

Manual for registering as user.

This manual aims to help you to register as user of our database. If questions arises that are not answered in this manual or you have other feed back please send it to info@fishpathogens.eu.

Advantages gained by registering

1. You get access to “my page”. From this page you can add isolate and sequence reports, browse your own reports, and update your profile.
2. When you add reports they will be registered in your own name (see [”Manual for adding an isolate report”](#) and [”Manual for adding a sequence report”](#) for further information on how to do this). This means that you are able to edit them later and access them from “my page”.
3. It is possible to add restricted data. Either the whole report can be restricted or confidential data can be added in a “private notes” field that will only be visible for you (we would encourage people to as often as possible use the last option).
4. It is possible to receive news updates via e-mail.

Manual

1. Go to the home site (<http://www.fishpathogens.eu>)
2. Click on “Create an account”.
3. Fill in your details and press “submit”
4. If things were filled in correctly you are now a registered user.
5. To log in go to the database of interest and press "log in" in upper right corner.
For VHSV (<http://www.fishpathogens.eu/vhsv>)
For IHNV (<http://www.fishpathogens.eu/ihnv>)
6. Fill in your username and password and press "Log in". This will bring you to your personal page "My page" and will let you add isolate reports and sequence reports in your own name.
If you have forgotten your log in details it is also possible to get a new password sent to your e-mail address.
7. From “My page” it is possible to make a personal configuration of how you want to view data in the database. This configuration can be set from “account profile” page under “update my account profile”.

Manual for adding an isolate report.

This manual aims to help you adding new isolates and their characteristics to our database. We are very grateful for your contribution. If you want to add a sequence of an isolate already in the database please go to "[Manual for adding a sequence report](#)". If questions arises that is not answered in this manual or you have other feed back please send it to info@fishpathogens.eu.

1. Go to the relevant home site:

If VHSV data is added follow (<http://www.fishpathogens.eu/vhsv/>)

If IHNV data is added follow (<http://www.fishpathogens.eu/ihnv/>)

2. (Only relevant for unregistered users) If you are not already a registered user we advice you to register by following instructions in "[Manual for registering as user](#)". *If you do not wish to register (and get the advantages connected with this) it is still possible to add data by clicking "add isolate report" under "Isolates" in the top menu. Start by adding a contact name and e-mail (go to step 6).*

3. If you are already a registered user press "log in" in upper right corner.

4. Fill in your username and password and press "Log in". This will bring you to your personal page "My page".

If you have forgotten your log in details it is also possible to get a new password sent to your e-mail address.

5. Under the headline "Add reports" press "Isolate".

Alternatively click "add isolate report" under "Isolates" in the top menu.

6. Before adding an isolate report please make sure that it is not already in the database, since only one report should exist pr. isolate. See "[Manual for searching reports](#)" for further information on how to do this. If an isolate has already been added that you have extra information about please contact us by info@fishpathogens.eu.

7. You will now be able to fill in data about your isolates. Fields marked with *** are required fields that need to be filled in. If you do not have the data it is often possible to answer "unknown" or similar. If you cannot find the pipe (|) on your keyboard just copy paste it in. Below are listed comments to some of the fields:

a. "Isolate name": Give new isolates a unique name that always should be used in publication etc.

b. "Synonyms": It is not preferable that an isolate is called different names, but especially old isolates sometimes have numerous names in the literature. These can be added here if they have arisen. But please do not invent new synonyms if possible

c. "Relevant References for Isolate": The reason for adding Pubmed Id's (PMID) is to be able to make a hyperlink directly to the abstract of the reference. Go to www.pubmed.com and find the reference you are looking for. The PMID can be found on the left below the abstract. If the reference is not in pubmed write relevant information about it under "general notes" in the end of the form.

d. "Geographical Location": Since several search functions in the database are map based we urge people to fill in information about "geographical Location" as precise as possible. If it is not possible to give the exact location maybe location of e.g. nearest big town can be given (If a precise location is not given please write a comment under "Comments on exact location"). Please give coordinates in both "Latitude Longitude" and "Universal Transverse Mercator" (NOTE: If you only have one of these the database can convert to the other by clicking the "convert to..." button after filling in one of them). After filling in geographical co-ordinates press "show location on world map" to make sure that the location is correctly given.

