google EarthTM. In order to be able to include all isolates when doing map based searches we encourage all uploaders of data to provide geographical locations. It is also possible to search data using a text based search. This method of searching gives freedom to assign several search criteria in order to specify your search. This includes genotype, country, name of host, host aquatic environment, a time period for isolation or uploading of reports, isolates located within a given radius of a specified location, and many more criteria. When searching for sequences, a BLAST based function allows users to obtain sequences from any part of the genome defined by the user. Making it possible to retrieve sequences from a gene region of choice the database does not force people to use one specific method of genotyping. In order to help users become familiar with all the different search possibilities, manuals are provided at the website.

When a search result is obtained there are several ways data can be displayed (figure 1A). If required full isolate and sequence reports can be viewed. However, if the number of isolates in the search is large a summary table can be chosen for displaying data (figure 1B). Simple graph and statistic tools will provide an overview of the distribution of several isolate characteristics within the search (figure 1C) and map based tools can be used to give an overview of the retrieved isolates (figure 1D). If sequence reports are available in the search, FASTA files can be downloaded and used directly in most sequence alignment programs.

Three types of users can draw on the FishPathogens.eu/vhsv database: Unregistered users, registered users, and pathogen experts. The unregistered user can search the uploaded data as well as submit new data. If one wants to submit data we recommend people to register. This will add a "my page" to the website from where the registered users can add reports as well as browse and edit their own reports. As a registered user it is also possible to restrict access to reports not ready for public access. The last type of user is the appointed pathogen experts who can add data themselves but will also review data added by other users to ensure that added data is of high quality.

The FishPathogens.eu/vhsv database has been developed using LAMP (Linux, Apache, MySQL, and PHP), a software stack composed of free, open source software. BL2seq available from NCBI is used to perform sequence alignments (<ftp://ftp.ncbi.nlm.nih.gov/blast/executables>). The database can be viewed using

recent versions of common web browsers, including Internet Explorer, Safari, Firefox, Mozilla, and Opera. A detailed description of the scripts behind the database is beyond the scope of the paper. If interested in this please contact us.

Availability

The database is accessible at <http://www.fishpathogens.eu/vhsv> as a freely available tool. It is possible to browse and download data without being a registered user, but we encourage people to do so.

At the time of writing 225 VHSV isolates and 161 sequences are publicly available in the database. Since updating the database will be the responsibility of the European Community Reference Laboratory for Fish Diseases, regardless of which laboratory that conducts this function we expect that this database will be continuously updated in the future and that the number of added isolates and sequences will grow continuously. Queries and feedback on the database are welcomed (info@fishpathogens.eu).

Future development

The VHSV database is the first database available at <http://www.fishpathogens.eu>. However, in the future FishPathogens.eu is planned to be expanded to include the following pathogens beside VHSV: Infectious hematopoietic necrosis virus (IHNV), infectious salmon anaemia virus (ISAV), Koi herpes virus (KHV), and systemic RANA viruses. All viral pathogens included in lists of exotic and non-exotic fish diseases in the European Community legislation (Council Directive 2006/88/EC). Besides this the database will also be expanded with other important fish pathogens as infectious pancreatic necrosis virus (IPNV), spring viraemia of carp virus (SVCV) and related rhabdoviruses, aquatic alphaviruses (PDV and SDV), nodaviruses etc. For each pathogen a pathogen expert or a consortium of experts will be appointed to take the responsibility for proof reading and quality assurance of submitted data.

It is our aim that this database will be a valuable tool and a commonly used space for future studies on viral fish pathogens, e.g. for epidemiological studies and when designing new molecular diagnostic methods.

