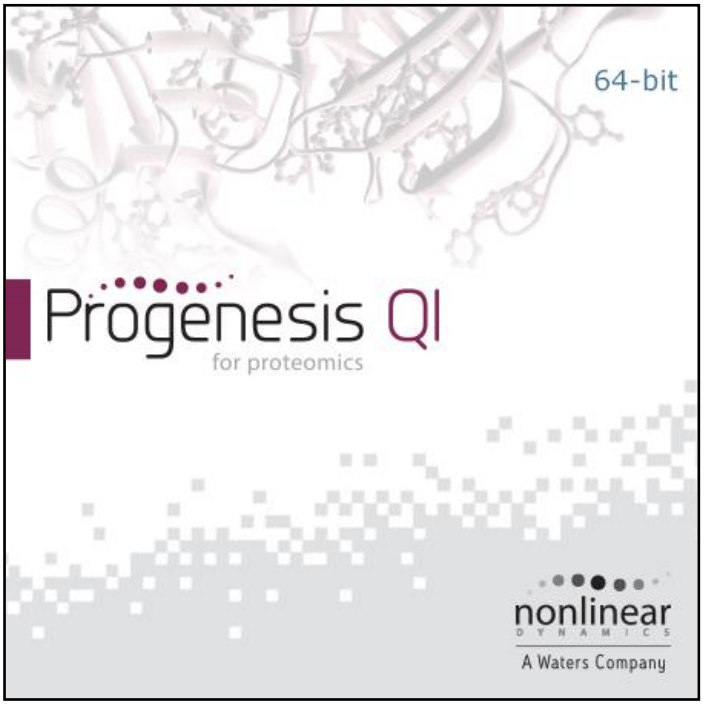


Progenesis QI for proteomics User Guide

Analysis workflow guidelines for HDMse and MSe data

Waters

THE SCIENCE OF WHAT'S POSSIBLE.™

Contents

Introduction

This user guide takes you through a complete analysis of 9 LC-MS runs with 3 groups (3 replicate runs per group) using the unique Progenesis QI for Proteomics workflow. It starts with LC-MS data loading then Alignment, followed by Peak Detection that creates a list of interesting peptide ions (peptides) which are explored within Peptide Ion Stats using multivariate statistical methods then onto Protein identity and Protein Stats.

To allow ease of use the tutorial is designed to start with the restoration of an Archived experiment where the data files have already been loaded. The document covers all the stages in the workflow, initially focusing on the Automatic Processing of the data then the use of the individual analysis stages.

If you are using your own data files please refer to Appendix 1 (page 76) then start at page 6.

How to use this document

You can print this user guide to help you work hands-on with the software. The complete user guide takes about 60 to 90 minutes. This means you can perform the first half focused on LC-MS run alignment and analysis then complete the second half of analysis exploring comparative differences and Protein identity at a convenient time.

If you experience any problems or require assistance, please contact us at support@nonlinear.com

How can I analyse my own runs using Progenesis QI for proteomics?

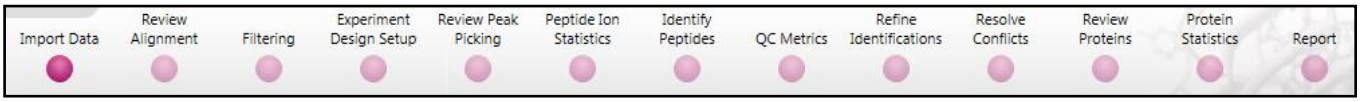
You can freely explore the quality of your LC-MS data using Data Import and then licence your own LC-MS runs using this evaluation copy of Progenesis QI for proteomics. Instructions on how to do this are included in a section at the end of the user guide document.

LC-MS Data used in this user guide

For the purposes of this data set the MS^E parameters were set to 250:125:1000 instead of the default settings as defined in Appendix 1 (page 76). This was to done to reduce the time taken to demo the data analysis.

Workflow approach to LC-MS run analysis

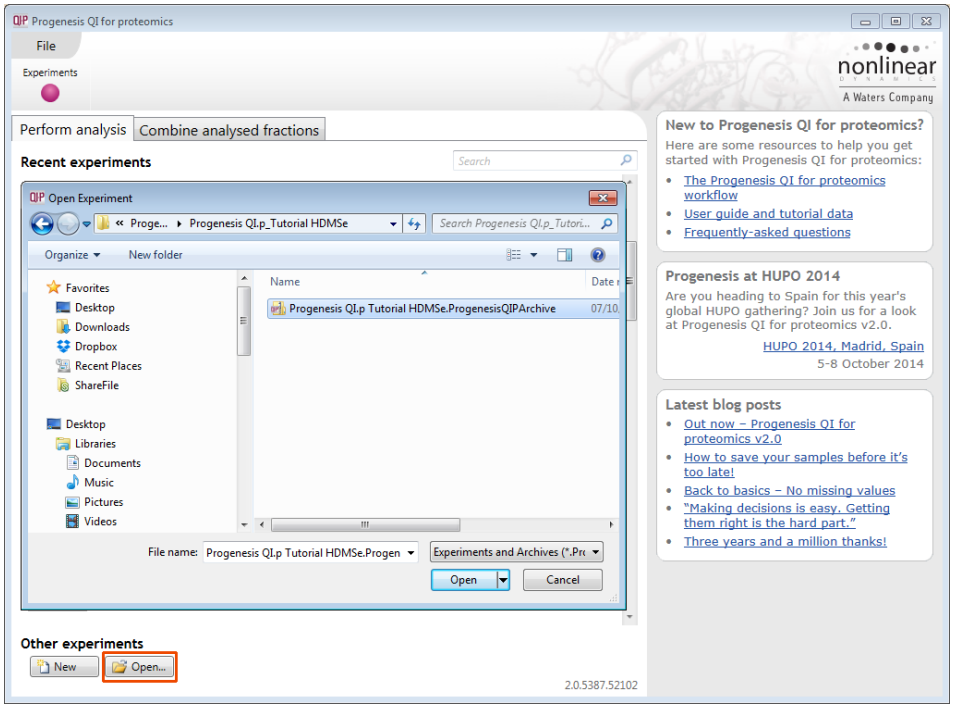
Progenesis QI for proteomics adopts an intuitive **Workflow** approach to performing comparative LC-MS data analysis. The following user guide describes the various stages of this workflow (see below) focusing mainly on the stages from Alignment to Report.



Restoring the Tutorial

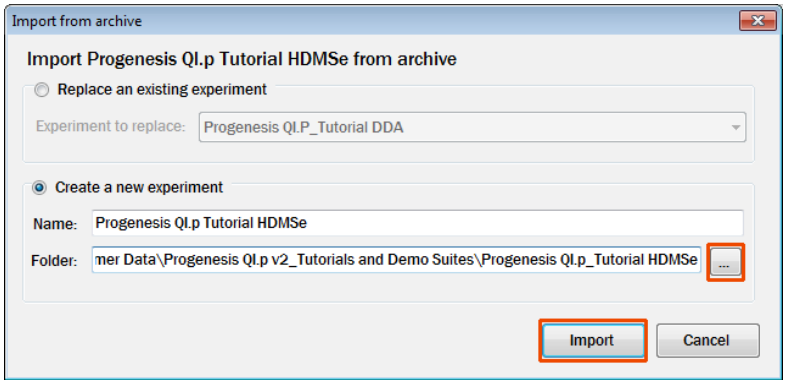
Open Progenesis QI for proteomics and download the Compressed (.zip) Tutorial Archive file from the 'User guide and tutorial' link shown below, placing it in a **new folder** on your desktop. Before restoring the tutorial in the software **you must** first right click on the (.zip) file and extract it to the same folder.

Now restore the uncompressed tutorial archive file. To do this, first locate the **Progenesis QI.P Tutorial HDMSe.Progenesis QIP Archive** file using the **Open** button and press Open.

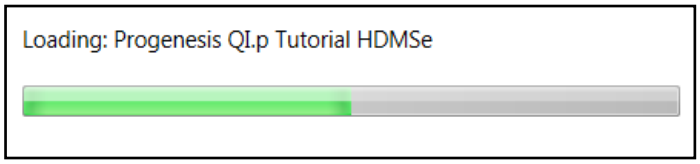


This opens the 'Import from archive' dialog.

Select the **Create a new experiment** option and select the folder in which you placed the archive, using the icon (to the right).



Then press **Import.**

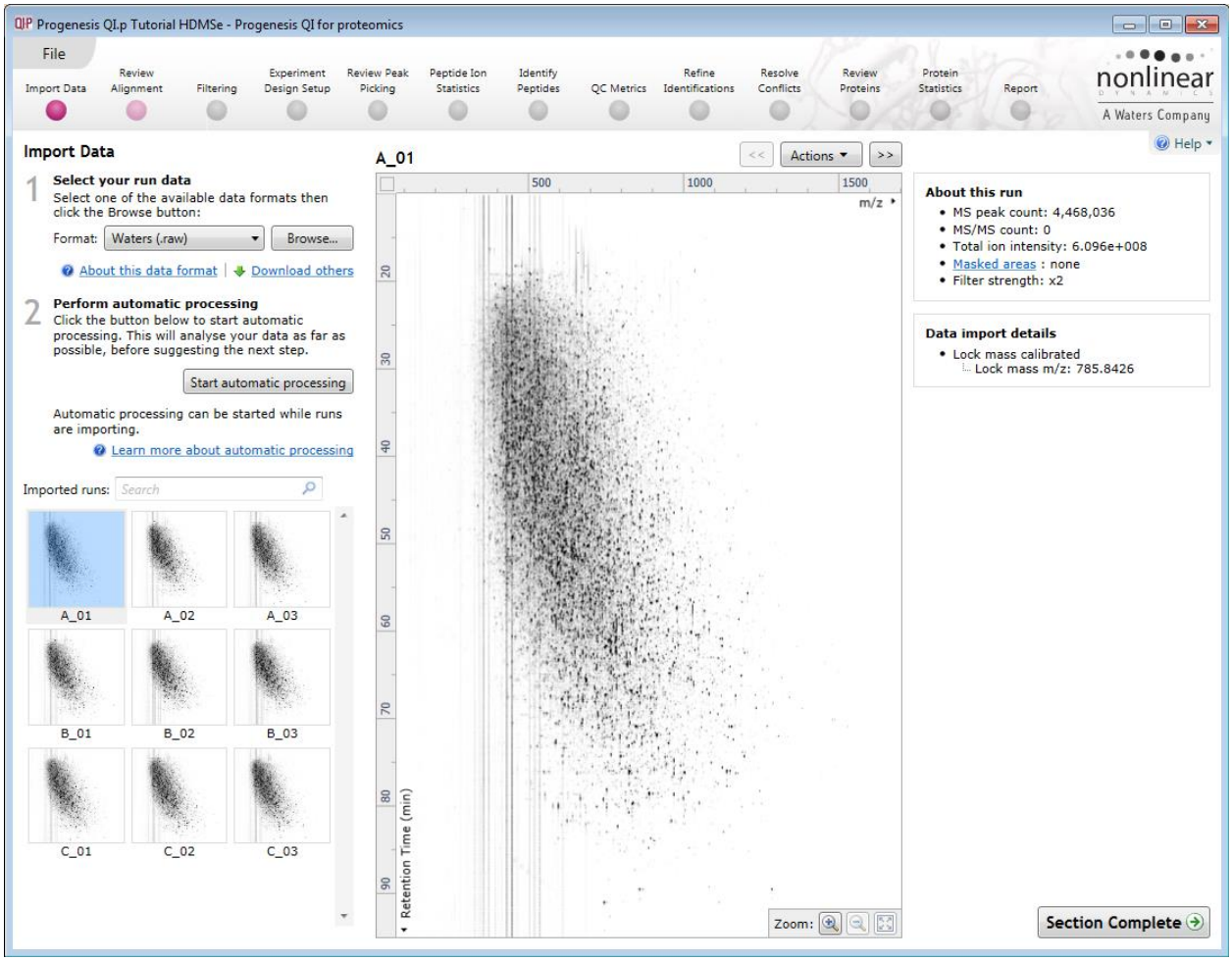


Note: use the **Replace an existing experiment** option if you want to over-write an existing version of the tutorial.

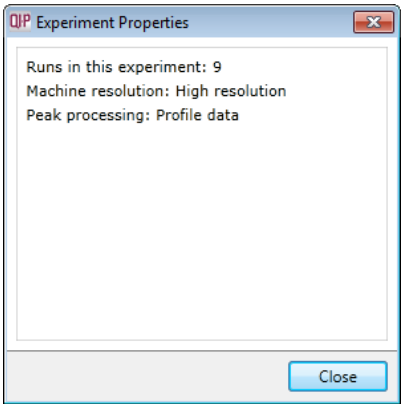
Tip: at each stage in the software there are links to more information and help on the website.

Stage 1: Import Data and QC review of LC-MS data set

The tutorial will now open at the **Import Data** stage (see below).

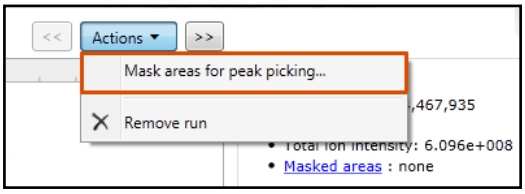


Each data file appears as a 2D representation of the run. At this stage you will be warned if any of the data files have been 'centroided' during the data acquisition and conversion process.



Note: the **Experiment Properties** are available from the File menu**.** These were selected when the experiment was created (see Appendix 1, page 76).

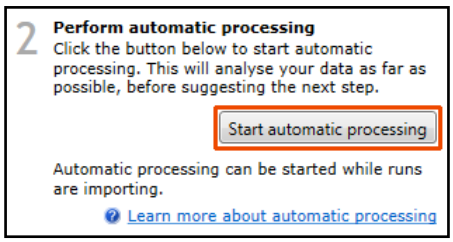
Tip: the **'Exclude areas from selected run'** facility allows you to examine and exclude areas (usually early and/or late in the LC dimension (Retention Time)) that appear excessively noisy due to capture of data during column regeneration. This is not required for this data set.



Note: use the **Remove Run** to remove run(s) from the current experiment.

Now start the Automatic processing.

Stage 2A: Automatic Processing of your data



The Automatic Processing of your data can be set up and started before the import of your data has been completed by clicking on **Start automatic processing.**

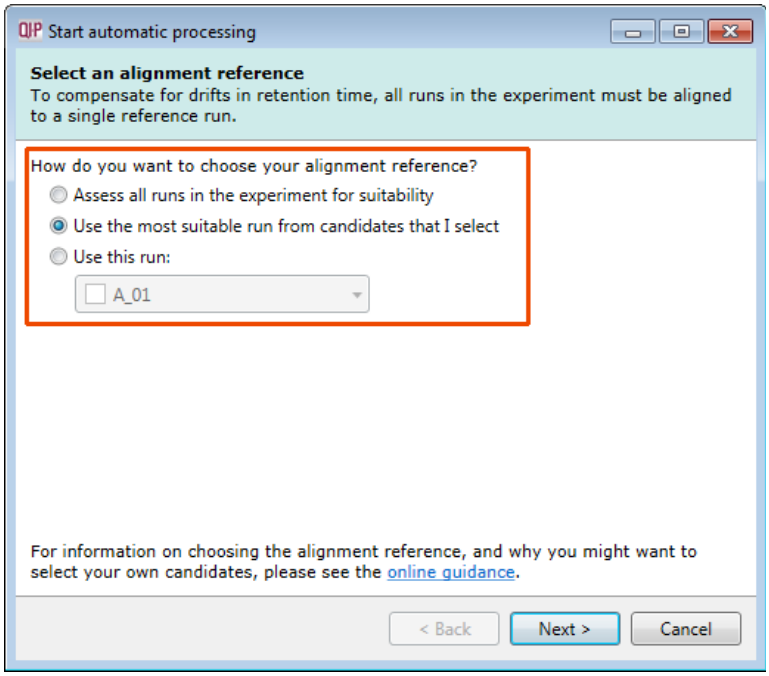
Note: for this tutorial the data has been imported already.

Setup of processing steps in the Analysis Workflow, up to and including Identify Peptides, can be performed in the Automatic Processing wizard, these include:

- Select an alignment reference
- Automatic alignment of all runs to a reference run
- Automatic peak picking for peptide ion detection
- Define an Experiment design
- Perform peptide and protein identification (for MS^E data only, as DDA data will require external database searching)
- Choosing a Quantitation method to be applied for abundance analysis

In this tutorial example you have 9 HDMSe LC-MS runs, so the automatic identification of peptides is available as part of the automatic processing.

As the runs have already been imported, click **Start automatic processing** to setup the Automatic processing wizard.



Progenesis QI for proteomics provides three methods for choosing the alignment reference run:

1. **Assess all runs in the experiment for suitability**
   This method compares every run in your experiment to every other run for similarity, then select the run with the greatest similarity to all other runs as the alignment reference.

If you have no prior knowledge about which of your runs would make a good reference, then this choice will normally produce a good alignment reference for you. This method, however, can take a long time for a large number of runs.

2. **Use the most suitable run from candidates that I select**
   This method asks you to make a selection of reference candidates; the automatic algorithm then chooses the best reference from this subset of runs.

This method is appropriate when you have some prior knowledge of your runs suitability as references:
   i.e when all the candidate runs are pooled samples or
   i.e when all the candidates are from a condition that displays the largest set of common peptide ions.

3. **Use this run**

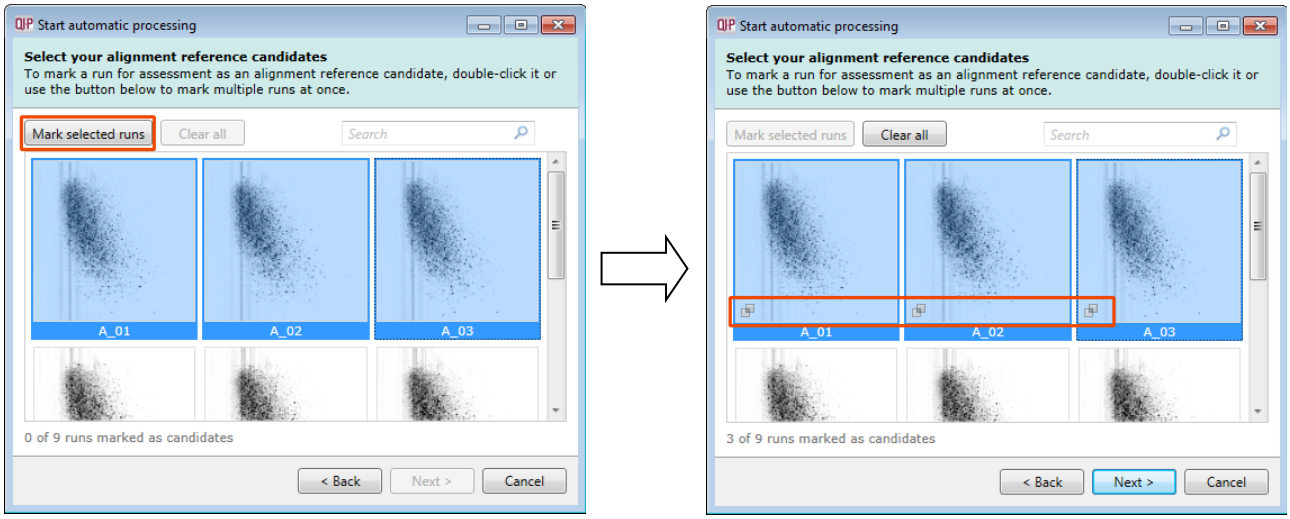
This method allows you to manually choose the reference run.

Manual selection gives you full control, but there are a couple of risks to note:

- If you choose a pending run which subsequently fails to load, alignment will not be performed.
- If you choose a run before it fully loads, and it turns out to have chromatography issues, alignment will be negatively affected (for this reason we recommend that you let your reference run fully load and assess its chromatography before loading further runs).

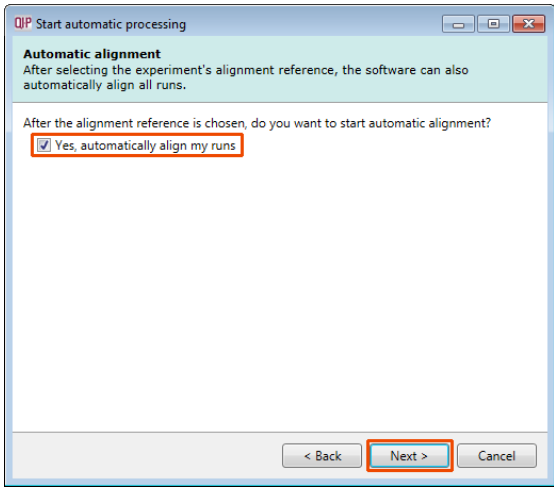
For this tutorial we will select the second option.

Select a subset of your runs as 'alignment reference candidates' and click **Mark selected runs.** An icon appears on the bottom left of each run to indicate that it is selected as a candidate.