HELP FOR GETTING COORDINATES: Follow this link:

<http://www.fishpathogens.eu/vhsv/search.php>. Find the map in the bottom of the page. Find and zoom in on your location. Click on your location and the latitude/longitude coordinates appear to the left of the map.

e. "Host Aquatic Environment": Since borderlines between fresh water, brackish water, and sea water are not easily defined in real life. It will in several cases be an estimate which type of environment the host originates from.

f. "Host Species Latin Name": When this field is filled "Host NCBI Taxid" and "Host Species Common Name" will fill out automatically. If the relevant species is not on the list please contact us on info@fishpathogens.eu.

g. "General Notes": This is a free text field where relevant information that does not fit in any other category can be added.

h. "Private Notes": If the report is made public confidential information can be written here. Information in this field is only visible for the user who has added the report.

i. "Report Permissions": We urge people to make their reports publicly available and put all confidential information in the "private notes" field. However, if restricted access is needed it can be defined here by choosing "restricted" and under "User accounts with permission to view report" it can be chosen which users are able to see the report.

8. After filling in data press "Submit form" in the bottom or top of the form. If the form is not filled correctly you will get an error message telling you what to correct.

9. If everything is filled in correctly the message "the isolate report has been added" will show.

10. The report will now show in your list of added reports. To view this you click "my page" in upper right corner (only available if logged in) and under "View My Reports" click "Browse Reports".

11. After a report has been added a mail will be sent to a pathogen expert who will have to approve the isolate report before it is public available.

Manual for adding a sequence report.

This manual aims to help you adding new sequences to our database. We are very grateful for your contribution. If questions arises that is not answered in this manual or you have other feed back please send it to info@fishpathogens.eu.

1. Go to the relevant home site:

If VHSV data is added follow (<http://www.fishpathogens.eu/vhsv/>)

If IHNV data is added follow (<http://www.fishpathogens.eu/ihnv/>)

2. (Only relevant for unregistered users) If you are not already a registered user we advice you to register by following instructions in "[Manual for registering as user](#)". *If you do not wish to register (and get the advantages connected with this) it is still possible to add data by clicking "add sequence report" under "Isolates" in the top menu. Start by adding a contact name and e-mail (go to step 6).*

3. If you are already a registered user press "log in" in upper right corner.

4. Fill in your username and password and press "Log in". This will bring you to your personal page "My page".

If you have forgotten your log in details it is also possible to get a new password sent to your e-mail address.

5. Under the headline "Add reports" press "Sequence".

Alternatively click "add sequence report" under "Isolates" in the top menu.

6. Before adding a sequence report please make sure that an isolate report for the sequenced isolate is already in the database. See "[Manual for searching reports](#)" for further information on how to do this. If no isolate report has been added see "[Manual for adding an isolate report](#)" on how to do this. If you are not capable of providing this data please contact us by info@fishpathogens.eu and we will try to help you find someone who can.

7. You will now be able to fill in data about your isolates. Fields marked with *** are required fields that need to be filled in. Below are listed comments to some of the fields. When reaching the "Nucleotide Sequence" field go to point 8 (after reading point 7b below):

- a. "Isolate name": Make sure that the isolate name written is identical to the isolate name in the isolate report that you want the sequence report to link to.
- b. "Nucleotide Sequence": Add your nucleotide sequence here. Make sure that the sequence is in the correct orientation and not the reverse complement. Make sure that added sequence is of high quality. Remove all surrounding sequence and add only sequence of high quality otherwise the identification of gene regions and protein sequences will introduce mistakes.
- c. "Sequencing General Comments": This is a free text field were relevant information that does not fit in any other category can be added.
- d. "Primer Publication (Pubmed Id)": The reason for adding Pubmed Id (PMID) is to be able to make a hyperlink directly to the abstract of the reference. Go to www.pubmed.com and find the reference you are looking for. The PMID can be found on the left below the abstract. If the reference is not in

pubmed write relevant information about it under "Sequencing Method Comments".

e: "Report Permissions": We urge people to make their reports publicly available. However, if restricted access is needed it can be defined here by choosing "restricted" and under "User accounts with permission to view report" it can be chosen which users are able to see the report.