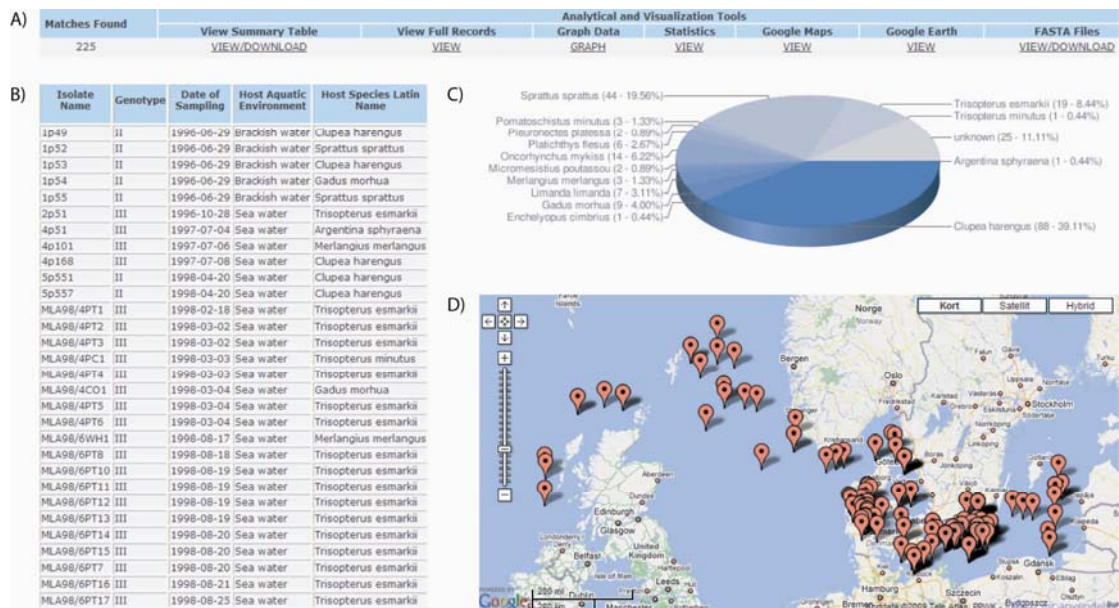
References

- Benmansour, A., Basurco, B., Monnier, A.F., Vende, P., Winton, J.R. & De Kinkelin, P. (1997) Sequence variation of the glycoprotein gene identifies three distinct lineages within field isolates of viral haemorrhagic septicaemia virus, a fish rhabdovirus. *J Gen Virol*, 78 (Pt 11), 2837-2846.
- Einer-Jensen, K., Ahrens, P., Forsberg, R. & Lorenzen, N. (2004) Evolution of the fish rhabdovirus viral haemorrhagic septicaemia virus. *J Gen Virol*, 85, 1167-1179.
- Einer-Jensen, K., Ahrens, P. & Lorenzen, N. (2005) Parallel phylogenetic analyses using the N, G or Nv gene from a fixed group of VHSV isolates reveal the same overall genetic typing. *Dis Aquat Organ*, 67, 39-45.
- Elsayed, E., Faisal, M., Thomas, M., Whelan, G., Batts, W. & Winton, J. (2006) Isolation of viral haemorrhagic septicaemia virus from muskellunge, *Esox*

- masquinongy (Mitchill), in Lake St Clair, Michigan, USA reveals a new sublineage of the North American genotype. *J Fish Dis*, 29, 611-619.
- Galperin, M.Y. (2008) The Molecular Biology Database Collection: 2008 update. *Nucleic Acids Res*, 36, D2-4.
- Nishizawa, T., Iida, H., Takano, R., Isshiki, T., Nakajima, K. & Muroga, K. (2002) Genetic relatedness among Japanese, American and European isolates of viral hemorrhagic septicemia virus (VHSV) based on partial G and P genes. *Dis Aquat Organ*, 48, 143-148.
- No_Author (2007) The database revolution. *Nature*, 445, 229-230.
- Schutze, H., Mundt, E. & Mettenleiter, T.C. (1999) Complete genomic sequence of viral hemorrhagic septicemia virus, a fish rhabdovirus. *Virus Genes*, 19, 59-65.
- Skall, H.F., Olesen, N.J. & Møllergaard, S. (2005) Viral haemorrhagic septicaemia virus in marine fish and its implications for fish farming--a review. *J Fish Dis*, 28, 509-529.
- Snow, M., Bain, N., Black, J., Taupin, V., Cunningham, C.O., King, J.A., Skall, H.F. & Raynard, R.S. (2004) Genetic population structure of marine viral haemorrhagic septicaemia virus (VHSV). *Dis Aquat Organ*, 61, 11-21.
- Snow, M., Cunningham, C.O., Melvin, W.T. & Kurath, G. (1999) Analysis of the nucleoprotein gene identifies distinct lineages of viral haemorrhagic septicaemia virus within the European marine environment. *Virus Res*, 63, 35-44.
- Stone, D.M., Way, K. & Dixon, P.F. (1997) Nucleotide sequence of the glycoprotein gene of viral haemorrhagic septicaemia (VHS) viruses from different geographical areas: a link between VHS in farmed fish species and viruses isolated from North Sea cod (*Gadus morhua* L.). *J Gen Virol*, 78 (Pt 6), 1319-1326.
- Thiery, R., De Boisseson, C., Jeffroy, J., Castric, J., De Kinkelin, P. & Benmansour, A. (2002) Phylogenetic analysis of viral haemorrhagic septicaemia virus (VHSV) isolates from France (1971-1999). *Dis Aquat Organ*, 52, 29-37.

Figure legends:

Figure 1. An example on how search results can be visualized using the FishPathogens.eu/vhsv database. A) When a search result is obtained there are several ways to display the data. Text based reports can be retrieved, several tools for visualization can be applied or FASTA files of sequences can be downloaded. B) Example on how a summary table can be displayed. The data shown in the table can be adjusted according to the users needs. The displayed data can also be downloaded in an Excel compatible format. C) Example on one of the many small visualization tools that are implemented to give a quick overview of the distribution of various parameters in the search result. Here is shown a circle diagram of the distribution of hosts within the search result. D) It is possible to visualize the geographical distribution of the isolates in the search results.



Acknowledgements

We would like to acknowledge colleagues involved in the first developmental stage of the database in the FP5 project Development of European Resource on the Origins of Pathogens of Aquaculture (EUROPA) contract no. QLRT-2001-02819. Especially K. Einar-Jensen, P. Enzmann, A. Nylund and P.Devold are acknowledged for their significant contributions to the present database. We would also like to thank all the people who have reviewed the database and helped us to improve it. This work was funded through FP6-2004-Food-3-A project EPIZONE and the European Commission financial aid for running the Community Reference Laboratory for fish Diseases.