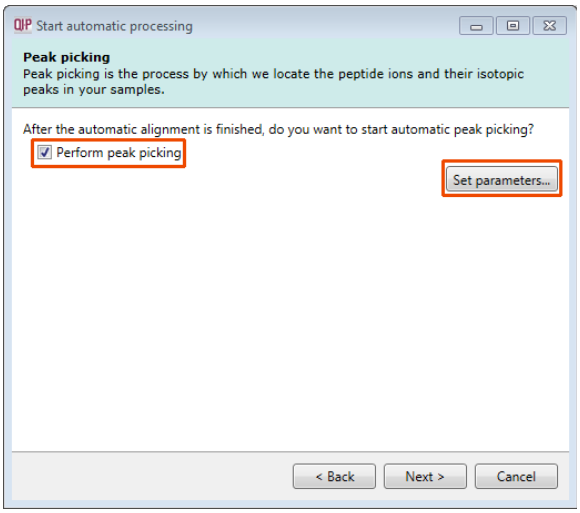


On clicking **Next** you will be asked if you want to align your runs automatically.

The default is for automatic alignment, click **Next**.



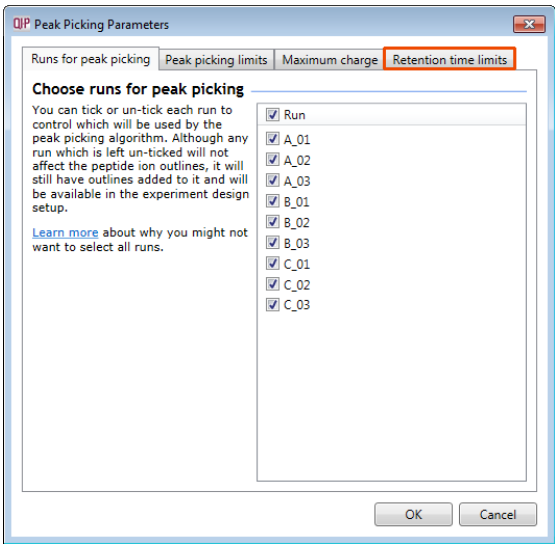
The next page of the processing wizard will ask you if you want to **Perform peak picking** and allow you to set appropriate parameters.

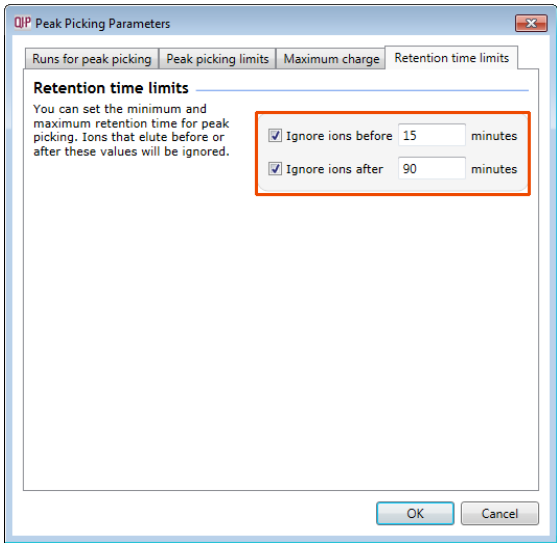


For the purposes of this User guide we will use the default settings for peak picking except in the case of defining Retention Time limits where we will limit the peak picking to between 15 and 90 min.

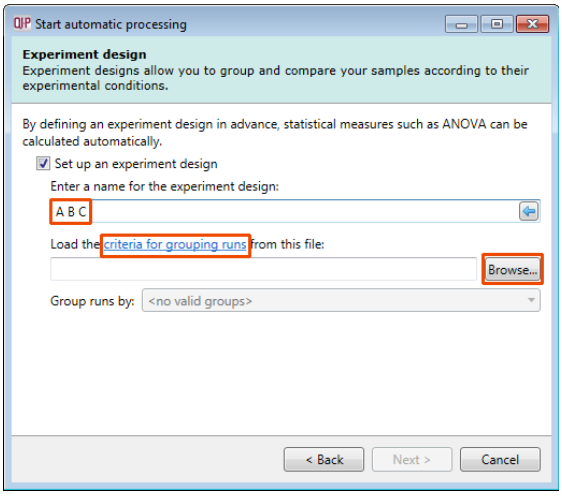
Note: for more details on setting Peak Picking parameters refer to the section on Filtering (page 22)

Click on **Set parameters** and select the fourth tab to set **Retention time limits** for the detection. The default limits are displayed in the unticked boxes and detection will use these values whenever the box(s) are unticked.





Enter values of 15 and 90 min and tick the boxes as shown above.



Click **OK** to return to the Automatic Processing Wizard and click **Next** to Define an Experiment design.

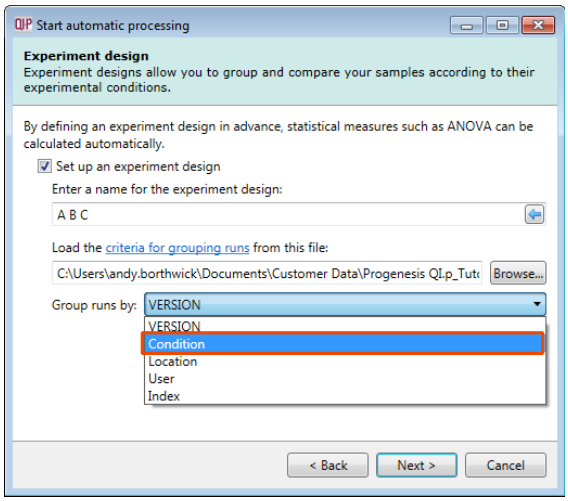
To handle the grouping of your run data you can make use of **sample tracking information** that has been stored in a spread sheet at the time of sample collection and/or preparation.

Note: if a spreadsheet file of your data is not available you can create your experiment designs after the automatic analysis of the runs

For this example there is a **QIP_Conditions.spl** file available with the Experiment Archive you restored at the beginning of this tutorial exercise.

Give the experiment design a name (i.e A B C) and then use the **Browse** function to locate the **QIP_Conditions.spl** file.

To use this approach select the **Import design from file** option from the New Experiment Design dialog. Then locate the QIP_Conditions file and select what to **Group runs by**, for example: **Condition**.

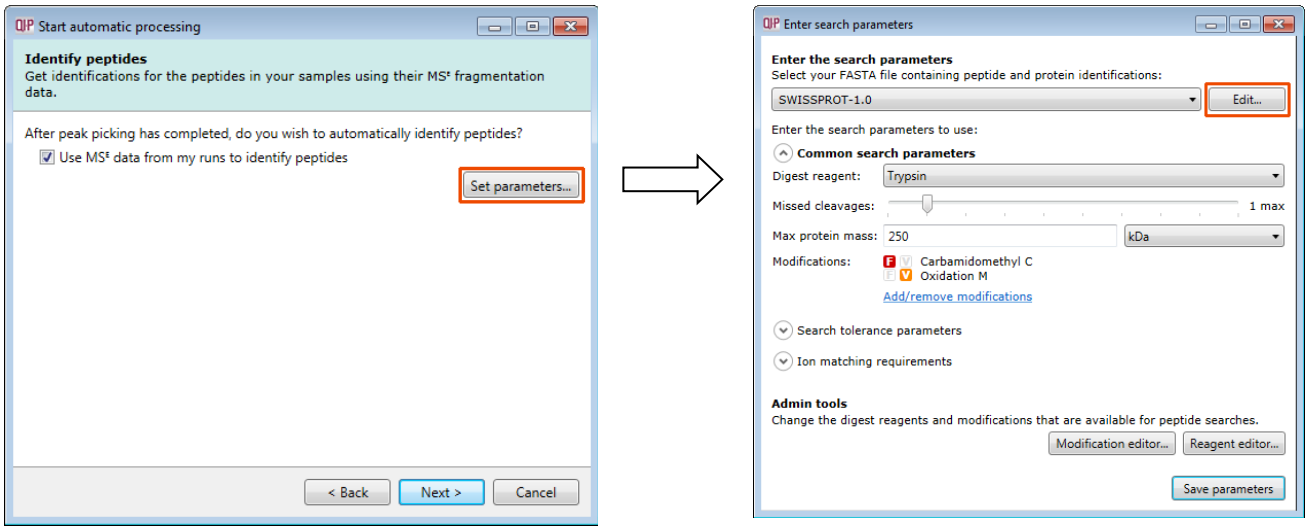


Note: you can create additional experiment designs following the completion of automatic processing.

Select Conditions and then click **Next**.

For MS^E fragmentation data you can set up the peptide identification to be performed automatically.

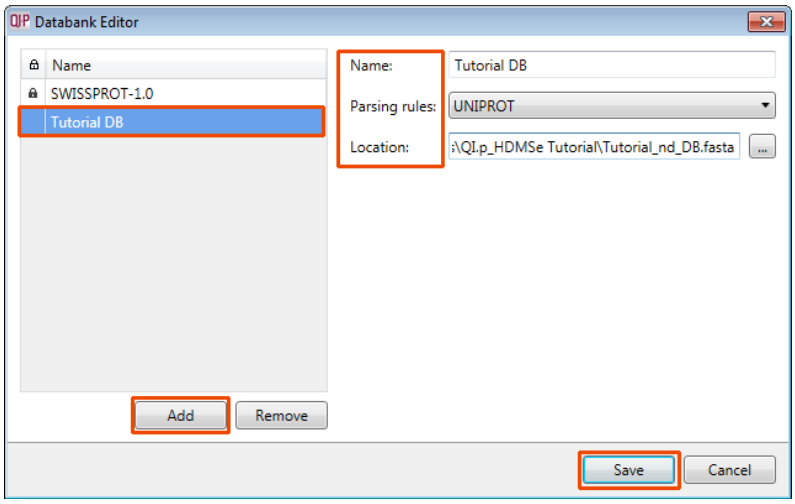
Note: if the software has detected MS^E data then this option will be ticked by default.



Click on **Set Parameters**. The default Databank is for Swissprot-1.0 (which is a locked example).

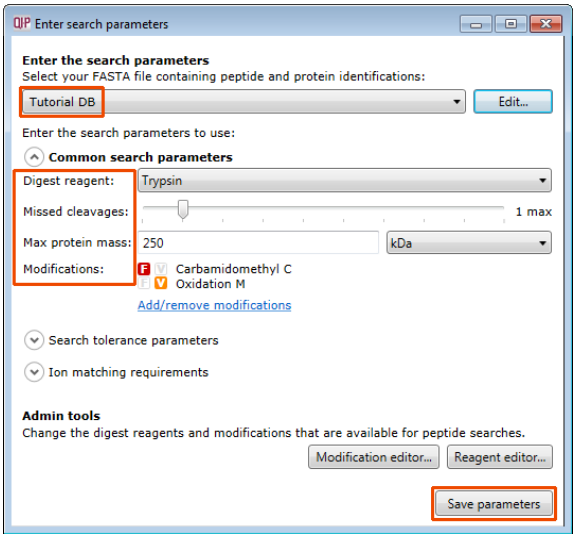
To create a new Databank from a Fasta file click on **Edit** and then create a new one using the example FASTA file that is available with the Experiment Archive you restored at the beginning of this tutorial exercise (**Tutorial_nd_DB.fasta**).

For a new Databank click **Add.** Then give it name (i.e. Tutorial DB), select the parsing rules (UNIPROT) and specify the location of the FASTA file, see the example below.



The new Data bank will appear in the left panel now click **Save** to return to the Search parameters.

If your databank is not already displayed then select it from the drop down list.



Check the **Common search parameters**

The default settings are displayed:

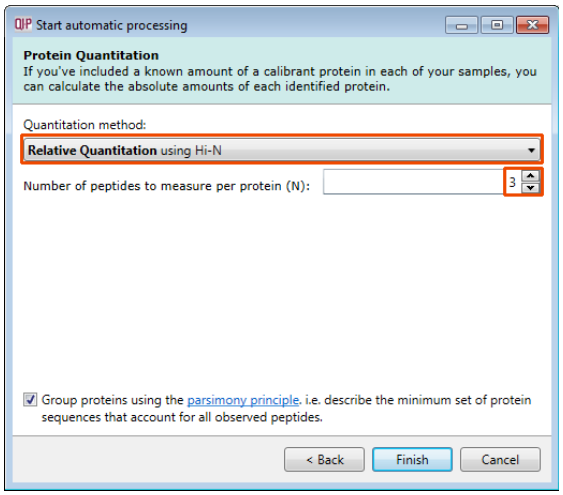
Digest reagent: is set as Trypsin**.** Alternative Digest reagents are available from the list and additional ones can be added to the list using the **Reagent editor...**

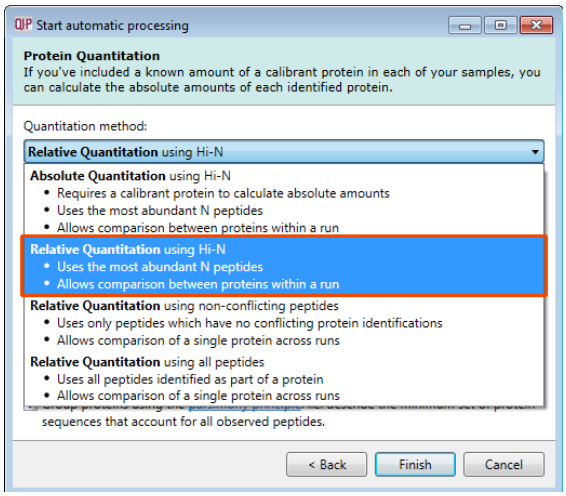
Missed cleavages: is set as 1.

Maximum protein mass: is set at 250kDa

Modifications: are set Carbamidomethyl C (Fixed) and Oxidation M (Variable). More modifications are available from the list and additional ones can be added to the list using the **Modification editor…**

Click **Save parameters** and then **Next.**



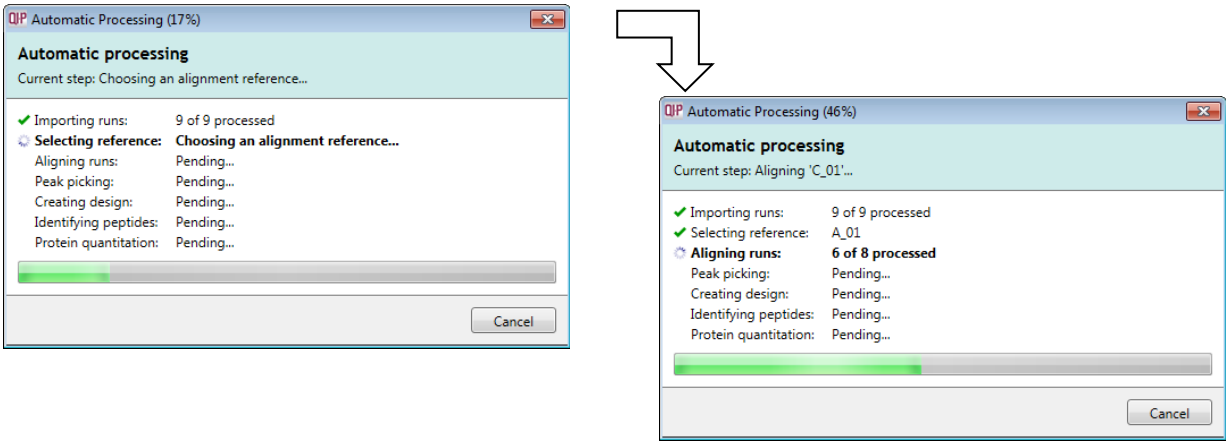


The **Protein Quantitation** dialog opens displaying the default method, **Relative Quantitation using Hi-N** which uses up to 3 peptides per protein to compute the relative amount of each 'identified' protein.

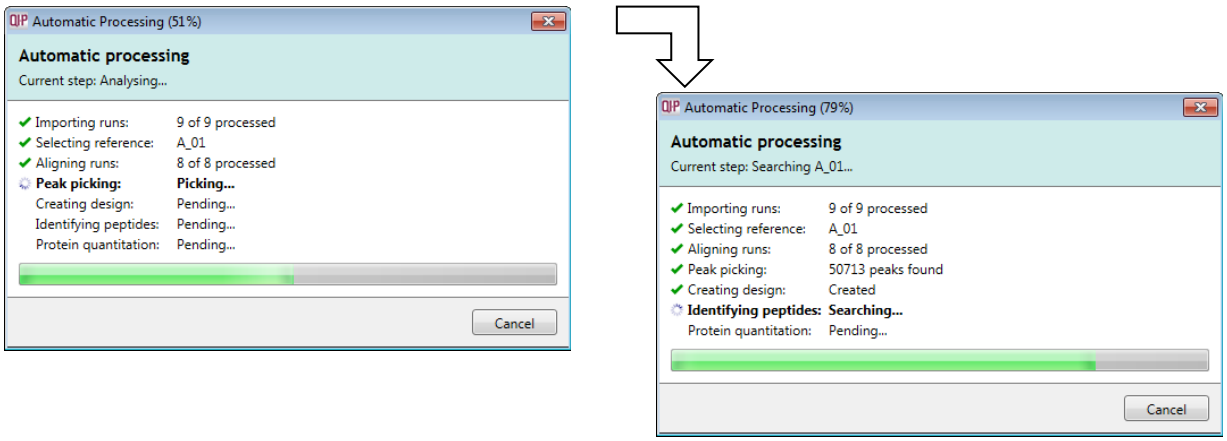
Use the drop down to reveal the alternative methods for protein quantitation

Select the Default option and click **Finish.**

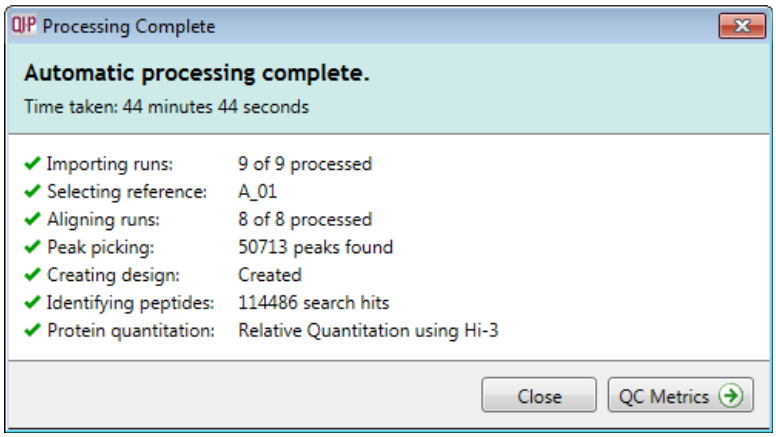
The Alignment process starts with the automatic selection of A_01 as the alignment reference



Once Alignment completes Peak Picking commences followed by Design and Identification.



Finally the Automatic Processing completes with the Quantitation being performed

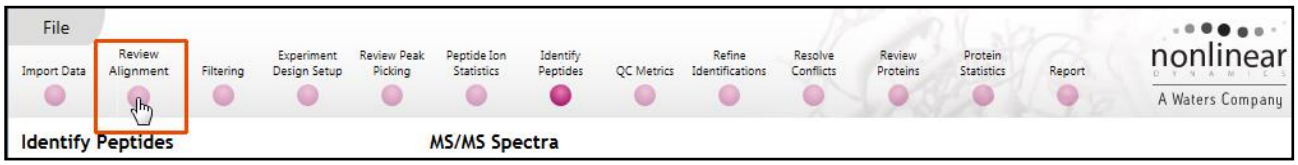


As the whole process proceeds you get information on what stage has been performed and also the % of the process that has been completed.

When Processing completes, depending on what stages you selected to perform, the Wizard displays what stage the workflow will open at. In this example it will open at QC Metrics.

You can either:

- Continue with the analysis, as the Processing dialog is not displaying any warnings, and review QC Metrics. In which case you can go to page 52
- Open the analysis at QC Metrics and immediately return to the Review Alignment stage by clicking on it in the Workflow to review alignment quality (page 16)
- Alternatively you can **Close** the dialog. This will not move you to a later stage in the workflow but instead allow you to navigate through all the stages yourself or jump from stage to stage.



Note: if processing fails to complete successfully there are a number of suggested strategies you can use to proceed with your analysis. These are out lined in Appendix 2 (page 80)

Please Note: the time taken to automatically process this data through all the steps including Peptide Identification takes approximately 45 minutes (as shown on the previous page). This is dependent on the specification of your PC.

Note: this does not include the time to load the data

You can explore and re-perform the steps, sequentially and/or as part of the automatic processing as described in this guide.

For this data set, the quality of the alignment and detection following automatic processing does not require to be re-performed. However, in the course of exploring the data you may choose to edit and re-perform the alignment and or the peak detection as described in this document.

You will be warned that you are about to lose the analysis performed after the step that you are about to 'unlock' to re-perform.