8. After filling in the sequence (See point 7b above for quality requirements) in the "Nucleotide Sequence" field press "Identify Gene Regions and Protein Sequence" just above. This will open a new window.

9. In the new window your sequence will pop-up together with the reference sequence of the database (do not alter these sequences). Press "Run Sequence Alignment" to identify your gene regions. Your sequence will now be blasted to the reference sequence and a result screen will open.

If you would like to change the BLAST parameters it is possible to do so below the sequences. However the parameters have been set to work in at least most cases.

10. The result screen show several features. Make sure that the data shown corresponds with what you expected:

Nucleotide sequence: Shows a copy of the input sequence.

Gene Regions Identified: Show information on which gene regions are identified in your sequence in the format **x;y;z;u;v** where **x** and **y** defines the start and end positions of a region in your added sequence identified as belonging to gene region **z**. The value **u** is either "no value" if a full length gene is identified or "partial" if only part of a certain gene region is identified. The value **v** is the reading frame applied to translate the gene to protein. The reading frame value can be 1, 2, or 3. If more than one gene region is identified they are listed separated by pipes (|).

Proteins Identified: Lists the protein names of the identified gene regions. No separation is made between full length and partial proteins.

Gene regions Identified with Corresponding Nucleotide and protein

Sequences: Here the gene regions identified above are shown where sequence **x** to **y** from the added sequence has been cut out. Below the sequence is a translation into protein based on the reading frame **v** assigned above.

bl2seq program output: Shows the BLAST result of the added sequence (Sbjct) to the reference sequence (Query). The database uses this BLAST output assign gene regions the following way. The best BLAST result is assumed to define the correct position of the added sequence in the VHSV genome. However since mismatches between the added sequence and the reference sequence in the ends of the added sequence might lead to an imperfect BLAST result the following rule is applied. If not all of the added sequence BLASTs to the reference sequence the BLAST result is extended to cover the whole region of the added sequence before gene regions are identified. This eliminates the problem of imperfect BLAST results but it is VERY IMPORTANT that only high quality sequences are added, since this function will otherwise introduce mistakes. Since added data is reviewed by a pathogen expert such mistakes will be in most cases be rejected during review, but we urge people to do their very best not to overload the pathogen expert with bad sequences.

11. If everything looks OK press "Copy Values to Sequence Input Form". This will close the window and automatically fill out the fields "Gene Region(s)" and "Protein Name(s)". Do not change these values.

12. Return to step 7 to fill out the rest of the form. Afterwards go to step 13.

13. After filling in data press "Submit form" in the bottom or top of the form. If the form is not filled correctly you will get an error message telling you what to correct.

14. If everything is filled in correctly the message "the sequence report has been added" will show.

15. The report will now show in your list of added reports. To view this you click "my page" in upper right corner (only available if logged in) and under "View My Reports" click "Browse Reports". To show sequence reports click "sequence".

16. After a report has been added a mail will be sent to a pathogen expert who will have to approve the isolate report before it is public available.

Appendix 1

The reference sequence used by the VHSV database is DK-Hededam (Z93412). The gene-regions are defined as followed:

Sequences spanning at least nt. 1 to nt. 10839 will be assigned as "Genome (from N-gene ORF start to L-gene ORF end)".

Sequences spanning nt. 1 to nt. 1212 will be assigned as "N-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 1314 to nt. 1979 will be assigned as "P (M1)-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 2100 to nt. 2702 will be assigned as "M (M2)-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 2794 to nt. 4314 will be assigned as "G-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 4392 to nt. 4757 will be assigned as "Nv-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 4888 to nt. 10839 will be assigned as "L-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

The reference sequence used by the VHSV database is WRAC (L40883). The gene-regions are defined as followed:

Sequences spanning at least nt. 1 to nt. 11131 will be assigned as "Complete genome".

