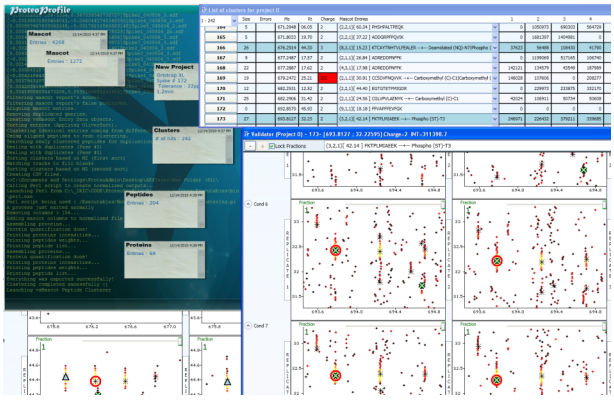
Olivier Caron-Lizotte

June 2011

Installation

What is ProteoProfile?

ProteoProfile is a software suit that assists scientists by extracting information from Mass Spectrometers data. Users can create projects, store them in a database and associate different mass spectrometer files to their different conditions / replicates / fractions. ProteoProfile can extract ion peaks from maps and cluster them together. If used with a Mascot report, ProteoProfile can extract the peptides and proteins found in your experiment and export different results to assist in quantitative proteomic analysis. To have more information on ProteoProfile, read the Technical documentation or download the binaries to try it out youself.



How to install ProteoProfile

ProteoProfile is an installation free executable written for Windows platforms. Its only dependency is .Net 4.0, which can be found in the PreRequisite folder. Installation is done in three steps:

1. If your computer doesn't have it, install **.Net 4.0** from this folder (you will need administrator access to your machine for this step).

2. Copy the ProteoProfile folder (and its content!) to **your computer**. The most up-to-date version will always be in here.

3. You can start ProteoProfile by launching « ***ProteoProfile.exe*** » from your copy of the « -= *ProteoProfile Latest* =- » folder.

When launching the application, if you see a message box telling you that your version of ProteoProfile **needs to be updated**, close the application and do step 2 of the installation process. Try not to overwrite the « Log » folder when doing an update, it could prove usefull if you find errors in the software.

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

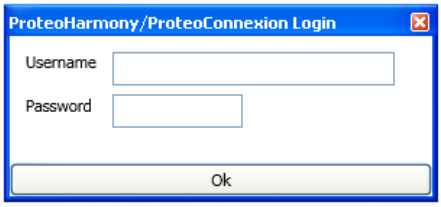
May 2012

Login

How to log to ProteoProfile

ProteoProfile uses ProteoConnection's credentials as user information. This means that in order to start a ProteoProfile session, **you need a ProteoConnection user** (username and password). Currently, creation of new users is done by Mathieu Courcelles of the Proteomic platform. You can find him in his office 2371.

When launching ProteoProfile, a small window appears asking for your username (which is your email) and password. Entering those will grant you access to the database and your list of existing projects (if any), will appear.



If you fail to log in, a message box will appear asking you if you want to try again. It is possible to cancel the login process, in which case you will not be able to retrieve your list of project, and any progress made will not be saved. Most features are still available, but keep in mind that information exported to csv files cannot be loaded back to ProteoProfile.

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

July 2010

Project list

What is a project?