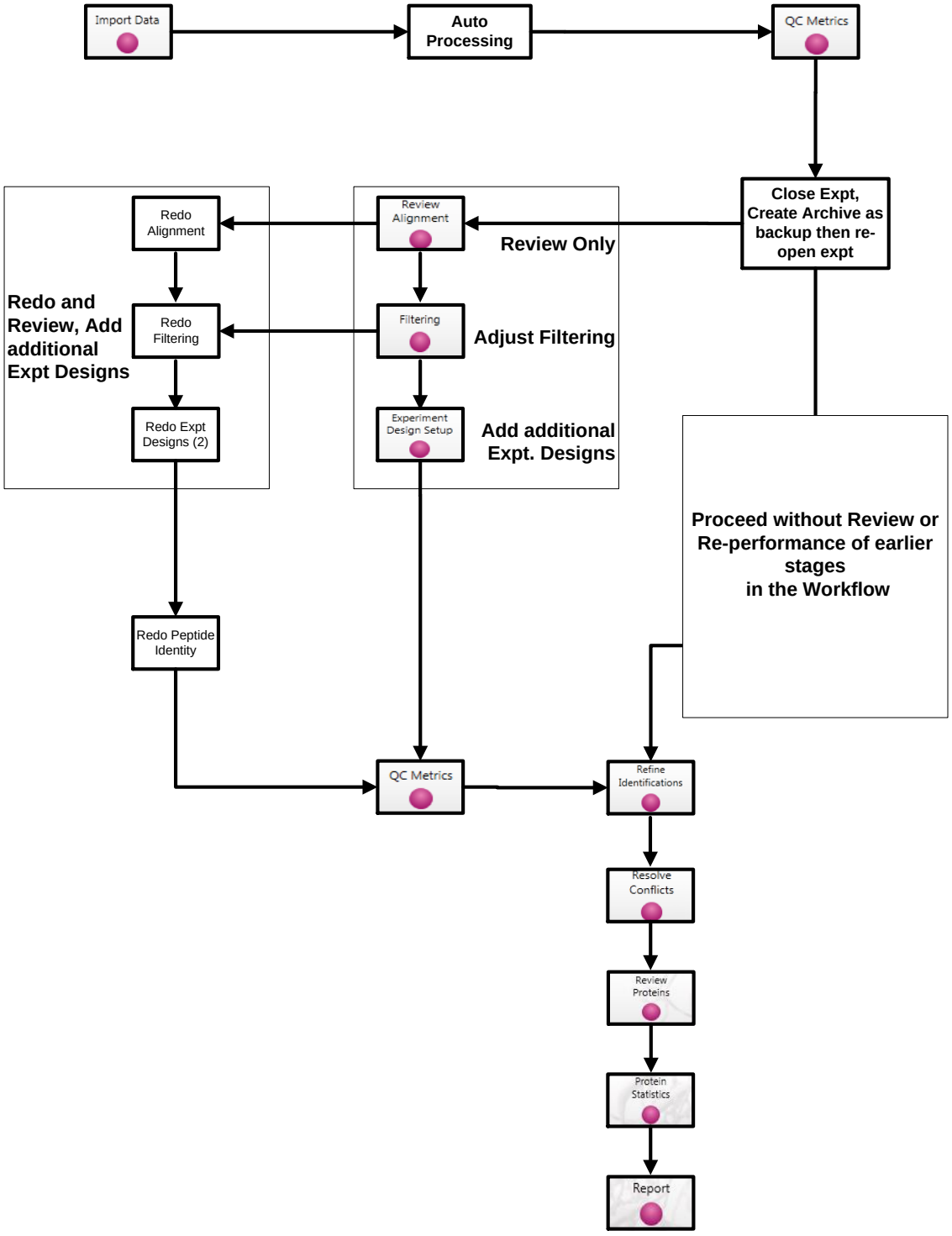
An example of the time it takes to perform each step is outlined in Appendix 8 (page 93).

The next stage in this document gives you a graphical view of how to proceed with your analysis following automatic processing.

Stage 2B: After Automatic Processing

When Processing completes, depending on what stages you selected to perform, the Automatic Processing Dialog displays (on the right) what stage the workflow will open at.

The flow chart gives you an overview of the various steps you can take to proceed with your automatically processed data.



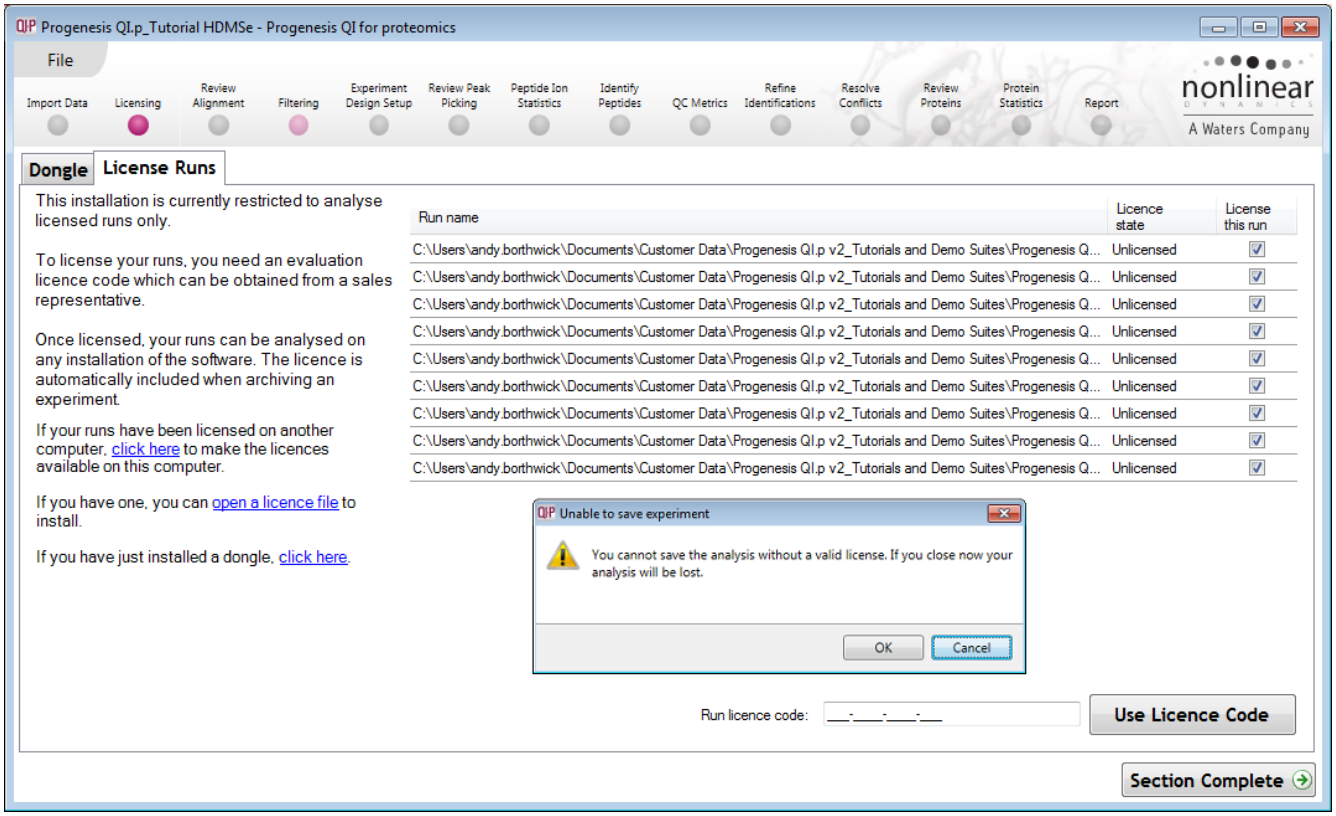
In this example, as the data is HDMSe (Waters) it is going to open at QC Metrics if you chose to perform all the steps in the Automatic Processing Wizard so you can either:

- Continue with the analysis, as the Processing dialog is not displaying any warnings, review the analysis at QC Metrics before proceeding with Refine Identifications. In which case you can go to page 52
- If warnings are displayed or on Review of the Quality Metrics you require to review/redo earlier stages in the analysis then proceed to pages 16 and Appendices 1 and 2.

Stage 3: Licensing

This stage in the analysis workflow will **only** appear if you are using 'Unlicensed' data files to evaluate the software and have no dongle attached.

If you have performed an analysis using Automatic Processing without a valid dongle or do not have the appropriate code to licence your runs, if you close Progenesis QI for proteomics you will be warned that the analysis will be lost.



Note: although the analysis will be lost (Alignment Peak Picking etc) the experiment file and all the 'loaded pre-processed runs are retained'. This allows you to reopen the experiment, when you have a licenced code and/or dongle available, and redo the analysis steps without having to reload the data.

For more details on managing the licencing of your data refer to Appendix 3 (page 82)

If you are using the tutorial archive, this page will not appear as the data files are licensed.

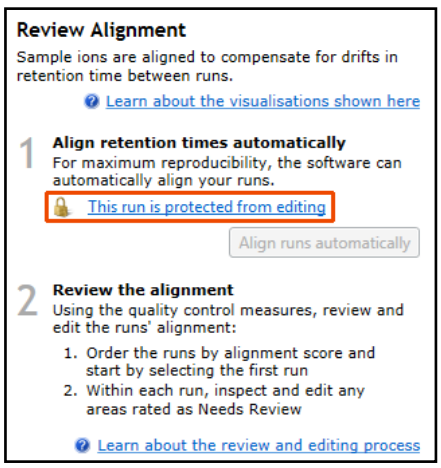
Stage 4: Review Alignment

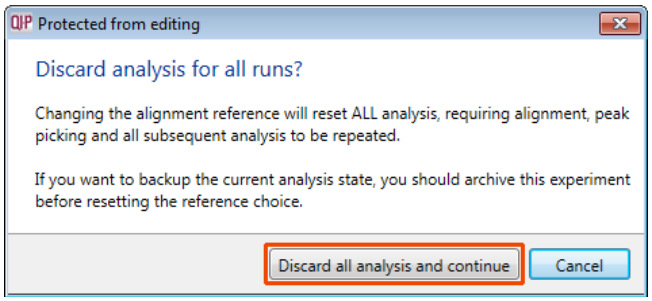
At this stage Progenesis QI for proteomics **Review Alignment** opens displaying the alignment of the runs to the Reference run (A_01).

Having performed the analysis automatically there will be icons next to each run in the vectors table indicating that the run is protected from editing.

In the course of reviewing the quality of alignment you may decide that the alignment requires editing.

To do this click on the **This run is protected from editing** link above the 'greyed out' **Align runs automatically** button.





Note: as you click on the link, you will be warned that you are discarding the current analysis (all steps beyond alignment).

Details on editing alignment are described in Appendix 5 (page 82)

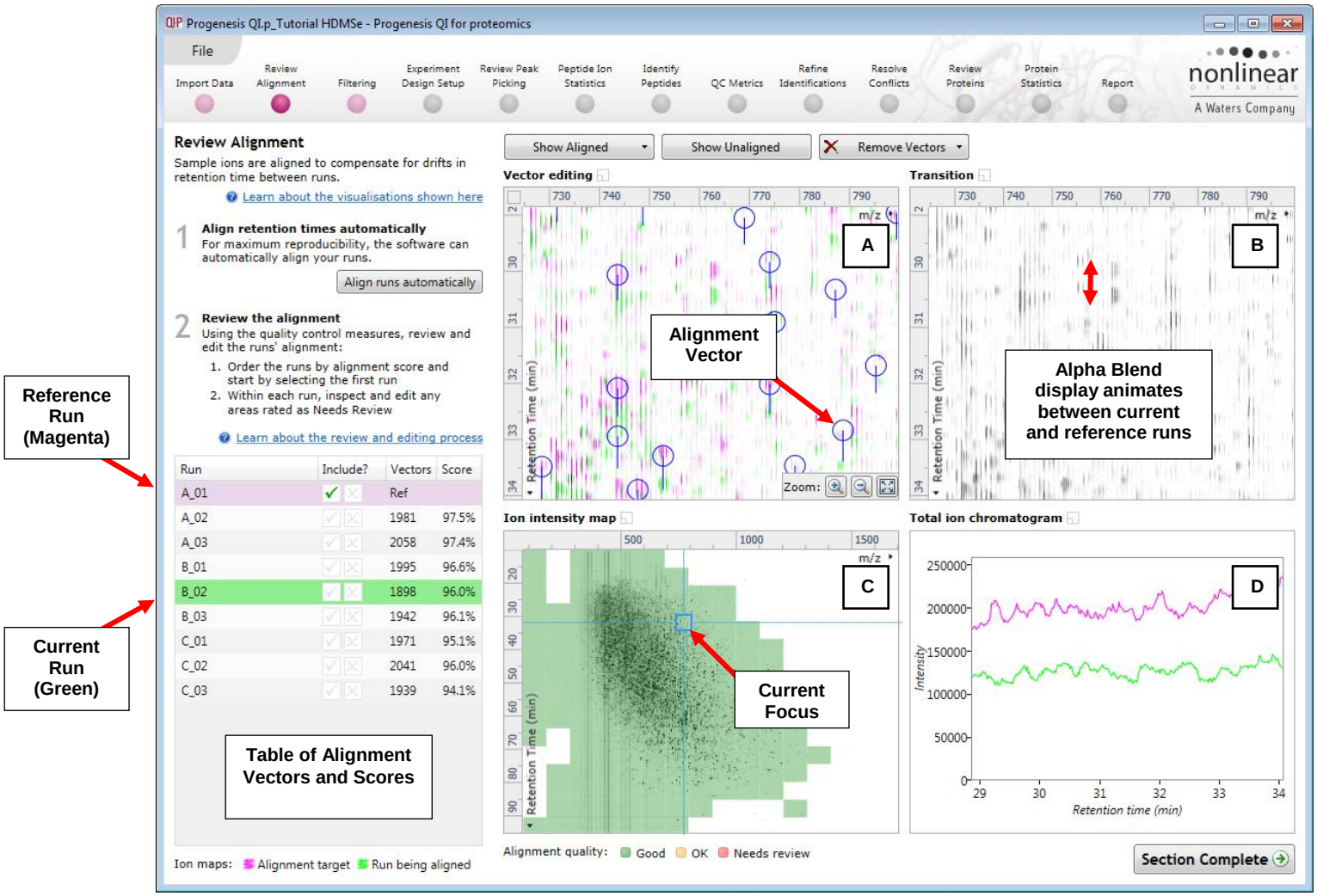
Layout of Alignment

To familiarize you with Progenesis QI for proteomics Alignment, this section describes the various views used in the alignment of the LC-MS runs

Run	Include?	Vectors	Score
A_01	✓	Ref	
A_02		1981	97.5%
A_03		2058	97.4%
B_01		1995	96.6%
B_02		1898	96.0%
B_03		1942	96.1%
C_01		1971	95.1%
C_02		2041	96.0%
C_03		1939	94.1%

To setup the display so that it looks similar to the one below:

- In the Run table click on Run B_02 to make it current. You will now be looking at the alignment of B_02 to A_01 in the Unaligned view. Now drag out an area to review on the **Ion intensity map**. The other 3 views will update to reflect the new focus.



Vector Editing (Window A): is the main alignment area and displays the area defined by the current **focus** rectangle shown in Window C. The current run is displayed in green and the reference run is displayed in magenta. Here is where you can review in detail the vectors and also place the manual alignment vectors when required.

Transition (Window B): uses an **alpha blend** to animate between the current and reference runs. Before the runs are aligned, the peptide ions appear to move up and down. Once correctly aligned, they will appear to pulse. During the process of adding vectors, this view can be used to zoom in and also navigate thus helping with accurate placement of manual vectors.

Ion Intensity Map (Window C): shows the **focus** for the other windows. When you click on the view the blue rectangle will move to the selected area. The focus can be moved systematically across the view using the cursor keys. The focus area size can be altered by clicking and dragging out a new area with the mouse. This view also provides a visual quality metric for the Alignment of the runs, from green through yellow to red.

Total Ion Chromatogram (Window D): shows the current **total ion** chromatogram (green) overlaid on the Reference chromatogram (magenta). As the peptide ions are aligned in the **Vector Editing** view the chromatograms become aligned. The retention time range displayed is the vertical dimension of the Focus Grid currently displayed in the **Ion Intensity map** (Window C).

Reviewing quality of alignment vectors

After **Automatic alignment** the number of vectors and Quality Scores will be updated on the **Runs** panel and the vectors will appear (in blue) on the view.

Where the alignment has worked well, the alignment views will look as below with the Ion Intensity Map showing green indicating good quality alignment and the Transition view showing peptide ions pulsing slightly but not moving up and down.

To simulate poor alignment, place a single manual vector on the Vector editing view (Window A). To do this click and drag out a single vector then release the mouse button. By doing this a single manual vector will appear with a length corresponding to the 'drag'.

Note: the manual vector is **red**, to distinguish it from the automatic vectors (blue)

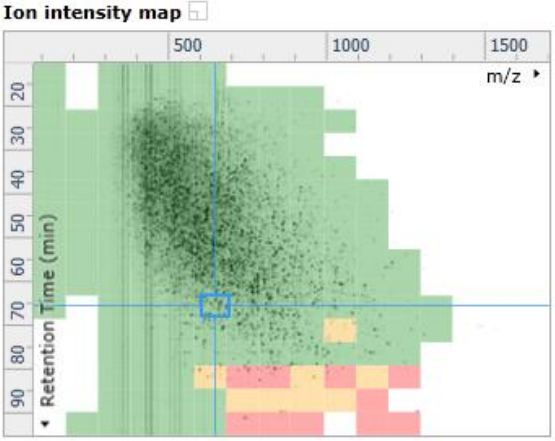
The effect of adding this incorrect manual vector is to reduce the Alignment score and also cause a significant proportion of the Alignment quality squares to turn red on the Ion Intensity Map (as shown below).

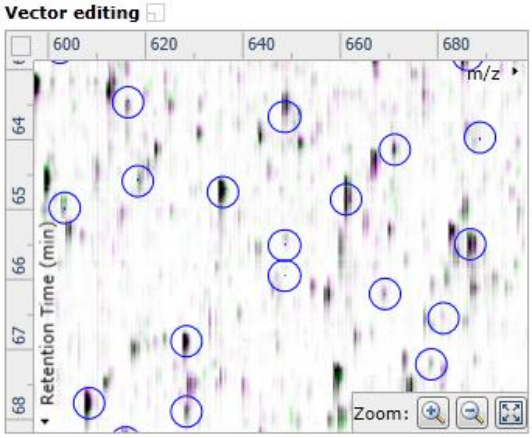
Using a **Simulated** miss-aligned example to explain the review process for alignment, the alignment looks as below with a region of poor alignment (highlighted in red).

Reviewing Quality of Alignment

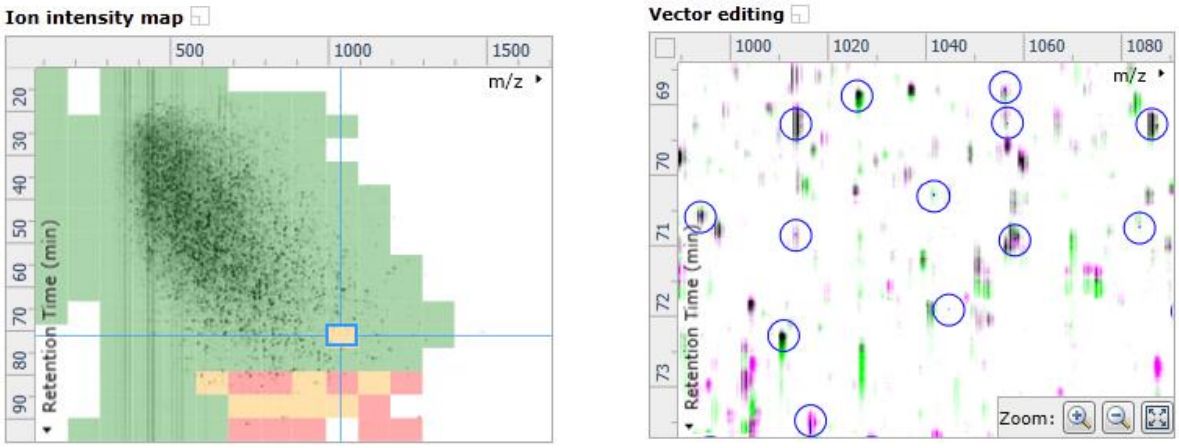
At this point the quality metric, overlaid on the Ion Intensity Map as coloured squares, acts as a guide drawing your attention to areas of the alignment. These range from Good (Green) through OK (Yellow) to Needs review (Red). Drag out a 'Focus' area that corresponds to one of the coloured squares. Three example squares are examined here.

For a 'green' square, the majority of the data appears overlapped (black) indicating good alignment. When viewed in the Transition view the data appears to pulse.

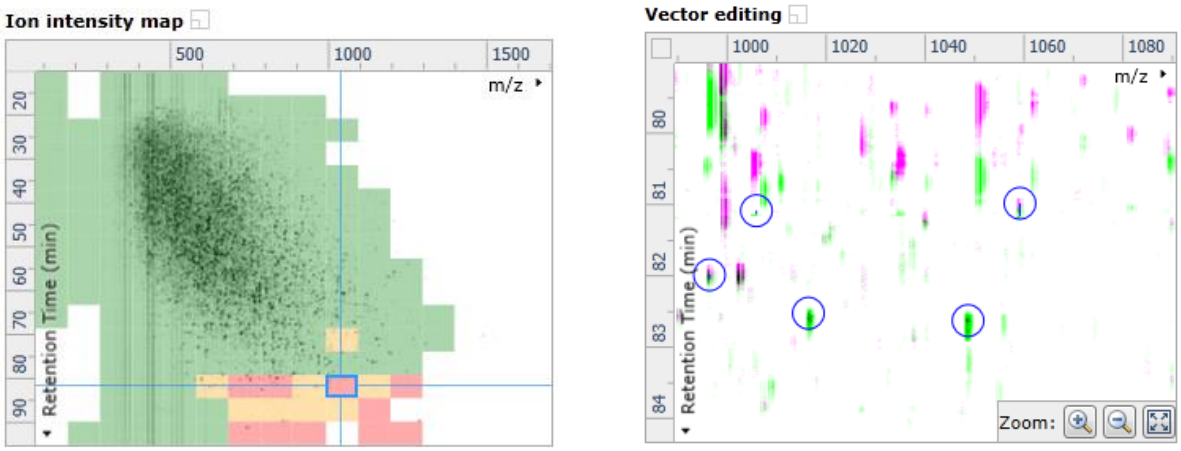




For a ‘yellow’ square some of the data appears overlapped (black) indicating OK alignment. When viewed in the Transition view some of the data appears to pulse.