Sequences spanning at least nt. 175 to nt. 10976 will be assigned as "Genome (from N-gene ORF start to L-gene ORF end)".

Sequences spanning nt. 175 to nt. 1350 will be assigned as "N-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 1466 to nt. 2158 will be assigned as "P (M1)-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 2255 to nt. 2842 will be assigned as "M (M2)-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 2999 to nt. 4525 will be assigned as "G-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning at least nt. 3638 to nt. 3940 will be assigned as "mid-G (303 nt)".

Sequences spanning nt. 4595 to nt. 4930 will be assigned as "Nv-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Sequences spanning nt. 5016 to nt. 10976 will be assigned as "L-gene ORF". Sequence spanning only parts of this region will be assigned as partial.

Manual for searching reports.

This manual aims to help you search for isolates and their attached sequence reports. For sequence search please also see "[Manual for searching reports using BLAST](#)" for an alternative procedure. If questions arises that is not answered in this manual or you have other feed back please send it to info@fishpathogens.eu.

Remember: If you have access to reports with restricted information it is necessary to login in order to include these reports in the search.

Primary search form.

1. Go to the relevant home site:
If VHSV data is to be searched follow (<http://www.fishpathogens.eu/vhsv/>)
If IHNV data is to be searched follow (<http://www.fishpathogens.eu/ihnv/>)
2. Choose "search reports" under "search" in the top menu or follow the link:
VHSV (<http://www.fishpathogens.eu/vhsv/search.php>).
IHNV (<http://www.fishpathogens.eu/ihnv/search.php>).
3. On this page several search mechanisms are available. They can either be used separately or in combination with each other. Below is a description of the different options available.
 - Isolate Report:** Here it is possible to make a free text search on the entire isolate report. To make search more specific it can also be chosen to search on a specific field in the isolate report (click on left drop-down menu to get the options available). Depending on whether the specific field is a free text field or a field with certain options the search field to the right will contain either a free text field or a drop-down list with available options. Click on the + to the right to add another search criterion.
 - Sequence Report:** Here it is possible to make a free text search on the entire sequence report. To make search more specific it can also be chosen to search on a specific field in the sequence report (click on left drop-down menu to get the options available). Depending on whether the specific field is a free text field or a field with certain options the search field to the right will contain either a free text field or a drop-down list with available options. Click on the + to the right to add another search criterion.
 - Gene region:** If you are interested in isolate reports with sequence reports spanning a certain gene region this can be specified here. It can also be specified whether only full length and/or partial reports should be included.
 - Sampling Date:** Here you can define a time period and get all isolates sampled in this period. Dates should be written as YYYY-MM-DD. If "From" field is left blank when searching you will get all isolates older than or equal to the date in the "To" field. If "To" field is left blank when searching you will get all isolates newer than or equal to the date in the "From" field. In stead of writing the date it is possible to click the "..." to the right of the fields. This will make a calendar pop-up and allow you to select a date.
 - Isolate Report Last Modified Date:** Here you can define a time period and get all isolates reports modified in this period. A new report added will to start with be recorded as last modified at the date it is added. Dates should be

written as YYYY-MM-DD. If “From” field is left blank when searching you will get all isolate reports modified before or equal to the date in the “To” field. If “To” field is left blank when searching you will get all isolate reports modified after or equal to the date in the “From” field. In stead of writing the date it is possible to click the “...” to the right of the fields. This will make a calendar pop-up and allow you to select a date.

Sequence Report Last Modified Date: Here you can define a time period and get all isolates reports containing sequence reports modified in this period. A new sequence report added will to start with be recorded as last modified at the date it is added. Dates should be written as YYYY-MM-DD. If “From” field is left blank when searching you will get all isolate reports containing sequence reports modified before or equal to the date in the “To” field. If “To” field is left blank when searching you will get all isolate reports containing sequence reports modified after or equal to the date in the “From” field. In stead of writing the date it is possible to click the “...” to the right of the fields. This will make a calendar pop-up and allow you to select a date.