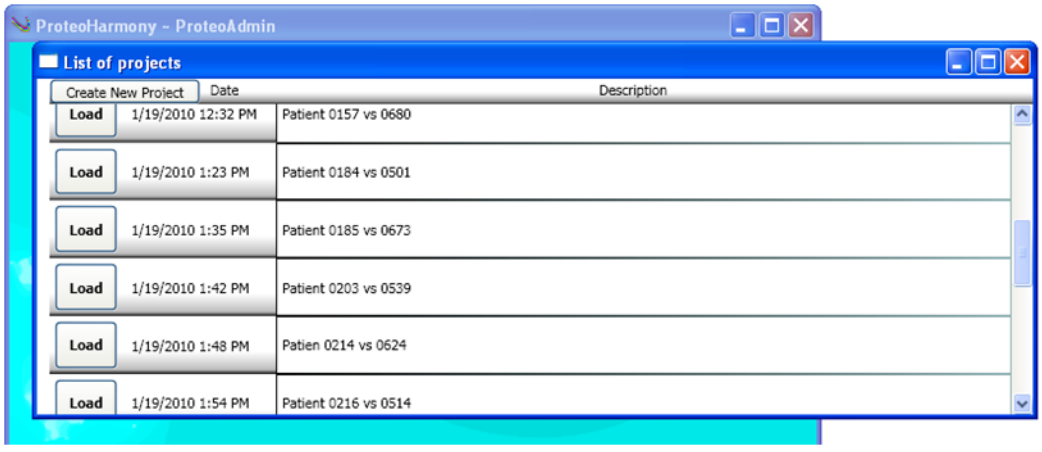
In ProteoProfile, a project regroups this information:

- A small description (or name) of the experiment;
- A list of conditions / replicates / fractions and their associated mass spectrometer file;

Currently, the user has to do a csv (comma separated value) file describing their experiment. This can be done with Notepad, Microsoft Excel or OpenOffice Calc using this file as a template.

The « List of projects » dialog

Once a user is logged in, a dialog that lists all the projects saved by the user in the database will appear. If the user is a « Power User » or an « Admin », shared projects will be listed in green. This dialog can also be opened by selecting « Show list of projects » from the contextual menu on the background of the main ProteoProfile window.



You can open previously saved projects by left clicking the corresponding project. The dialog will automatically be closed and the nodes of the project will pop one after another in the main window. The loading icon will disapear as soon as all the nodes are loaded. Also, the option to share the project among other ProteoProfile users is available by righ clicking on the corresponding project.

If you wish to start a new project, click on the « Create New Project » button and the ProteoProfile wizard will appear.

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

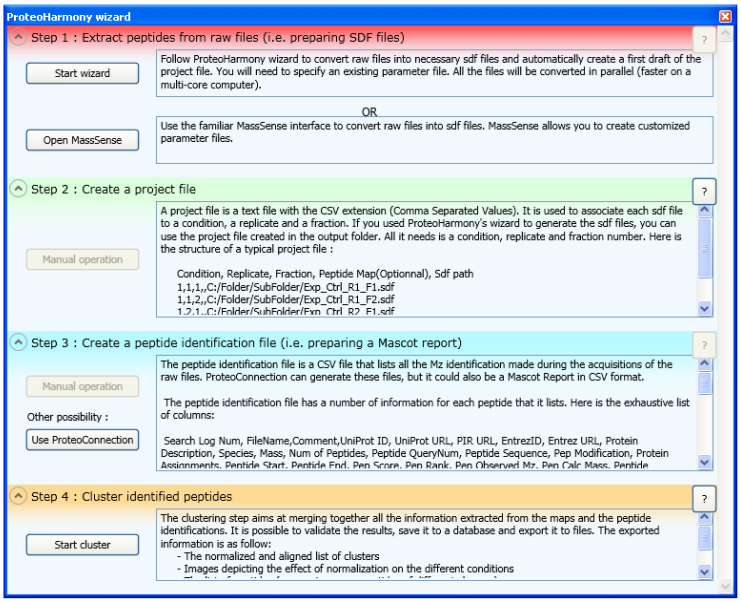
December 2010

Creating a new project

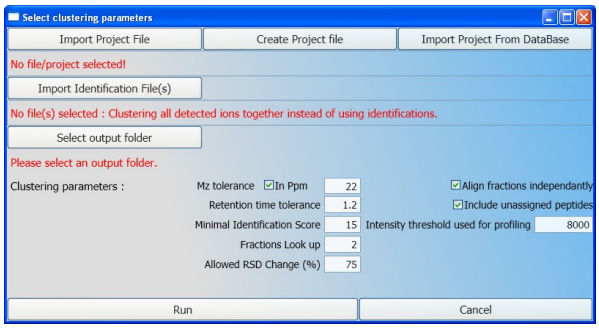
Using ProteoProfile wizard

When you first use ProteoProfile, after login in, you are greeted with a window listing your (currently empty) list of projects. To start using ProteoProfile, click on the « Create New Project » button. This action displays a new window called the ProteoProfile wizard.