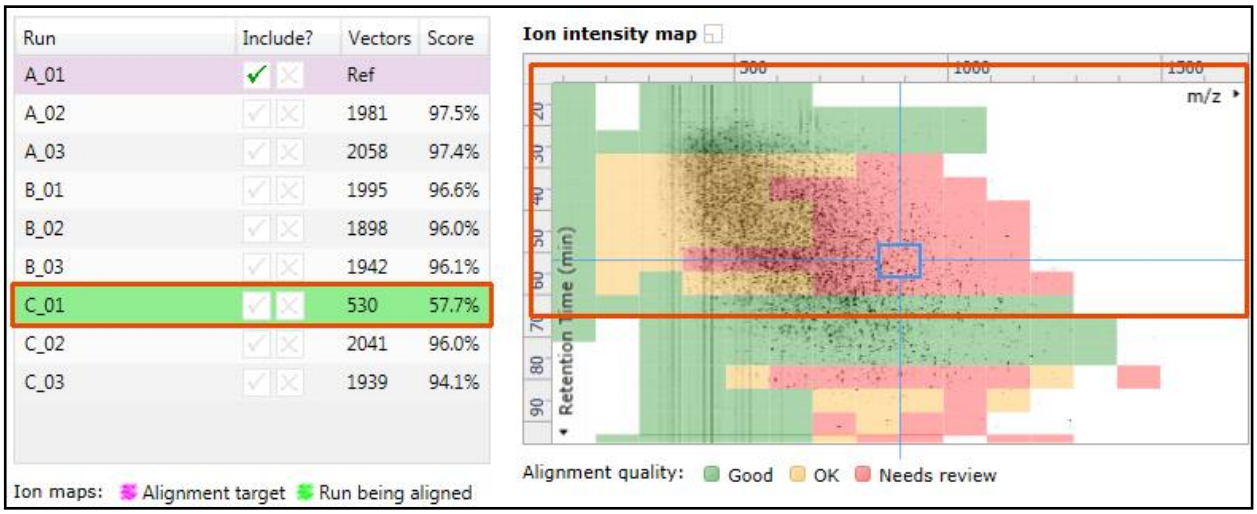


For a ‘red’ square little if any of the data appears overlapped (black) indicating questionable alignment. When viewed in the Transition view little data appears to pulse.



Note: the coloured metric **should be used as a guide**. In cases where there are a few ‘isolated’ red squares this can also be indicative of ‘real’ differences between the two runs being aligned and should be considered when examining the overall score and surrounding squares in the current alignment.

The weighted average of the individual squares gives the overall percentage score for each run’s alignment.



Note: a marked red area combined with a low score clearly indicates a ‘miss alignment’ and may require some manual intervention (see Appendix 4, page 83).

The alignment quality of this tutorial data set does not require any manual intervention so before going to the next section make sure you have removed all manual vectors and re-performed the Automatic alignment.

To do this for C_01 first select 'All vectors in the whole run' from the Remove Vectors and then click Align runs automatically.

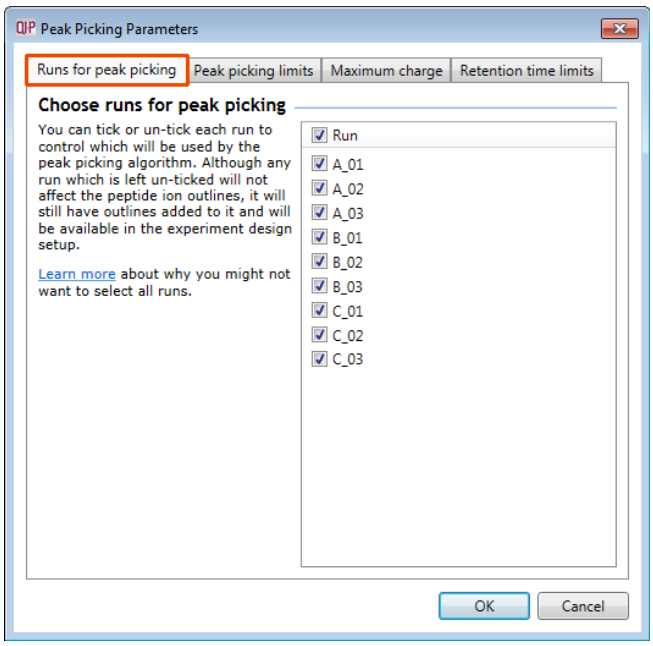
Having aligned the runs automatically, click **Section Complete** to move to Filtering.

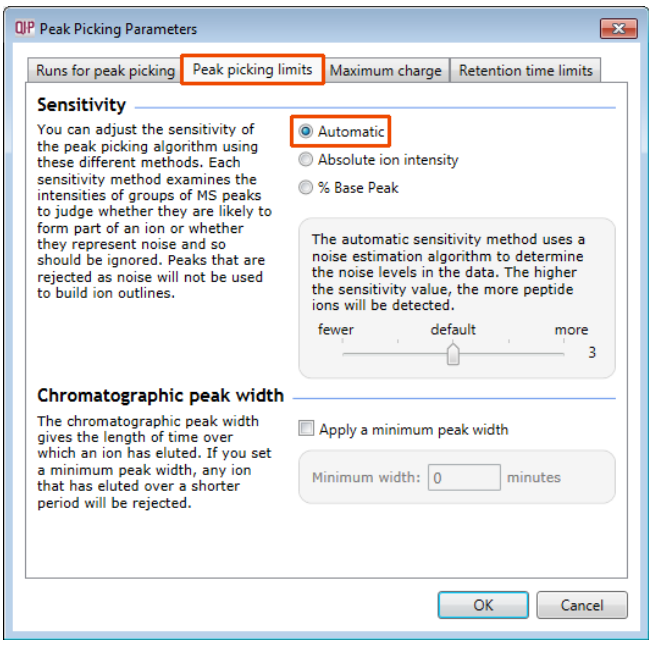
Stage 5A: Filtering

The Peak picking Parameters dialog opens if Picking has not been performed. If it has been performed move to section 5B.

Peak Picking Parameters

The Peak Picking Parameters dialog opens, showing all the runs in the current experiment and a tick against each run. This is the default setting, where the peak picking algorithm uses information from all of the runs to contribute to the pattern of peptide ion outlines.





Tip: It may be appropriate **only** to pick peaks that are present in a limited number of your runs. In which case un-tick the runs that you do **NOT** want to contribute to the peptide ion detection pattern. This may be important when one or more of the runs appear noisy due to non-optimal chromatography or sample handling.

Note: peptide ions outlines will be added to 'un-ticked' runs; although these runs will not contribute to the peak picking pattern.

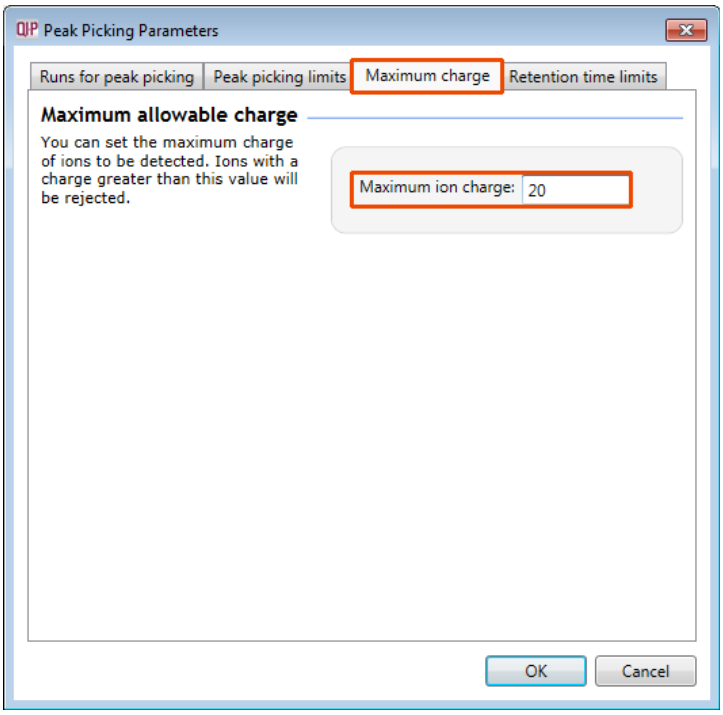
Tip: depending on run quality, a suggested minimum number of ticked runs should include at least one replicate of each experimental condition.

The sensitivity of the detection can be controlled by adjusting settings under the **Peak picking limits** tab.

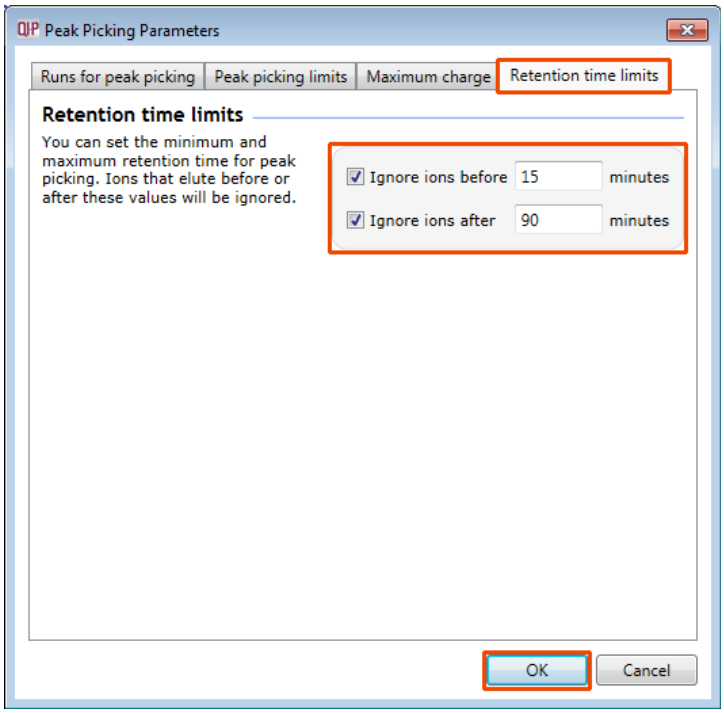
Note: for all 3 **Sensitivity** modes a Chromatographic peak width (Retention time window) for the peaks **can** be set by applying a minimum retention window or peak width in minutes.

For the runs in this user guide, we will use the default settings for the **Automatic** method.

The third tab allows you to set the **Maximum charge** of the ions, which will be detected. The default setting is a charge state of 20.



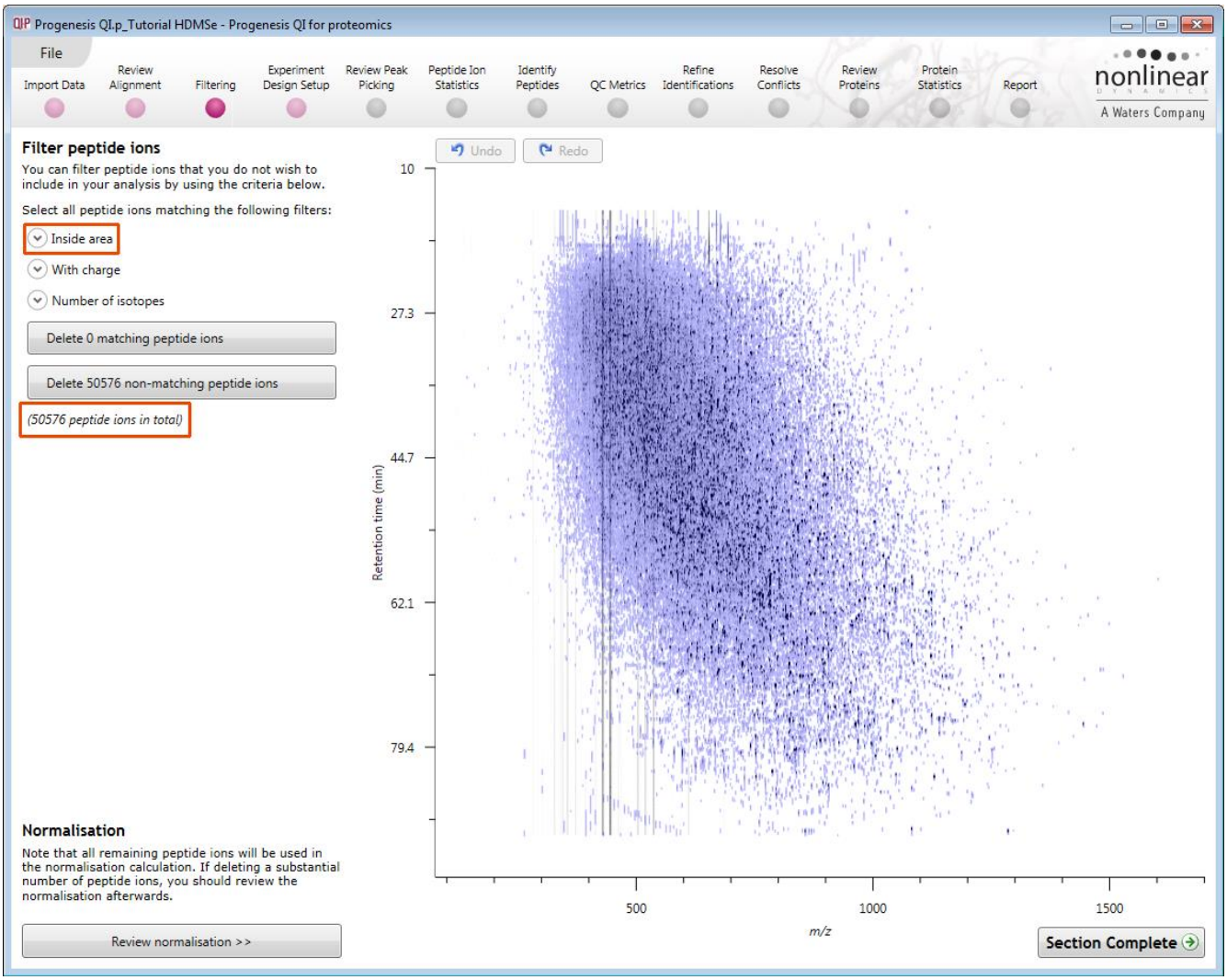
Finally, you can set **Retention time limits** for the detection. The default limits are displayed in the unticked boxes and detection will use these values whenever the box(s) are unticked.



For this dataset we will use the RT settings as shown above 15 and 90 minutes .

Press **OK** to start the detection process.

On completion of analysis, the Filtering stage will open displaying the number of peptide ions detected, in this example 50576.



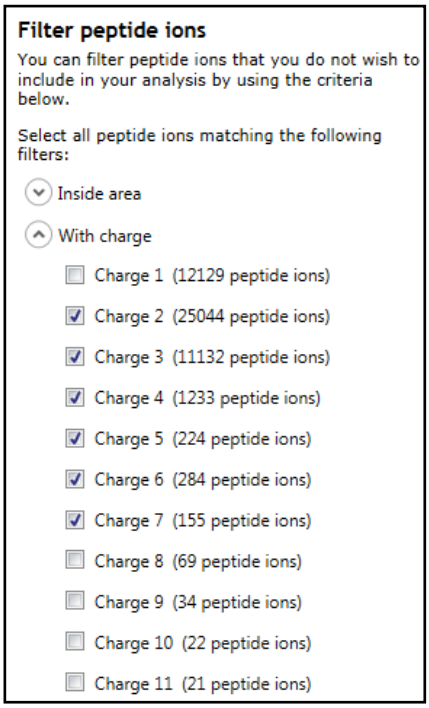
If required you can remove peptide ions based on position, charge state, number of isotopes or combinations of these peptide ion properties.

As an example, we will filter the peptide ions based on charge 'charge state'.

When **With charge** is selected the number of peptide ions present at each charge state is displayed, these can be selected accordingly. In this case we will retain peptide ions with a charge state of 2 to 7.

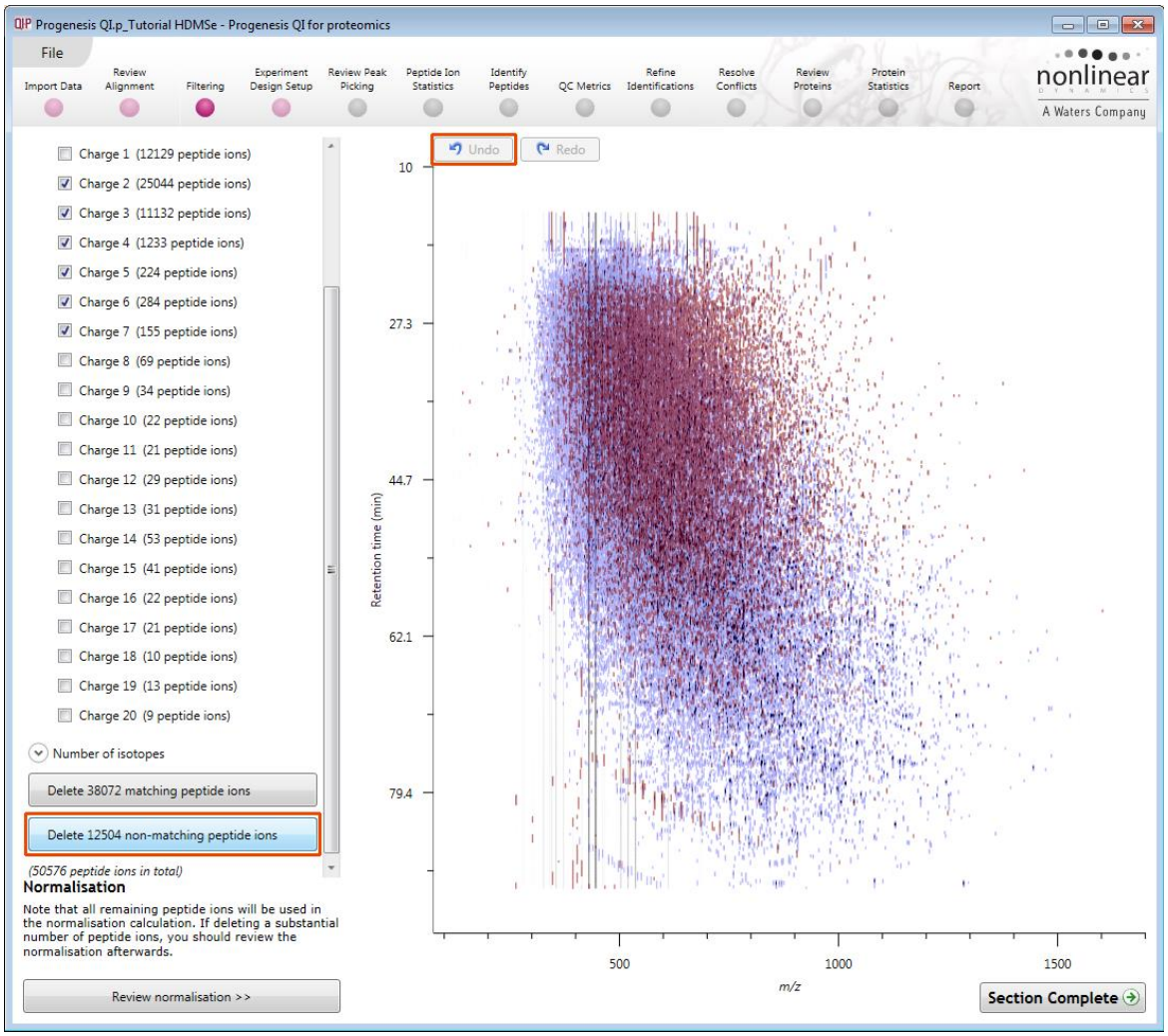
Area limits, charge state and number of isotopes can be combined to refine the peptide ion selection.

Tip: when filtering on one property of the peptide ion i.e. charge state, make sure you have 'collapsed' the other filters (see right)



Hence all peptide ions with a charge state of 1 or 8 and above will appear red on the main view as you hold the cursor over the appropriate delete button.

To remove these peptide ions press **Delete 12504 Non Matching Peptide ions.**



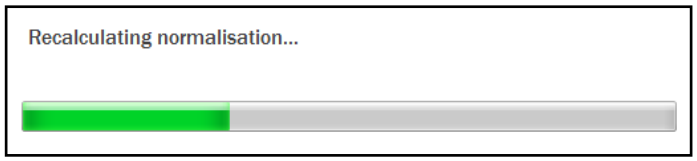
You can use the **Undo** button to bring back deleted peptide ions, however, when you move to the next section you will lose the capacity to undo the filter. Before moving on from filtering, you can review the normalisation of the experiment.