Geographical Area: Here it is possible to retrieve all isolates located within a given radius of a certain location. On the map in the right side of the screen it is possible to scroll and zoom in on your desired location. Click on the location with left mouse button. This will make the coordinates of this point pop-up in the latitude and longitude windows to the left of the map. Choose the maximum distance from this point you would like to get isolates from.

Alternative search forms

It is possible to manually browse through all isolate reports in the database. The page can be found by choosing “Browse Isolate Reports” under “Isolates” in the top menu or by following this link:

VHSV (http://www.fishpathogens.eu/vhsv/browse_reports.php?type=1)

IHNV (http://www.fishpathogens.eu/ihnv/browse_reports.php?type=1)

It is possible to manually browse through all sequence reports in the database. The page can be found by choosing “Browse curated sequence reports” on the front page or by following this link:

VHSV (http://www.fishpathogens.eu/vhsv/browse_reports.php?type=2)

IHNV (http://www.fishpathogens.eu/ihnv/browse_reports.php?type=2)

It is possible to find isolates by searching on a world map with all isolates plotted in. The page can also be found by choosing “Geographical Distribution” under “Isolates” in the top menu or by following this link:

VHSV (http://www.fishpathogens.eu/vhsv/geo_distribution.php)

IHNV (http://www.fishpathogens.eu/ihnv/geo_distribution.php)

Manual for searching reports using BLAST.

This manual aims to help you search for sequences in the database. For search on isolates and an alternative way of searching for sequences please refer to "[Manual for searching reports](#)". If questions arise that are not answered in this manual or you have other feedback please send it to info@fishpathogens.eu.

1. Go to the relevant home site:

If VHSV data is to be searched follow (<http://www.fishpathogens.eu/vhsv/>)

If IHNV data is to be searched follow (<http://www.fishpathogens.eu/ihnv/>)

If you are a registered user and would like to include your reports with restricted access in the search log in by clicking "log in" in upper right corner and fill in your username and password.

2. Choose "search reports using BLAST" under "search" in the top menu or follow the link:

VHSV (http://www.fishpathogens.eu/vhsv/search_blast.php).

IHNV (http://www.fishpathogens.eu/ihnv/search_blast.php).

3. Fill in the search form by following point 4-9.

4. Select "Gene region". Here you choose which gene region you are interested in retrieving sequences from.

5. After a gene region has been specified the reference sequence for this region will be shown in the window "reference sequence". For more information on the reference sequence used see appendix 1 in "[Manual for adding a sequence report](#)".

6. Fill out "Position of first base pair in reference sequence" by writing the number of the first nucleotide (according to the reference sequence selected above) of the region you would like to retrieve sequences from.

7. Fill out "Position of last base pair in reference sequence" by writing the number of the last nucleotide (according to the reference sequence selected above) of the region you would like to retrieve sequences from.

8. Press "Show selected sequence". This will make your selected sequence pop-up in "Selected sequence"-window below. Check that this was the sequence you expected. *Alternatively to following point 4-8 you can copy paste your own sequence into the "selected sequences"-window. It is still recommended to fill out "gene region" if using this approach, since defining the sequence to belong to a specific gene region will lower the chance of false positives.*

9. Choose the "Percentage of match required" (a number from 0 to 100). The default value is the value that for us in most cases gives no false positives and no false negatives. Increasing the value will lower risk of false positives but increase risk of false negatives. Decreasing the value will increase the risk of false positive but lower the risk of false negatives.

10. After filling out the form press "Search the Database" to get your result.

If you would like to alter the BLAST parameters it can be done below the “Search the Database”-button.

11. You will now enter the result page where retrieved sequences can be viewed. It is possible to download these sequences in a FASTA-file by clicking the relevant button below the window showing the sequences. These files can be imported directly into most standard sequence analysis programs.