The role of the wizard is to guide you through the process of clustering your experiment. Details for the 4 steps can be seen by using the « Expand » button. The 4 steps are necessary, but they do not require to be done through the wizard. Read the short texts associated with the steps to get a better idea of what to do next. You can also click on the « ? » buttons to get the corresponding page of the user manual.



When you click on the « Start cluster » button, a dialog appears where you need to enter the desired clustering options. These options include the path to the project file and the identification csv file (or Mascot report in csv format) and an output folder. Pay a particular attention to the « Parameters » zone as these parameters greatly influence the quality of the clustering results.



Click « Run » to start the clustering process. Follow the progress through the Console. Your results will eventually be written in the specified directory in csv format. You will also be asked to save your results in a database. If available, this option makes it easier to recover your results and validate them.

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

December 2010

Main window

What is a node?

In ProteoProfile, a node is a groupment of information associated to a project. Nodes are displayed on the main window as green (focused) or red (archived) rectangles. There are many types of nodes:

- Project;
- Clusters;
- Filtered clusters;
- Mascot entries;
- ...

Each nodes have there own options available throught their contextual menus (right-click). Nodes displayed in ProteoProfile's main window always represent the list of objects currently loaded in memory.

ProteoProfile's Main Window

The Main Window is where all the action takes place. Each time a node is loaded from the database, it is displayed there. Every action also spawns new objects that end up there. Nodes can be moved individually or as a whole to maintain the space a little more readable. They can also be archived, which puts them in red, and in the back. Archived node remain in memory and are still fully accessible. Focused nodes (green) serve simply as a reminder of which node a user is currently working on.

The dependency between nodes is represented by the shape of the lines connecting them. The parent node as a large line that gets thinner before it meets a child node. When exporting nodes to a database, this becomes an important fact: All the parent nodes of the node being exported will also be exported, but all its childs will be ignored.

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

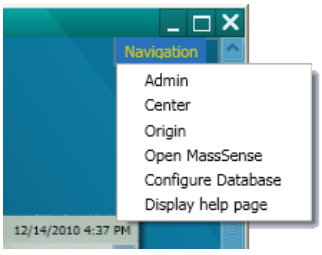
December 2010

Console

Using the main window's console

The console displays various information on what the software is actually doing. Keeping an eye on it for unusual messages is a good way of making use of the application.

Every interactions commanded by a user are displayed on the console, both bad news and good news are printed. Messages will also be written in a file, in the « *Log* » folder of the application's directory. This is of great help to speed the diagnostic process when a bug is found in the application.



Additionnally, it is possible to enter commands in the field bellow the console. Commands are directly interpreted as C# code. It is possible to use it to resolve simple equations. It can also be used to call methods that are part of the software but not accessible through the graphical user interface. Here are some examples of its use:

- Typing « 15+18*(22/3.5) » and the return key (Enter) will output the answer on the console : « 128.142857142857 ».
- Typing « CSV.FILTER » will launch an application that removes lines from a csv file based on the presence of a certain word at a specified column.