Tip: When you have reached the filtering stage, it is good practice to close the experiment and save an archive. This can be used to restore the unfiltered state if the filtering you have performed is too 'stringent'.

Stage 5B: Reviewing Normalisation

Normalisation review is accessed from the button at the bottom left corner of the filtering page

If you have filtered out a number of peptide ions from the original detection pattern then the normalisation will update.

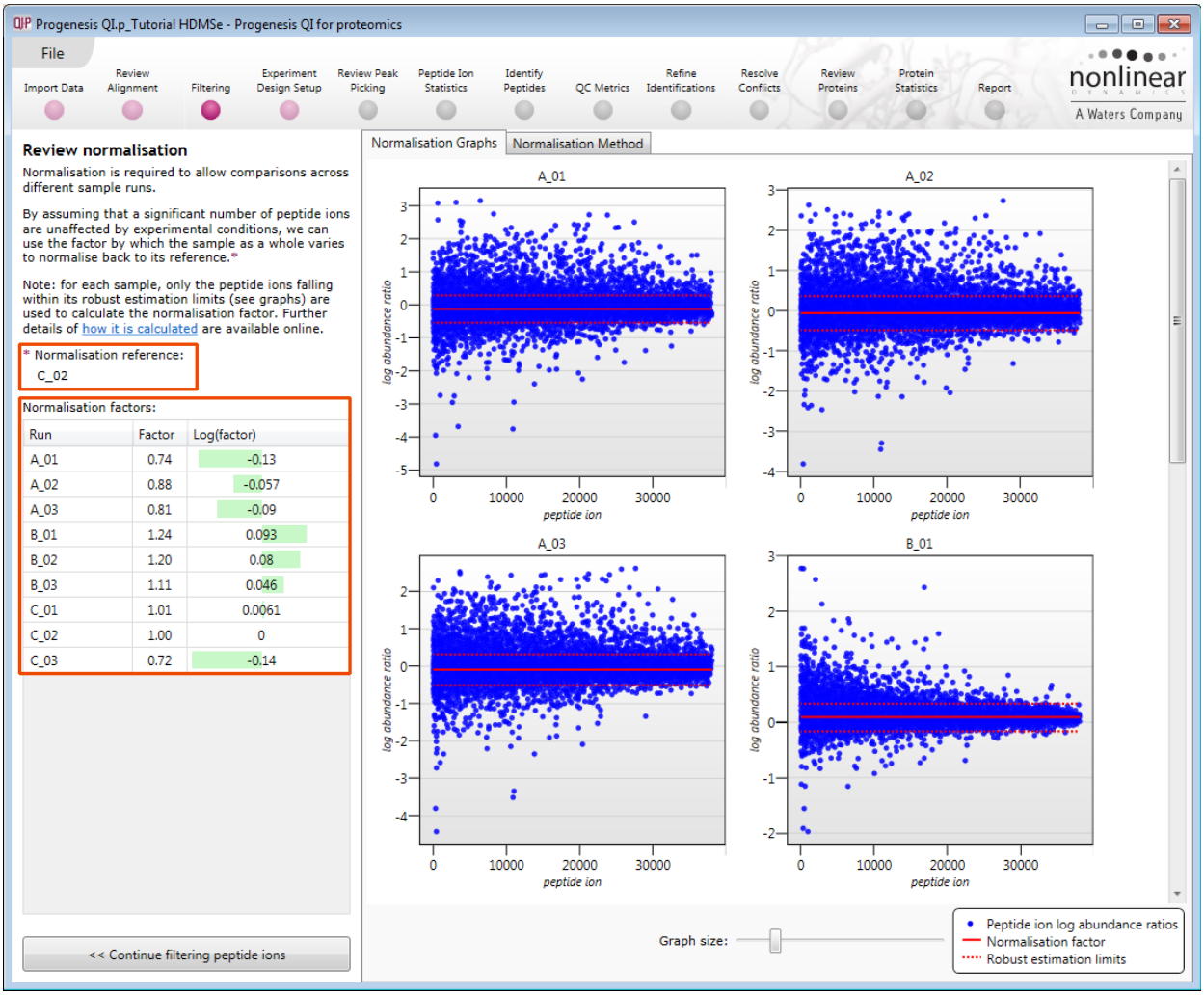


The **Review Normalisation** page will open displaying plots for the normalisation of all the peptide ions on each run.

This page in the workflow **does not** allow you to alter the Normalisation of your data but provides you with individual views for each run showing the data points used in the calculation of the normalisation factor for the run.

Alternatively, if you do not believe normalisation is necessary, you can opt to 'Don't use any normalisation' for the rest of the analysis (Normalisation Method tab).

Normalisation factors are reported in the table to the left of the plots.

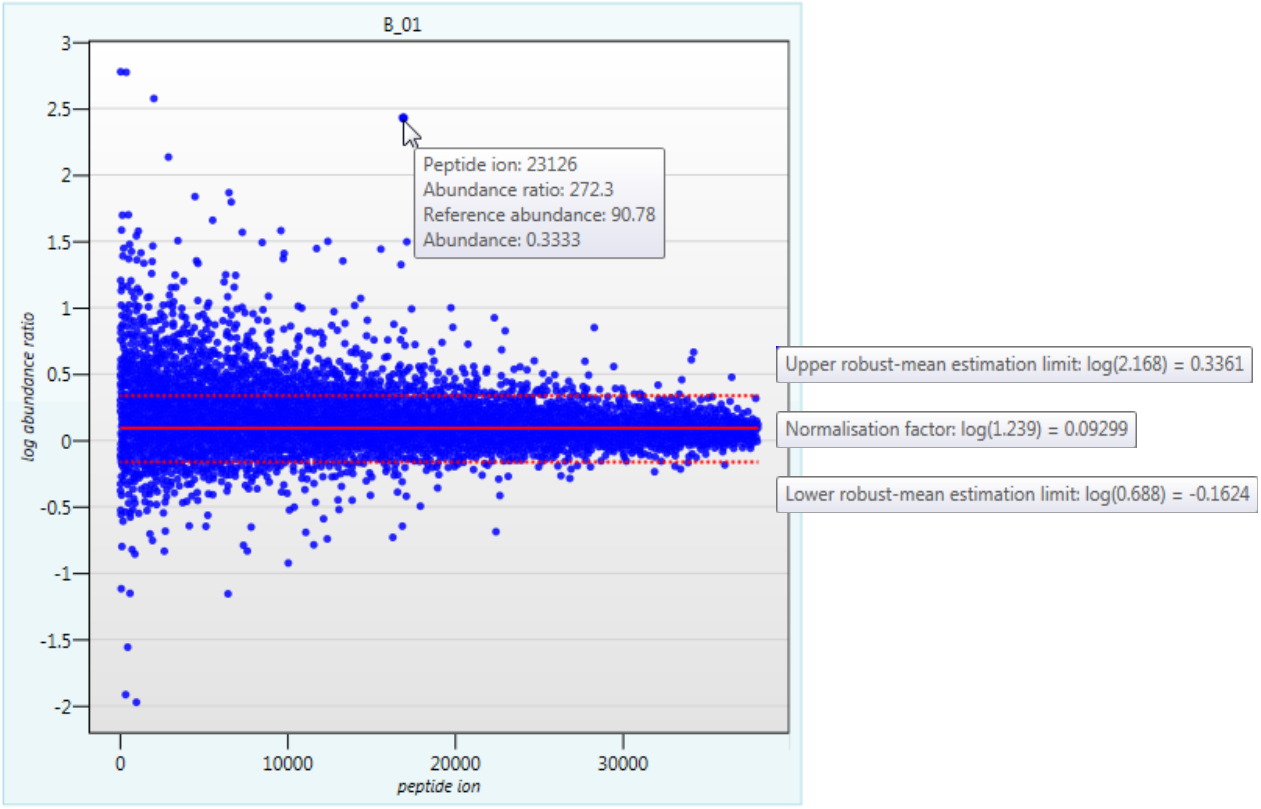


Calculation of Normalisation Factor:

Progenesis QI for proteomics will automatically select one of the runs that is 'least different' from all the other runs in the data and then set this to be the 'Normalising reference'. The run used, is shown above the table of Normalisation factors (in this example C_02).

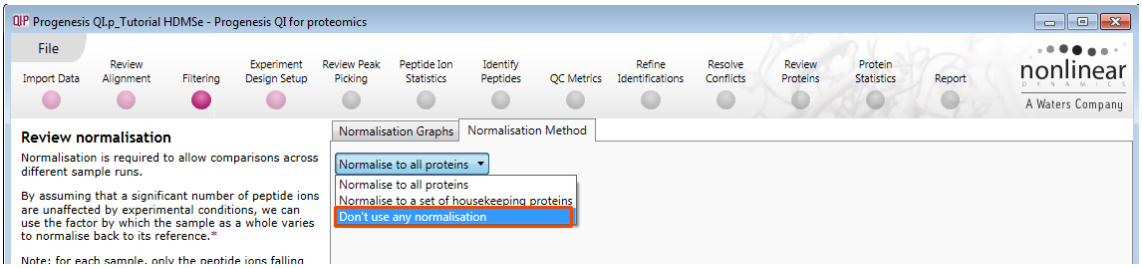
For each sample run, each blue dot shows the log of the abundance ratio for a different peptide ion (normalisation target abundance/run abundance).

The details for individual peptide ions can be viewed as you hold the cursor over the dots on the plot.



On the graph the peptide ions are shown ordered by ascending mean abundance. The normalisation factor is then calculated by finding the mean of the log abundance ratios of the peptide ions that fall within the 'robust estimated limits' (dotted red lines). Peptide ions outside these limits are considered to be outliers and therefore will not affect the normalisation.

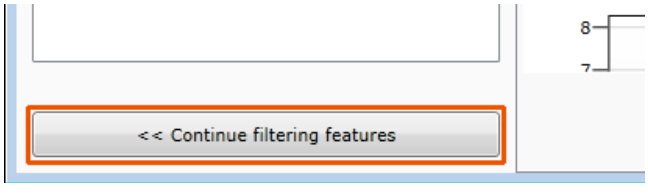
Finally, if you do **not** wish to work with normalised data then you can **use the raw abundances** by switching off the normalisation.



Note: once you have identified the peptide ions, you can then apply the **Normalise to a set of house keeping proteins** by using this option to locate and select the peptide ions.

For this experiment, you should leave the **Normalise to all proteins** option selected.

Now return to filtering by clicking on the button on the bottom left of the screen



For this example, we **DO NOT** do any additional Filtering so click on **Section complete**.

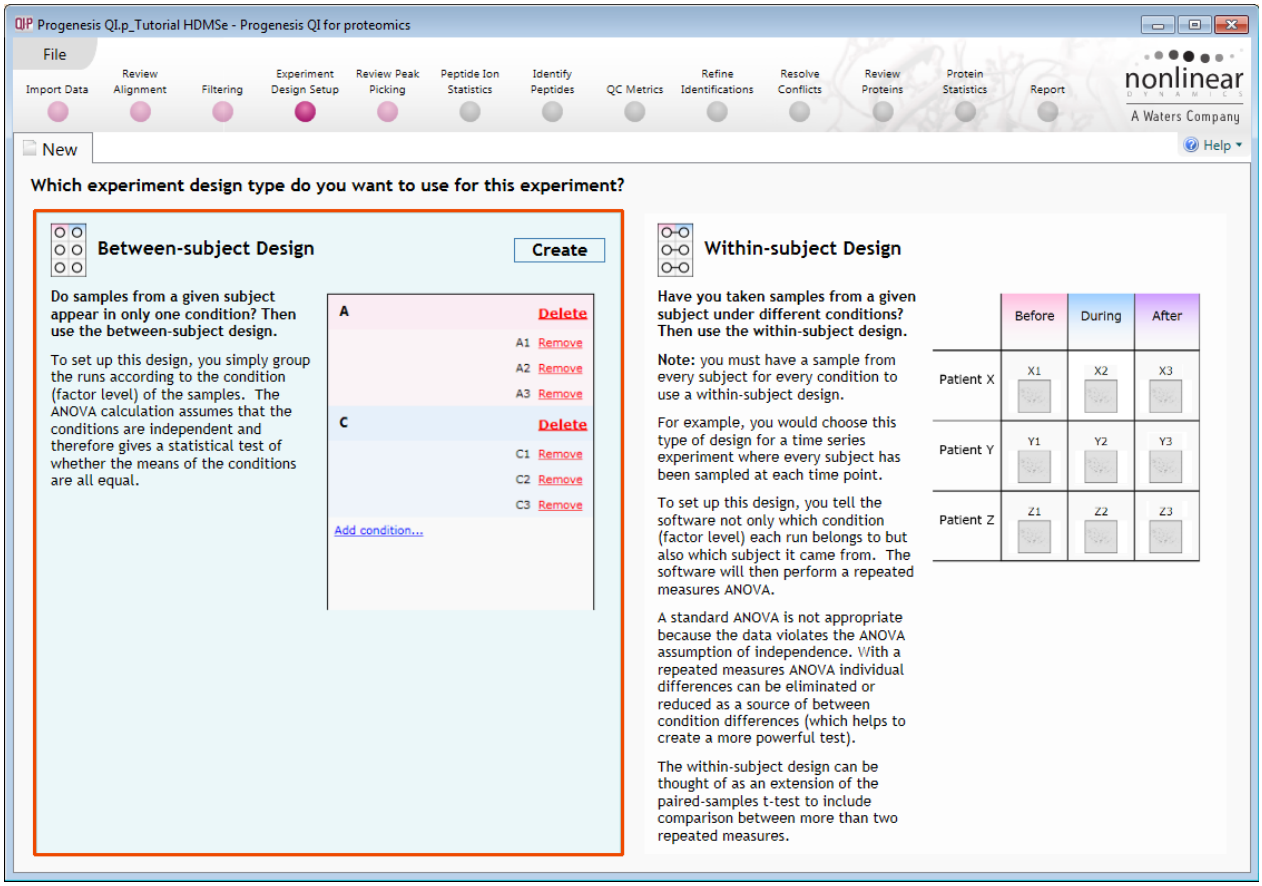
Note: if you do any extra filtering then **Normalisation recalculates** as you move to the next stage in the Workflow**.**

Stage 6: Experiment Design Setup for Analysed Runs

At this stage in the workflow you can setup one or more experimental designs for your data.

There are two basic types of experimental designs:

Between-subject design: here samples from any given subject appear in only one condition. (i.e. control versus various drug treatments).The ANOVA calculation assumes that the conditions are independent and applies the statistical test that assumes the means of the conditions are equal.

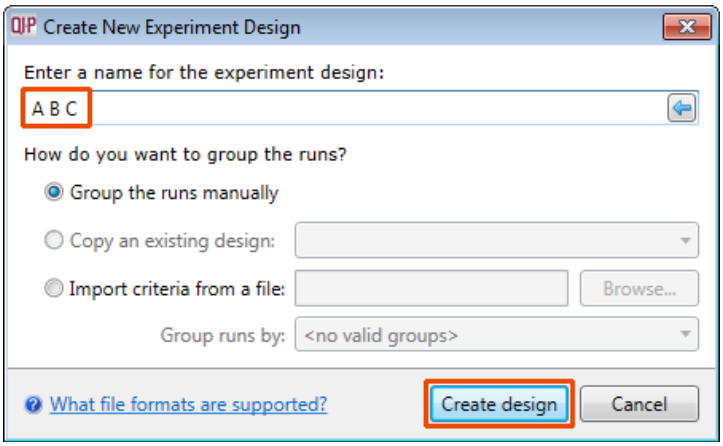


Within-subject design: here samples have been taken from a given subject under different conditions (i.e. the same subject has been sampled over a period of time or after one or more treatments). Here a standard ANOVA is not appropriate as the data violates the ANOVA assumption of independance. Therefore by using a *repeated measures* ANOVA, individual differences can be eliminated or reduced as a source of between condition differences. This within-subject design can be thought of as a extension of the paired samples t-test, including comparison between more than two repeated measures.

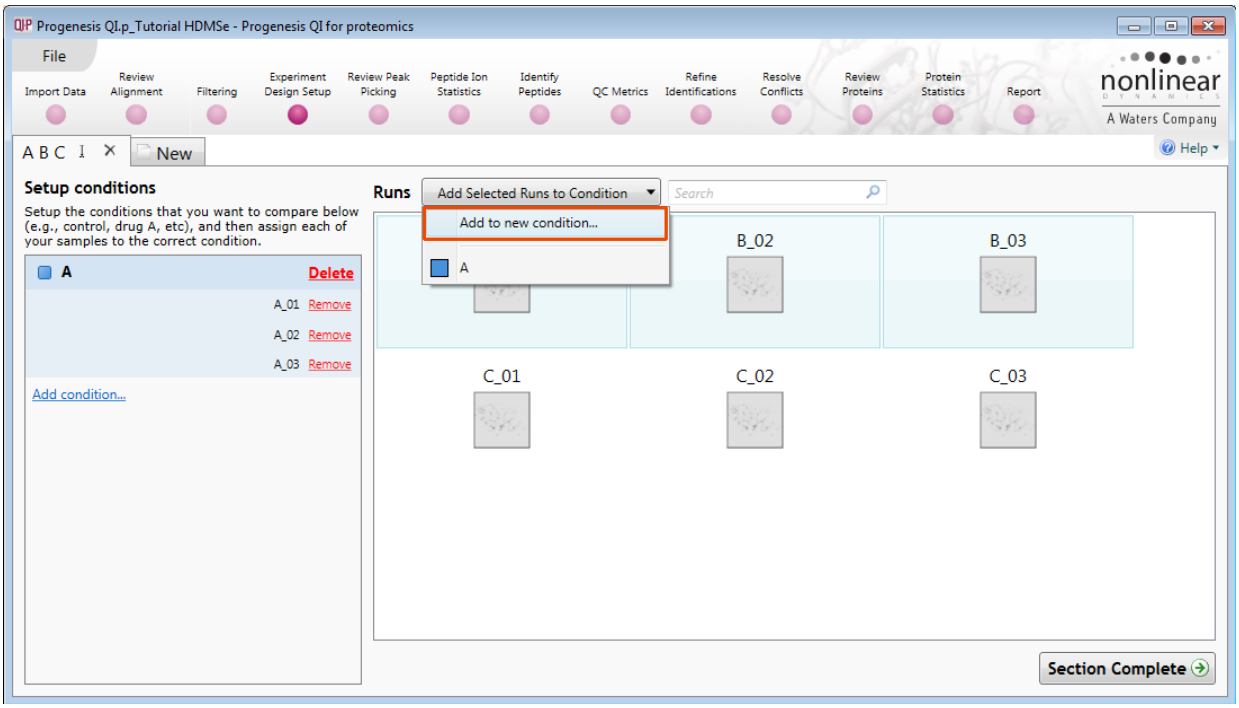
Additional information on how to apply the Within-subject Design is in Appendix 5 page 88

This experiment contains 2 conditions: Control and Treated and uses the **Between-subject design** to group the analysed runs to reflect the Biological conditions in the original study.

To create a new **Between-subject Design** hold the cursor over this option and click to open the dialog.



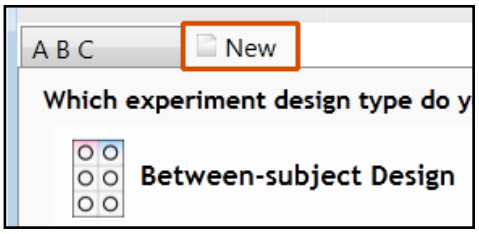
Give the new experiment design a name and then click **Create design**.



To create a new condition

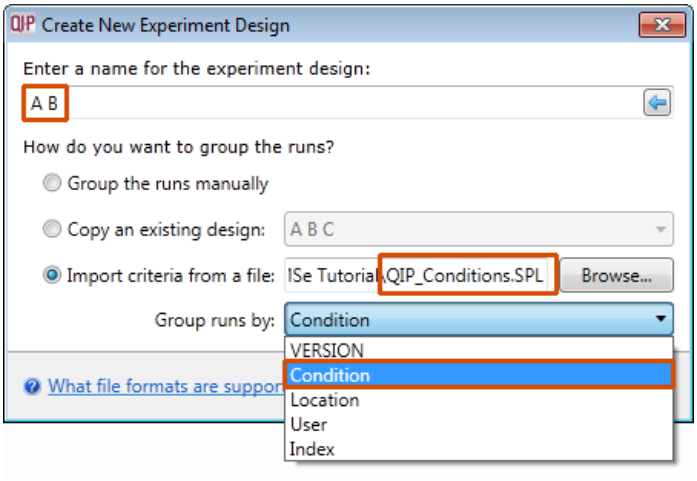
1. Select the runs for the condition by clicking on the required icon in the **Runs** panel, as shown.
2. Press the ‘black triangle’ next to the **Add Selected Runs to Condition** button on the main toolbar.
3. Select **Add to new condition**… from the drop down menu.
4. A new condition will appear in the **Conditions** panel on the left.
5. Rename the condition (e.g. C) by over typing the default name
6. Repeat steps 1 to 5 until all the runs are grouped into conditions.

An alternative way to handling the grouping of this set and other larger (and more complex) experimental designs is to make use of **sample tracking information** that has been stored in a spread sheet at the time of sample collection and/or preparation.