Note: Be aware that this search function only retrieves sequences spanning the complete region that you have defined. Sequences spanning only parts of the region are left out. Therefore the retrieved sequences should be of similar length unless insertions or deletions are present in some of the sequences.

Manual for looking at search results.

This manual aims to help you look at your search results. For manuals on searching data please refer to "[Manual for searching reports](#)" or "[Manual for searching reports using BLAST](#)". If questions arise that are not answered in this manual or you have other feedback please send it to info@fishpathogens.eu.

After doing a search:

On (<http://www.fishpathogens.eu/vhsv/search.php>) for VHSV

On (<http://www.fishpathogens.eu/ihnv/search.php>) for IHNV

a result page will show up offering you several ways to look at your results. Here the different options are described.

Matches Found

The number displayed here is the number of isolate reports included in your search result. The following options will be performed on this data-set.

View Summary Table

Click "view/download" to get a summary table of all isolates in the search. Clicking "select output" (or scrolling to the bottom of the page) will show a list of the different types of information that can be shown in the summary table. As default all information is shown, but if lesser information is wanted it is possible to deselect the unwanted information. After selecting what should be shown in the summary table click "Reload". This will give a new summary table where only selected information is shown.

To view a specific isolate report click its ID number in the left side of the screen in the column "view report".

It is possible to download the summary table in a format that can be imported in excel. Click "Download to excel (.csv)" and save the file on your computer. You can import by opening excel and select "Data/Import External Data/Import Data". Select the relevant .csv file. It is important to define that commas separate each field. It is also advisable to define all columns as texts to avoid excel making its own formatting in the data.

Press "Select a different search result" to get back to your search result page.

View Full Records

Click "view" to view a list showing the full isolate reports and sequence reports belonging to these of all isolates in the search. For every 20 isolates a new page is generated.

Press "Select a different search result" to get back to your search result page.

Graph Data

Click "graph" to enter a graph tool. Here it is possible to get an overview over the isolates in the search result by presenting data in a simple graph format. Choose which representation you would like to see and click "Graph" to visualize.

Press “Select a different search result” to get back to your search result page.

Statistics

Click “view” to view several statistics of the isolates in your search results. Open the type of statistic you would like to view by clicking the red headline. Close the statistics again by re-clicking the red headline.

Press “Select a different search result” to get back to your search result page.

Google maps

Click “view” to view the geographical distribution of the isolates in your search result using Google maps. Only isolates that has been given geographical coordinates will show up. It is possible to zoom and scroll using the tools in the left side of the map. In the upper right corner of the map it is possible to shift between map view and satellite view. Isolates are shown as red drops with a black dot. By clicking an isolate it is possible to get a short description of which isolate it is and by clicking on the “id” number it is possible to enter the isolate report for the specific isolate.

Press “Select a different search result” to get back to your search result page.

Google Earth

Click “view” to view the geographical distribution of the isolates in your search result using Google earth. For this function to work it is necessary to have Google earth installed on the computer. Only isolates that has been given geographical coordinates will show up. By clicking on an isolate followed by a click on its “id” number it is possible to display the isolate report for this isolate in the bottom of the screen. To learn more about the possibilities by using Google earth please see the help function in this program.

FASTA Files

Click “view/download” to open a new page where it is possible to view and/or download FASTA files of sequence reports attached to the isolate reports in the search result.

Start by choosing the gene region you are interested in. “Raw sequences” will give all sequences available in the search result in a non-trimmed form. If in stead a specific gene region is selected only sequences from this region is included in the output file. Raw sequences extending this gene region will be trimmed so that only the part of the sequence lying within the specific gene is returned. For protein coding gene regions it is possible to choose whether only sequences spanning the whole gene and/or partial sequences should be returned.

After selecting the region of interest press “Produce FASTA Files”. This will generate a FASTA file of all sequences in the search result fulfilling the criteria defined. If relevant also a FASTA file of protein sequences is generated. It is possible to download these files by clicking the relevant button below the window showing the sequences. These files can be imported directly into most standard sequence analysis programs.

Press “Select a different search result” to get back to your search result page.