What to do if you see an error in the console

ProteoProfile is design to be robust to errors. This is good to prevent loosing data for simple, non invasive errors. But these errors still need to be reported and fixed. This is where the console, and its associated log files, are handy. When error messages of unusual warning appear on the console, report it to Mathieu. To speed the debuging process, attach to your email the latest log file and a small description of the project / file you were working on. The log file can be found in your copy of the ProteoProfile folder and is named with the date of the error :

-= ProteoProfile Latest =-/Logs/YourDomain/ YourUser/ Day_Month_Year.txt

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

February 2010

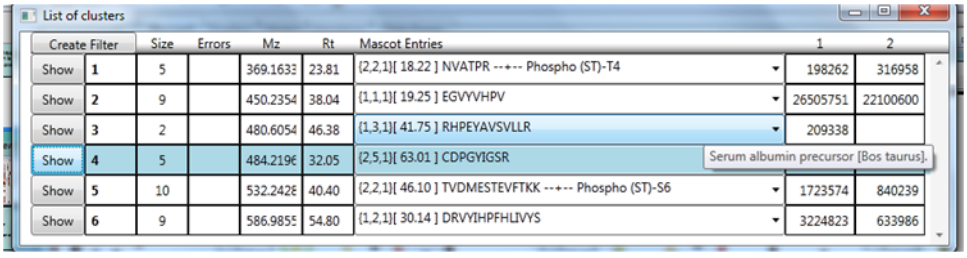
Validation

Importance of cluster validation

Clustering is a non-trivial operation that consists of grouping together elements between objects. In this case, it regroups common peptide ions found on different maps. Since it clusters different conditions together, peptides that are very similar are sometimes put in a cluster they do not belong to. So far, nothing beats the human eye to spot these errors. That is why it is very wise to look at important clusters (peptides) that show unpredicted behavior. It is also necessary to validate clusters that have duplicates between them to correct the clustering error that occured. To learn more about clustering, read ProteoProfile's technical documentation.

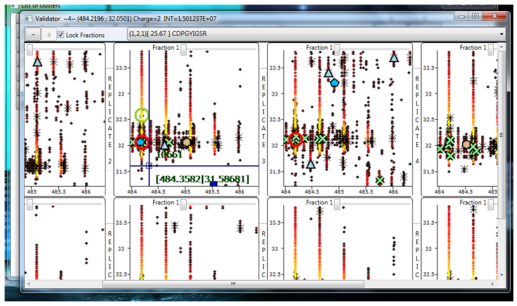
Using the Validator

To launch the validation tool, use the contextual menu on a cluster or filter node. This tool consists of multiple windows that can be reshaped and moved to fit the user's need. There are two types of windows.



The list of clusters displays information on all the clusters belonging to the node. Each line can be left click (to simply change its color) or right click to access the list of operations available on the cluster. Current operation list include a deletion option for the selected cluster. Visualisation of the actual content of the cluster is accessed throught the button at the left side of each cluster. It is also possible to open an arbitrary zone (mz/rt) by right clicking on the header of the list of clusters.

The cluster window displays the heat maps for all the conditions in the project. Any number of windows can be opened simultaneously. As long as the window is opened, the cluster line from the corresponding cluster list window will remain dark blue. When cycling between windows, this helps identify faster the cluster that are currently opened.



Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

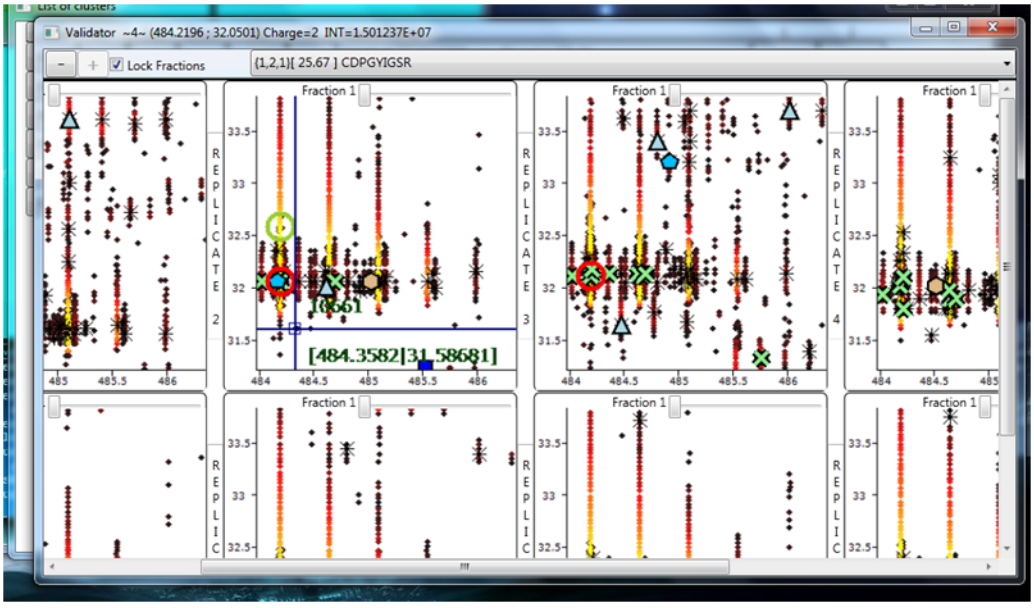
Olivier Caron-Lizotte

February 2010

Validator window

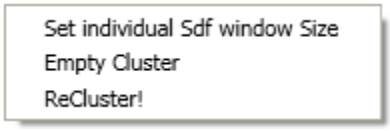
Using the validator window

The main purpose of the validator window is to allow the manual modification of clusters. In this regard, it displays as much information as possible to help users in their decision process. The title



of each window shows information about the current values of the cluster. Below the title, the zoom back and zoom forward buttons becomes available as soon as the user as zoomed in on one of the maps. If selected, the « Lock fractions » option lets the user control all the sliders (read below) at the same time. Furthermore, a drop-down list shows all the Mascot entries associated to this cluster.

A contextual menu is available by right-clicking in the background of the validator window. It offers three options to (respectively) : set the desired size of the heat maps, in pixel; remove all the peptides from this cluster; start a new clustering for these mascot entries, using the current mz/rt values of the cluster as guide.



The heat maps are disposed so that each line corresponds to a condition, and each column a replicate. Each fractions are then displayed on top of one another. To cycle through fraction, use the slider located at each right corner of heat maps. Two heat maps are displayed at the same time, with varying transparency depending on the slider's position. When interacting with the heats maps, pay attention to the fraction you are currently working on as it is not always trivial based only on the position of the slider.

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

February 2010

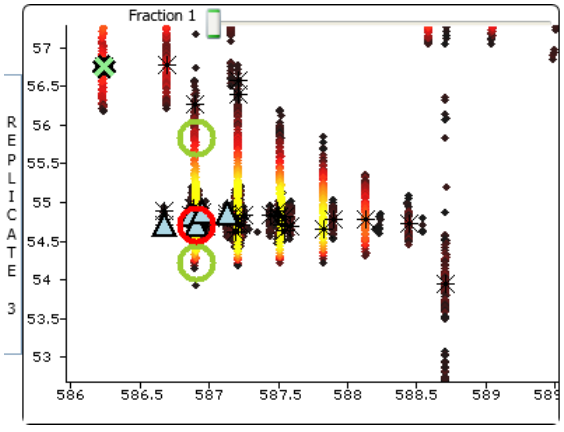
Heat maps

Interacting with the heat maps

Heat maps are mass spectrometer files rendered by dots of color : yellow for high intensity values, and red for medium values and black for low values. The maps are displayed using mass to charge ratio as X axis and retention time as Y axis. Detected peptides are drawn with different shapes:

- Star : Unknown charge
- Plain purple circle: 1
- Green cross : 2
- Triangle : 3

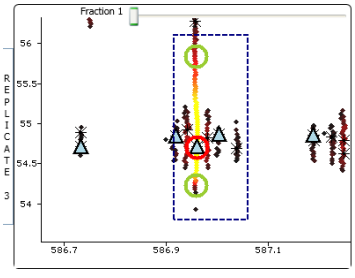
For every other shapes, the number of sides corresponds to the charge state.