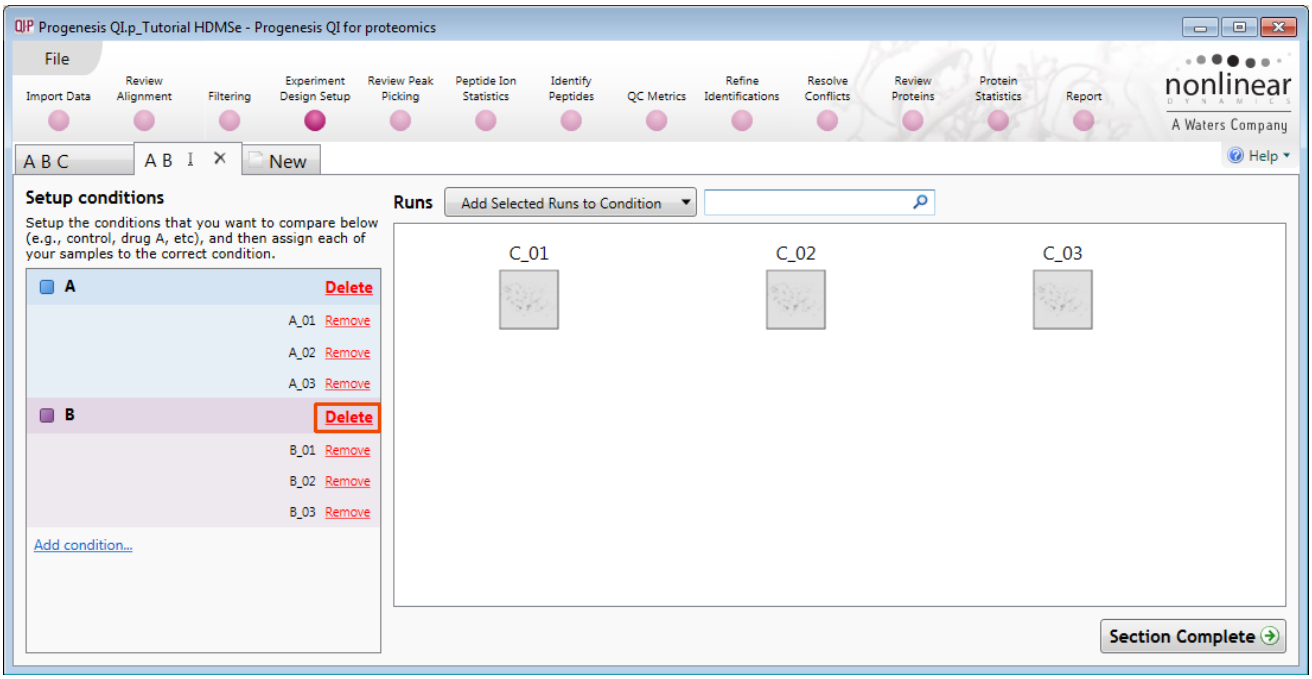
For this example there is a **QIP_Conditions.spl** file available in the Experiment Archive you restored at the beginning of this tutorial exercise.

To use this approach select the **Import design from file** option from the New Experiment Design dialog. Then locate the QIP_Conditions file and select what to **Group by**, for example: **Condition**.



When **Create design** is pressed the new tab refreshes to allow you to adjust the conditions.

Use **Delete** on the Conditions panel to remove conditions that are not required in this particular design.



Note: On deleting a condition the runs will reappear in the Runs window.

Note: both designs are available as separate tabs.

To move to the next stage in the workflow, **Review Peak Picking**, click **Section Complete.**

Stage 7: Review Peak Picking and editing of results

The purpose of this stage in the Workflow is to review the list of peptide ions using the visual tools provided and edit peptide ions if required.

The review stage has 5 display modes: 1D, 2D, 3D, Drift Time and Peptide ion Details controlled by the tabs on the bottom left of the display and the expander bar to the right of the table. Each display has multiple views to allow comparative exploration of the detected peptide ions on the aligned LC-MS runs.

Exploring analysed data using the Data displays

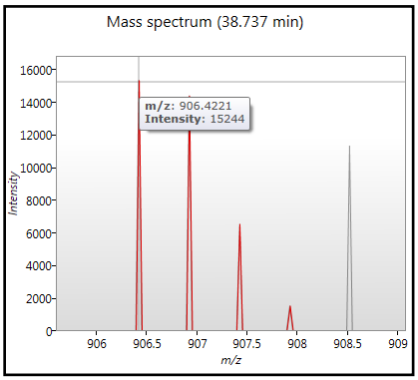
Window A: shows the list of peptide ions ranked by the p value for the one way **Anova** using the current Experiment Design (A B C).

Note: a value of 'Infinity' in the **Fold** column indicates 'Presence/Absence'

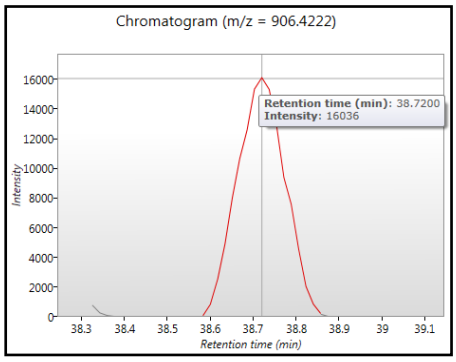
Note: by default all the peptide ions are included in the selection for the next section of the analysis.

To highlight a group of peptide ions drag out a selection on the table.

The 1D Display

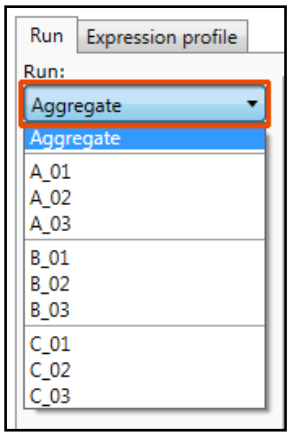


Window B: displays the Mass spectrum for the current peptide ion on the selected Run (in window D). Hold the cursor over the peak to indicate m/z and intensity



Window C: displays the Chromatogram for the current peptide ion on the selected Run (in window D). Hold the cursor over the peak to indicate Retention time and intensity.

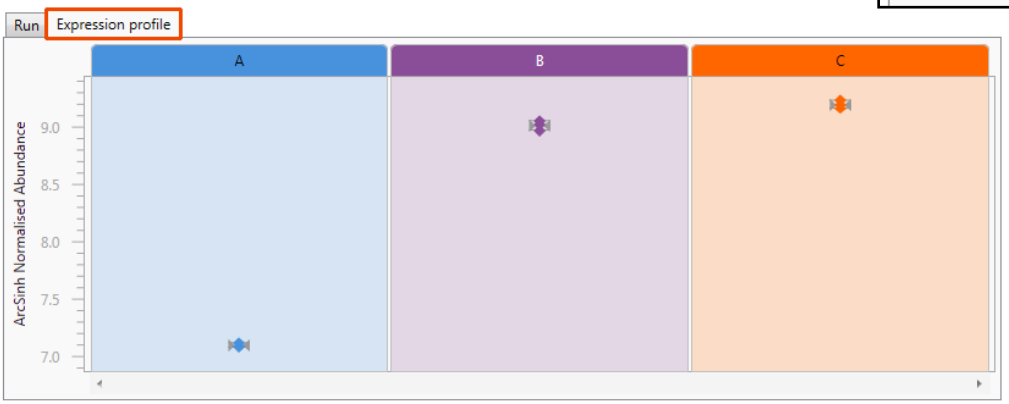
Window D: displays the details of the currently selected run. By default the selected run is an Aggregate view of all the aligned runs.



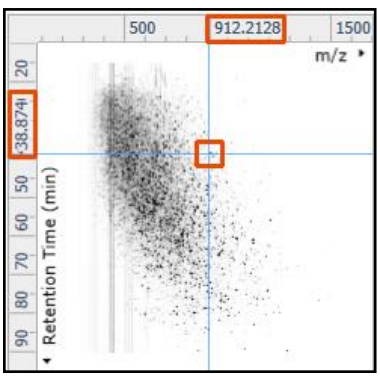
Details of individual runs can be viewed by using the 'Run' link and selecting the run you wish to view.

The peptide ion editing tools are located in this window (see page 41 for functional explanation).

Clicking on the Expression Profile tab in Window D shows the comparative behaviour of the peptide ion across the various biological groups based on group average normalised volume. The error bars show +/- 3 standard errors.



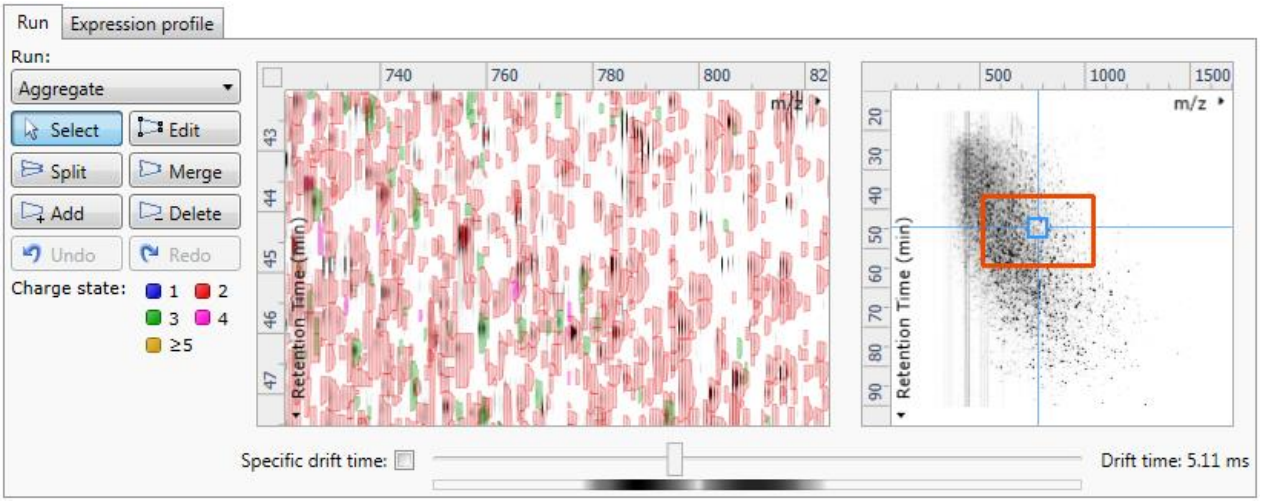
Window E: shows where the current peptide ion is located on the LC-MS run by means of the cross hairs.



To change the current location, click on the image of the run (**note**: the retention time and m/z values update as you move the cursor around this view).

Note: doing this updates the focus of all the other windows.

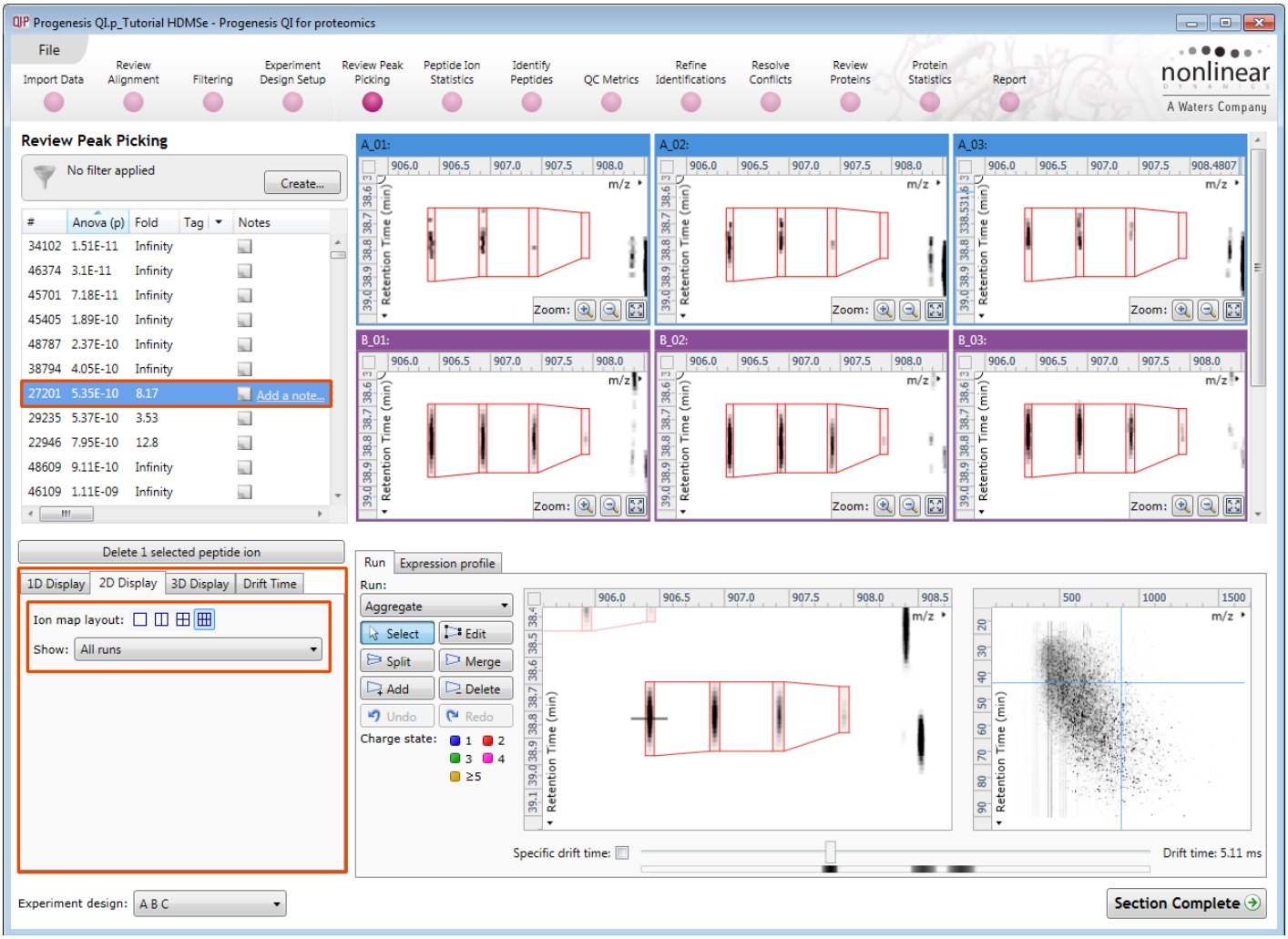
You can also drag out an area (blue square) on this view that will re-focus the other windows.



The 2D Display

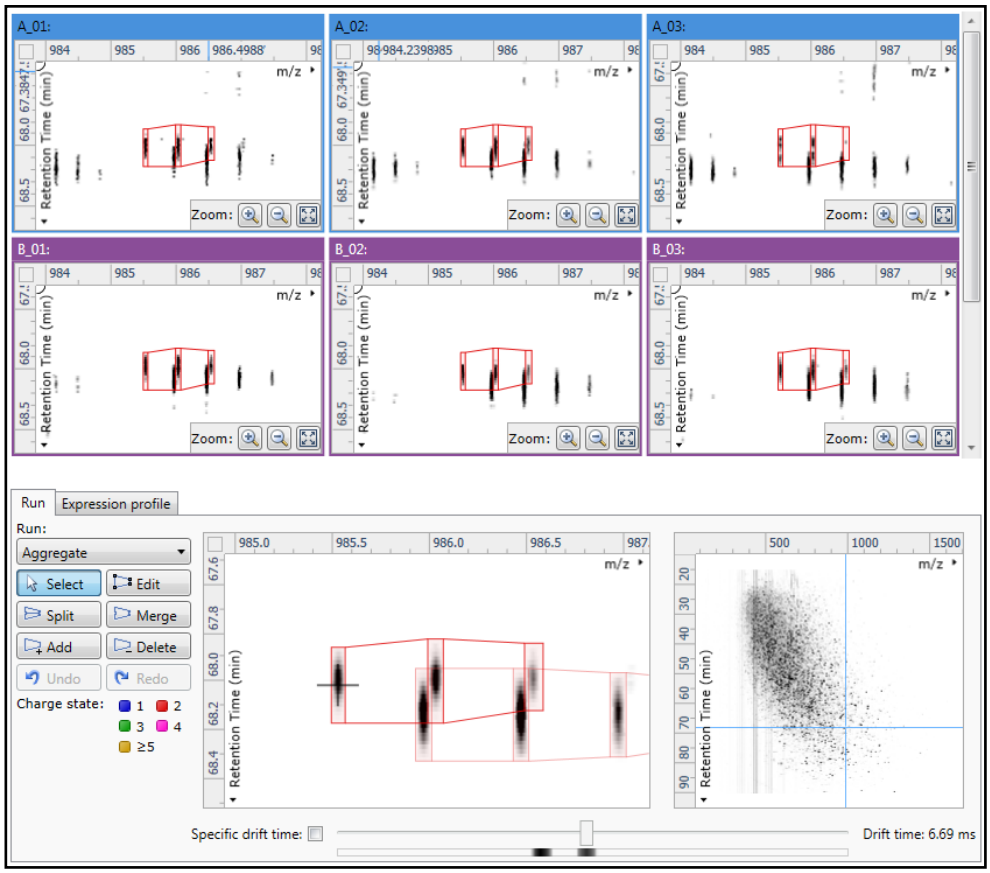
Windows **A, D and E**: perform the same functions across all 4 display modes.

In the 2D Montage mode, Window B displays a montage of the current peptide ion across all the aligned LC-MS runs.



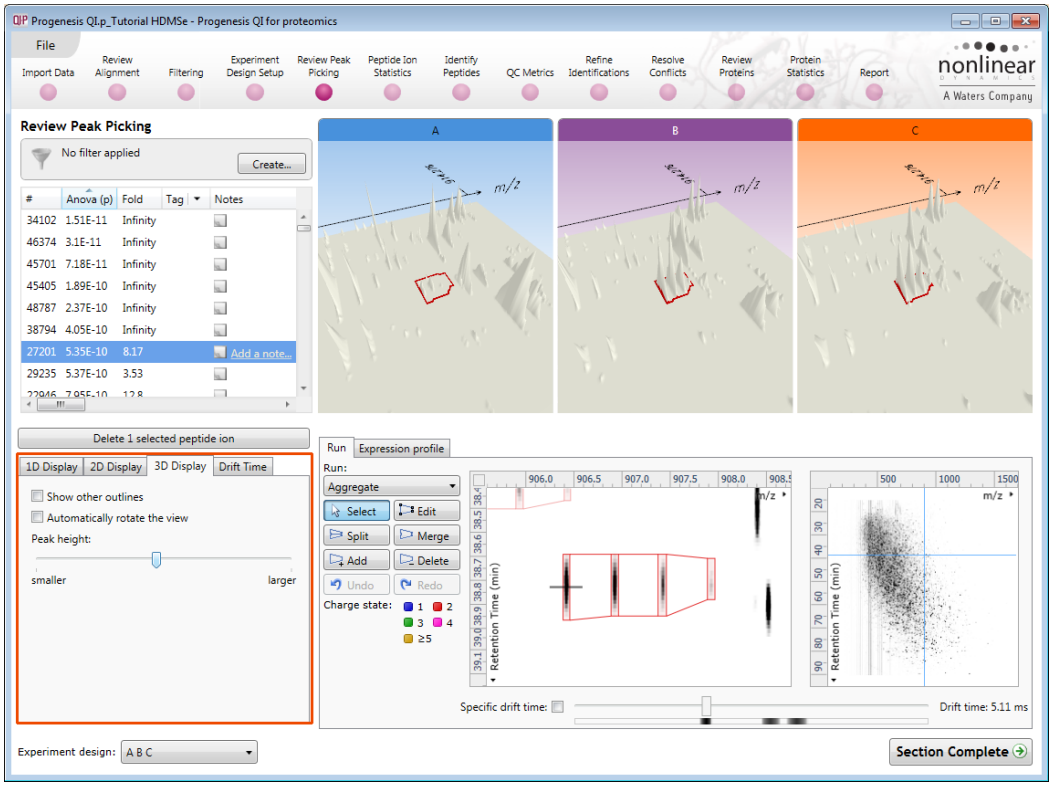
The appearance of the Montage (window B) is controlled by the panel on the bottom left of the display.

Using the the various views in the 2D display one can examine the peptide ion detection in detail to validate the correct detection of even fully overlapping peptide ions as shown above.



The 3D Display

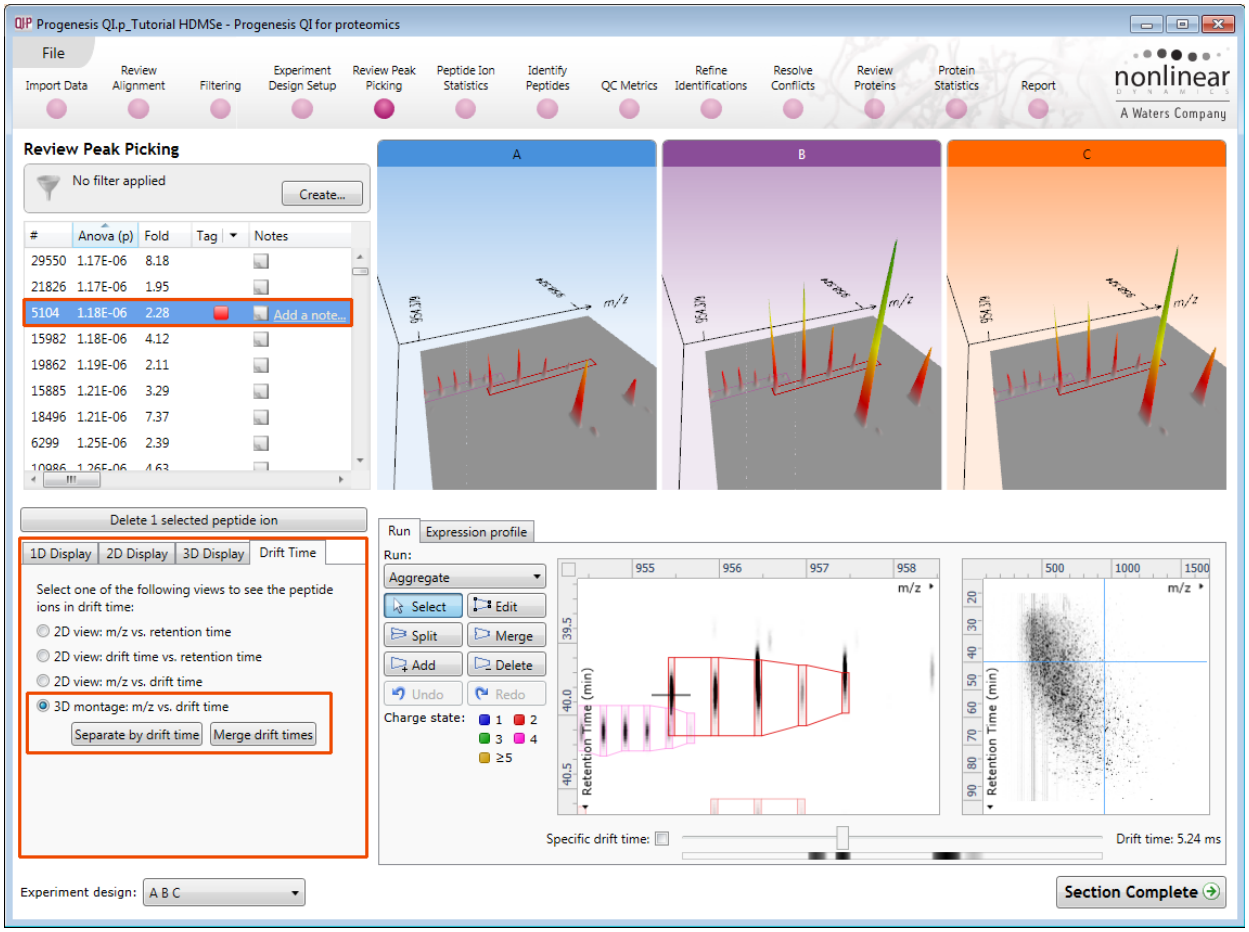
Window B changes into a 3D view by selecting the 3D Montage tab on the bottom left of the display.



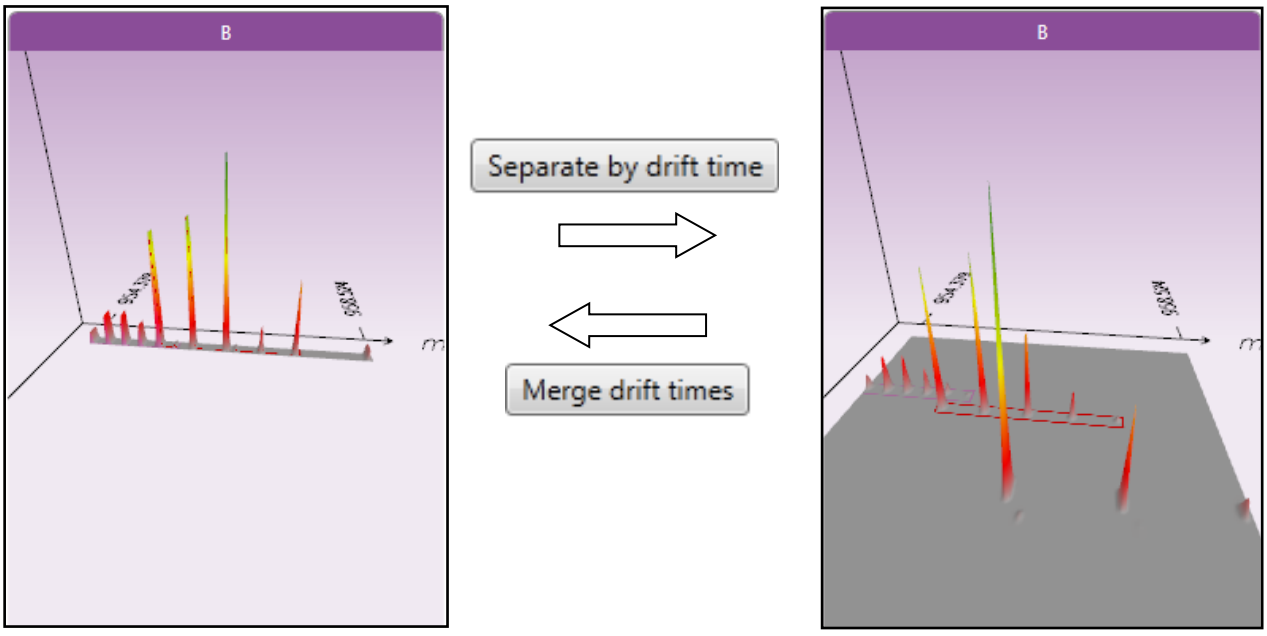
The views can be set to **Rotate** automatically or you can rotate them manually by clicking and dragging them with the mouse.

Drift Time Display

When the Drift Time tab (**F**) is selected in the bottom left of the display Window B changes into a 3D view displaying the Drift Time dimension.



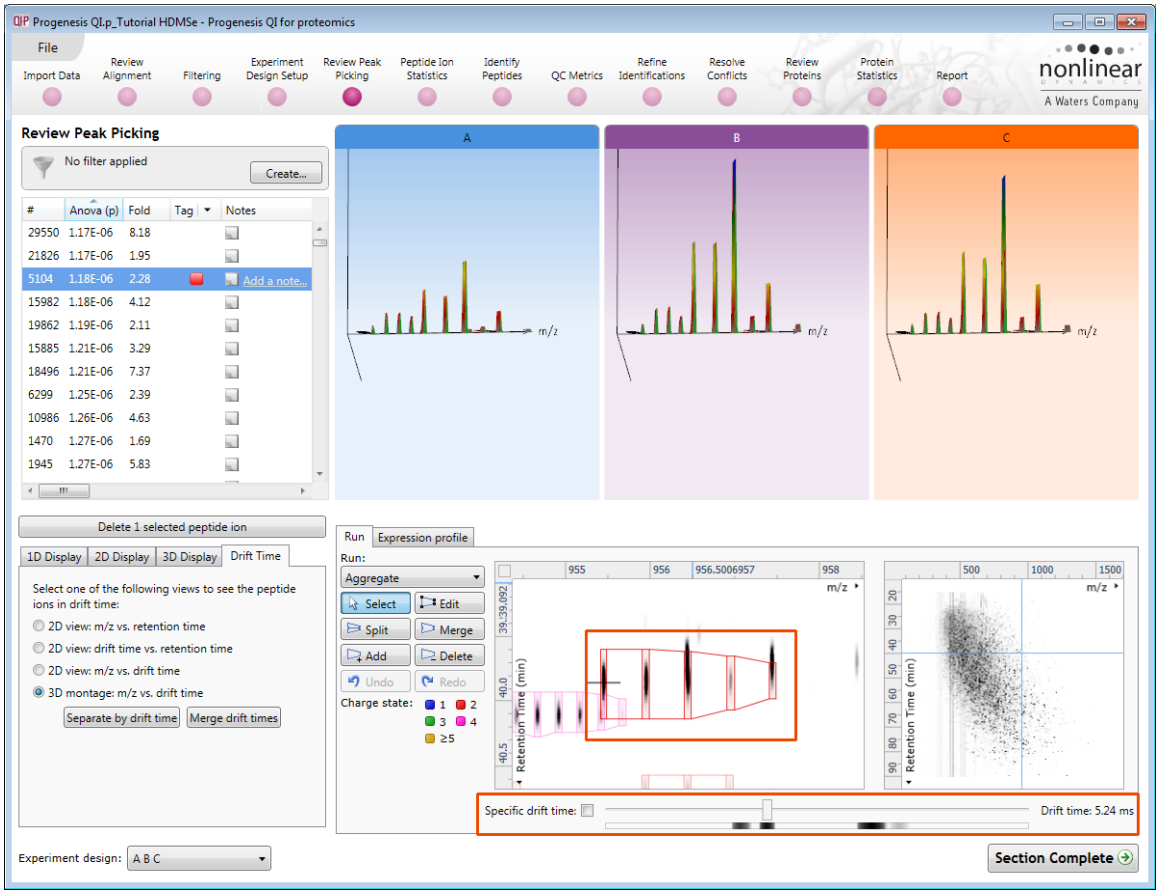
To view the drift time dimension for the current peptide ion, click on **Separate by drift time** on the Drift Time tab. This will expand the view showing the drift time separation between the detected peptide ions (below right)



The left hand view displays the merged Drift Times.

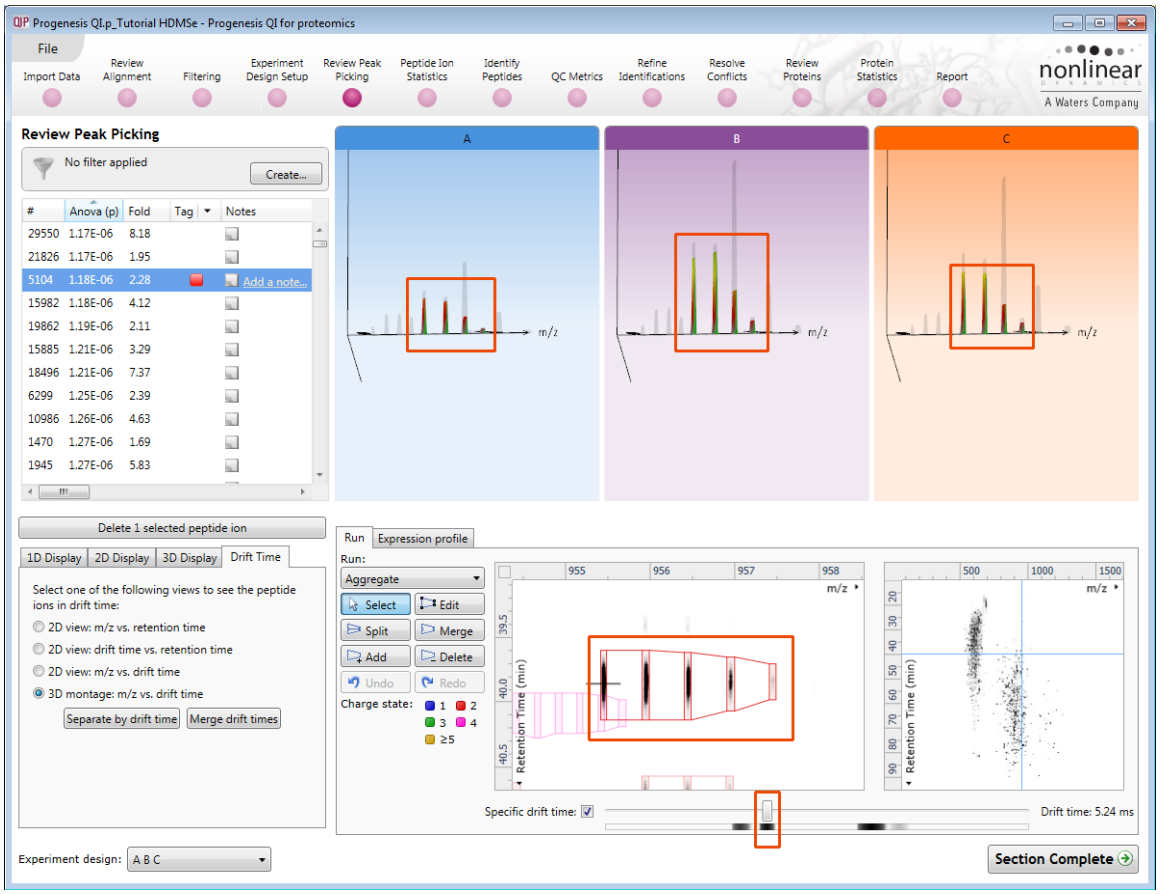
Note: you can manipulate the orientation of the views by clicking on them and dragging the display to the required orientation. You can also zoom in and out of the panels by using the 'scroll' wheel on your mouse

Note: you can step through the specific **Drift Times** (measured in milli seconds) for the current peptide ion by clicking on the **Specific drift time** tick box at the bottom of the display.

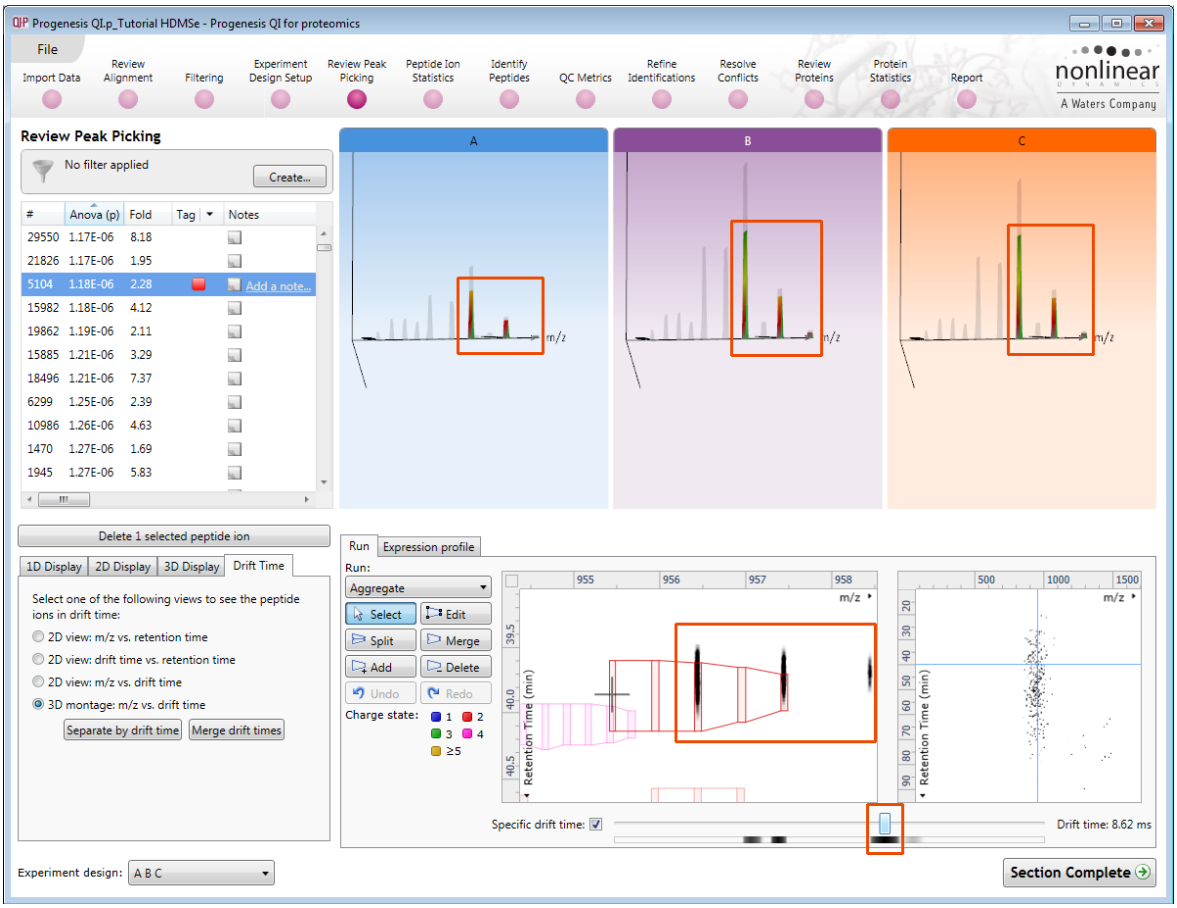


The ‘crosshairs’ on the peptide ion in the **Run** view identifies it as the current peptide ion in the table.

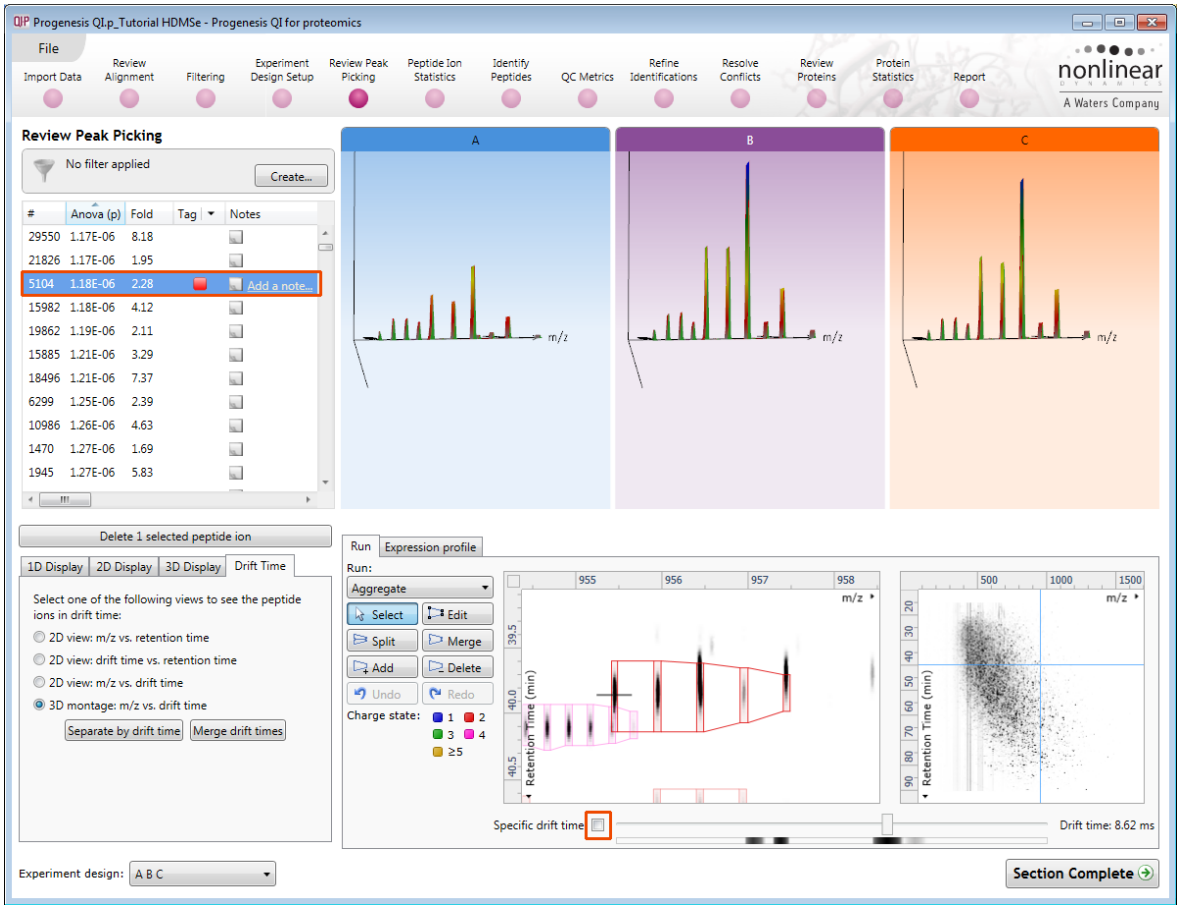
As you move the slider over the intense areas, indicated below, all the views update to the corresponding drift time.