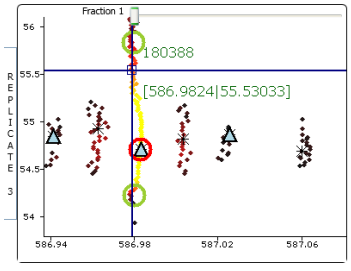


Peptides included in the cluster and mascot entries associated with this map are represented respectively by red empty circles and green empty circles. The highest scored mascot entry is displayed as a green diamond.

To add a peptide to a cluster, double click where you want it to be. If another peptide was already selected for that map, it will be replaced. If you wish to remove a peptide from one map, simply double click on the map, in a zone without any value (no colored dots).

To zoom in on a zone of the heat maps, drag the right mouse button on the map. This will draw a blue rectangle. Releasing the right mouse button will then display the content of the drawn rectangle, stretched to fill the space between the X and Y axis.





When dragging the left mouse button, a blue cross follows the cursor. This cross is accompanied with the values of the pointed artefacts. The cross is useful to see what is the currently edited fraction.

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

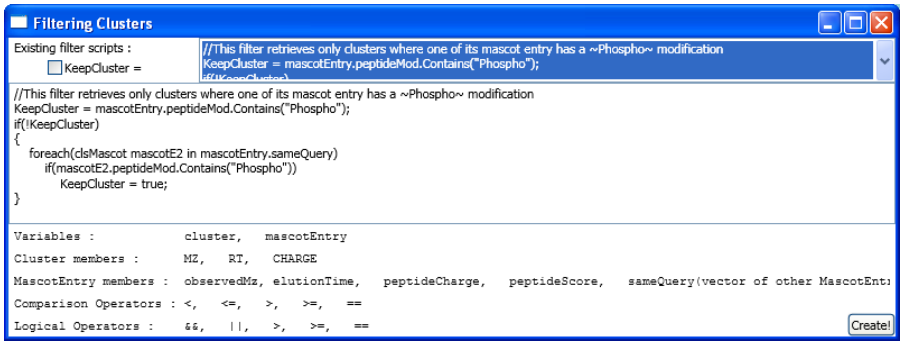
December 2010

Filters

Using filters

By using the contextual menu of a cluster node, you can access the « Create filter » option. This option is also available from a button on the list of cluster window (top left corner). The « Filtering clusters » window has a list box in the top right corner that shows the already written filters. These can be chosen and modified, at will. Each time a filter node is saved to the Database, it is added to this list.

New filters can be written in C#. The filter is run for each cluster that are kept only if the KeepCluster variable is set to « true ». At the bottom of the dialog, there is a small legend describing some of the available variables. This is not an exhaustive list, and you might have to refer to the code of ProteoProfile to write complex expressions. The code is available through the subversion server.



Examples of filters

The following filter retrieves only clusters where one of its mascot entry has a ~Phospho~ modification:

```
KeepCluster = mascotEntry.peptideMod.Contains("Phospho");
if(!KeepCluster)
{
   foreach(clsMascot mascotE2 in mascotEntry.sameQuery)
      if(mascotE2.peptideMod.Contains("Phospho"))
         KeepCluster = true;
}
```

The following filter retrieves only clusters where the highest scored peptide has the sequence « ETQGG » :

```
KeepCluster = mascotEntry.peptideSeq.Contains("EQTGG");
```

The following filter keeps clusters with MZ values between 600 and 800:

```
KeepCluster = cluster.MZ > 600 && cluster.MZ < 800
```

Useful links

- **User guide**
- **Technical documentation**
- **Subversion repository**
- **Software support**

Olivier Caron-Lizotte

February 2010

Something is missing?

Tell me about what is missing in here!

Useful links

- User guide
- Technical documentation
- Subversion repository
- Software support