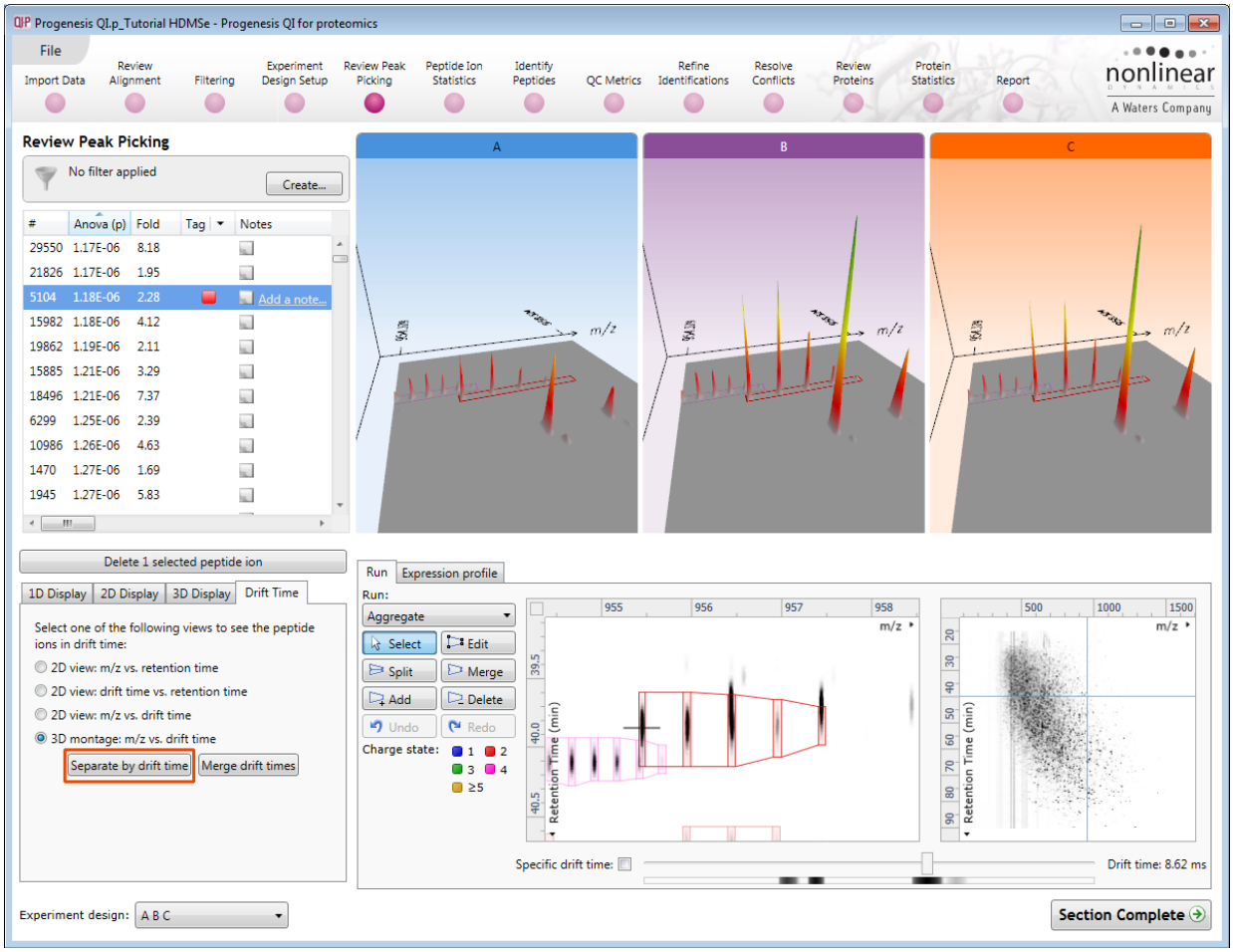
Note: the crosshairs will remain on the original peptide ion in the table as you explore the Specific Drift times



When you un-tick the ‘Specific drift time, tick box the 3D views will return to showing the Merged Views for the current peptide ion in the table.



Click Separate **by drift time** to view the drift times in 3D.

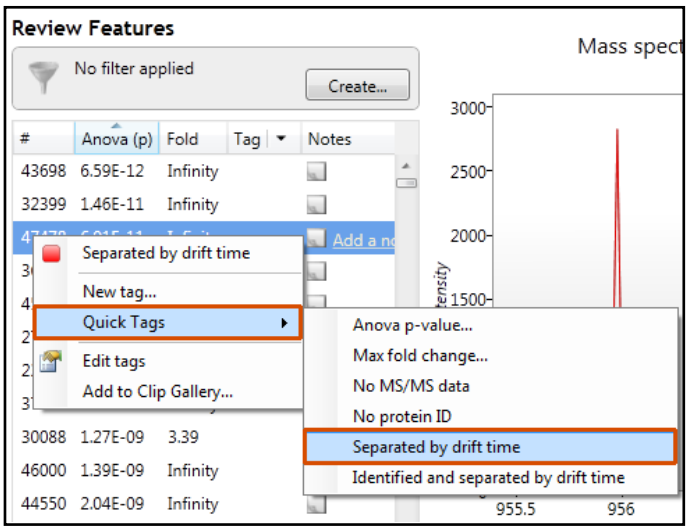


Note: you can use the arrow keys to review the peptide ions while maintaining the current view

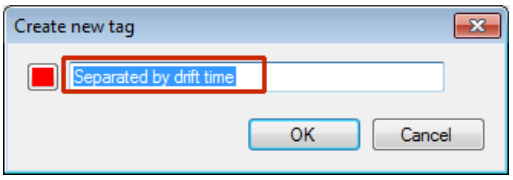
Using Quick Tags to locate examples of Drift Time

In the previous section, describing how to view Drift Time, you will have noticed the presence of a red 'Tag' in the table next to the peptide ion that we examined. Progenesis QI for proteomics allows you to assign tags based on the properties of detected peptide ions either through the manual sorting of the table or making use of the 'Quick Tags'. These tags can be used to filter the list of displayed data in order to aid exploration of the data.

To create a Quick tag for all peptide ions demonstrating separation by Drift time, right click on the table. Select **Quick Tags** then **Separated by drift time.**



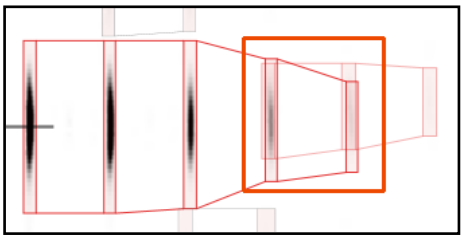
In the new tag dialog either accept or overtype the tag name.

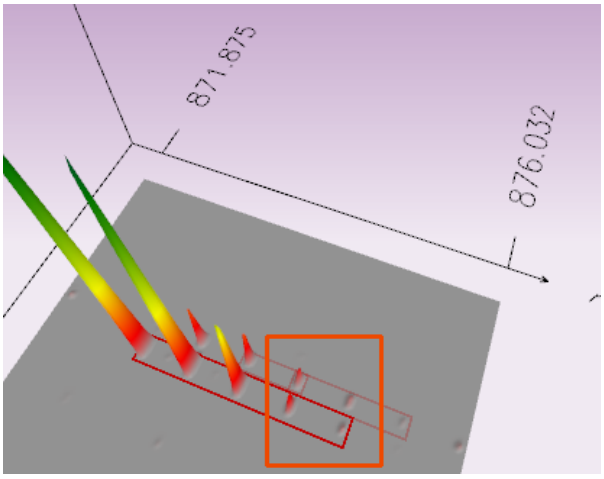


When the tag is created it will appear against those peptide ions that meet the criteria for the creation of the tag, in this case:

It tags peptide ions that overlap in both m/z and retention time but do not show an overlap in the drift time dimension i.e. those peptide ions that drift time has separated

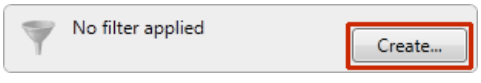
For example the peptide ions below is overlapping at the same m/z and RT but are **separated in drift time**





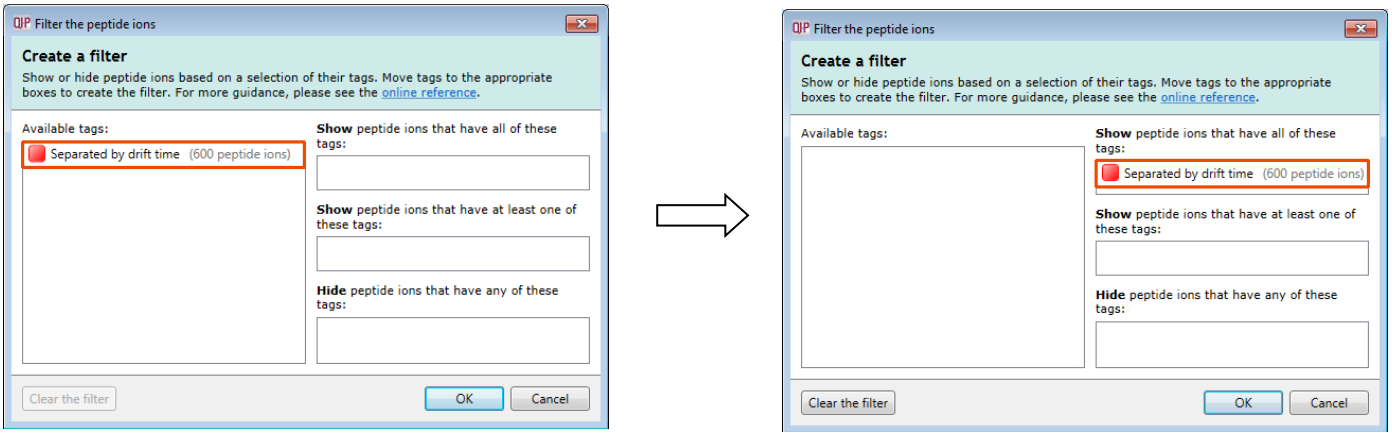
Now filter the table so that it currently only displays a list of peptide ions containing the **separated by drift time** tag.

Click on **Create** on the filter panel above the table.

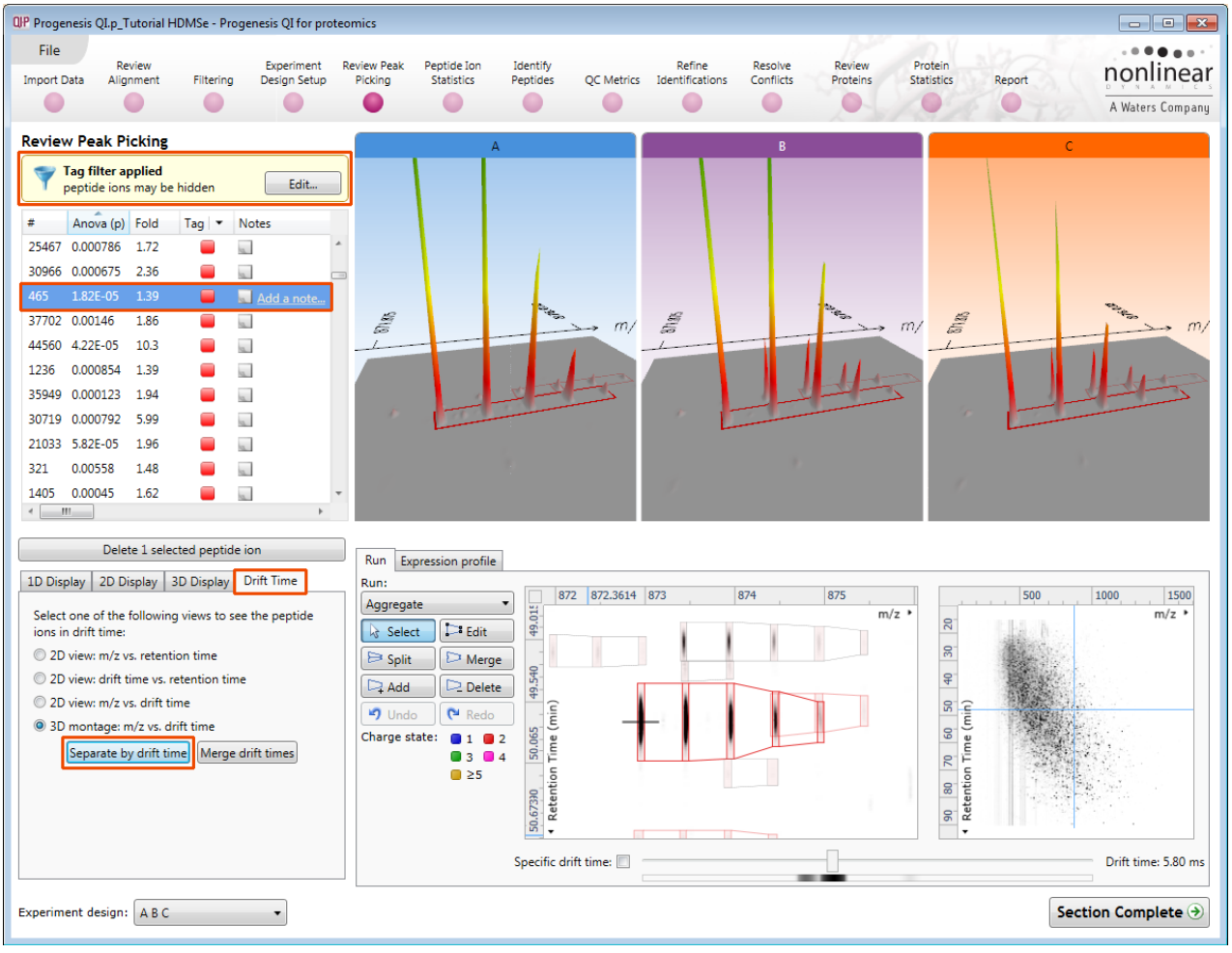


This will open a Tag filter dialog, in this example, displaying that you have created assigned the **Separated by drift time** tag to 600 peptide ions in your experiment.

To display only those peptide ions containing this tag drag the **Separated by drift time** tag on to the **Show** panel and click OK.



When you apply the tag filter the table will now only display the peptide ions with the appropriate tag(s).



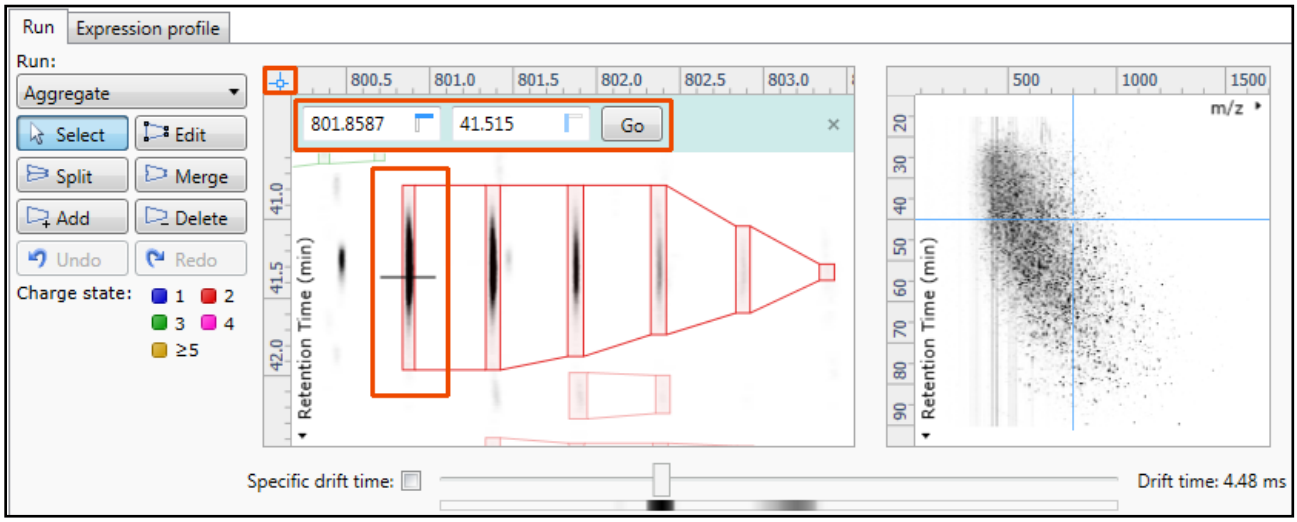
Note: with this **Tag filter applied** you can easily review the effect of Drift time separation for the peptide ions.

To remove the filter click on **Edit**, above the table, and **Clear the filter** followed by **OK**.

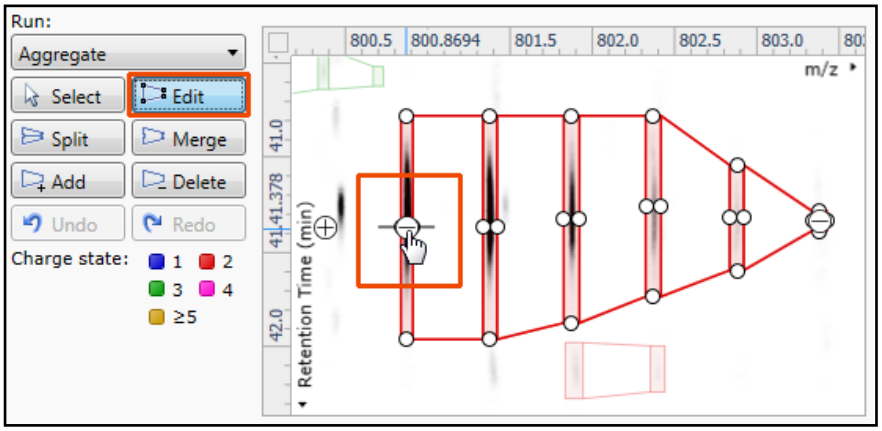
Editing of peptide ions in the View Results stage

As an example of using the editing tools which are located on the left of the LC-MS Run view, we will remove and add back the 'monoisotopic peak' for the detected peptide ion selected below. A peptide ion can be selected from the 'Peptide ions' list or located using the various views.

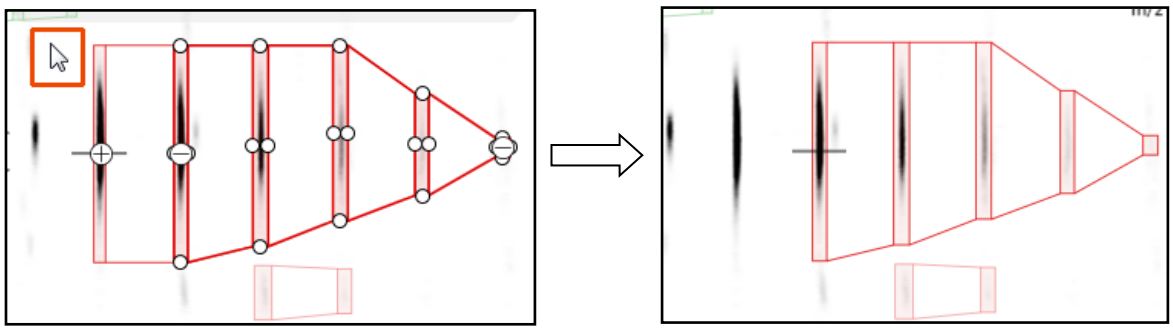
1. Locate the peptide ion at approx 801.86 m/z and 41.5 min using the **Go To Location** tool (top left of zoomed ion map), right click and zoom out and click the cross hairs on the monoisotopic peak to set the zoom.



2. Select the **Edit** tool and click on the peptide ion (in the Run view) to reveal the 'edit handles'

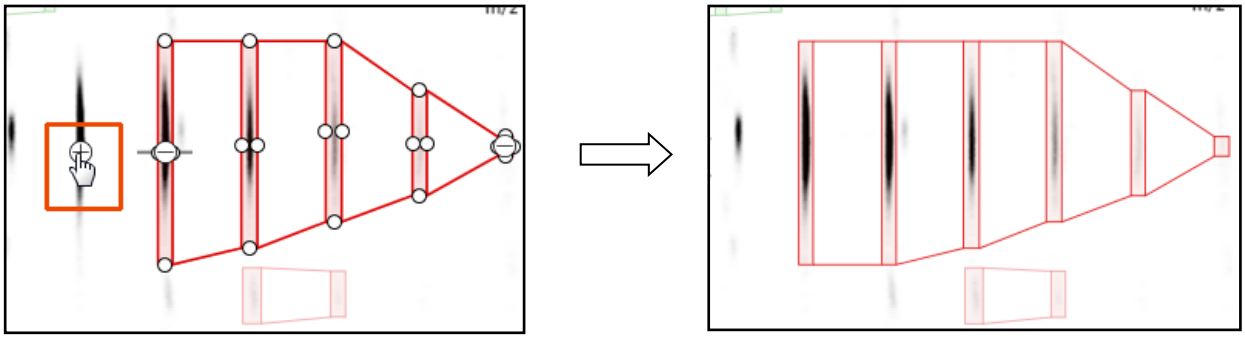


3. Click on the 'minus' handle over the monoisotopic peak to remove it.



4. Click outside the boundary of the peptide ion to update the view.
5. To add a peak to an existing peptide ion, ensure that **Edit** is selected then click inside the peptide ion to reveal the handles.

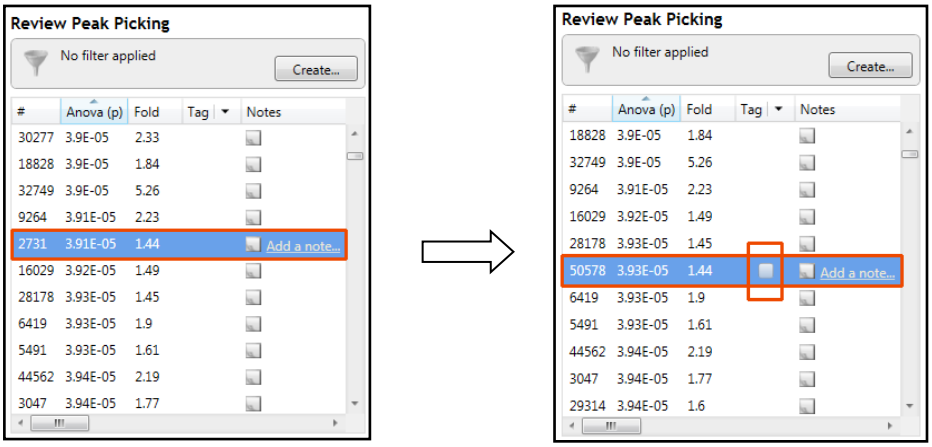
6. Click on the ‘plus’ handle on the peak to add it.



7. Then click outside the peptide ion to update the view.

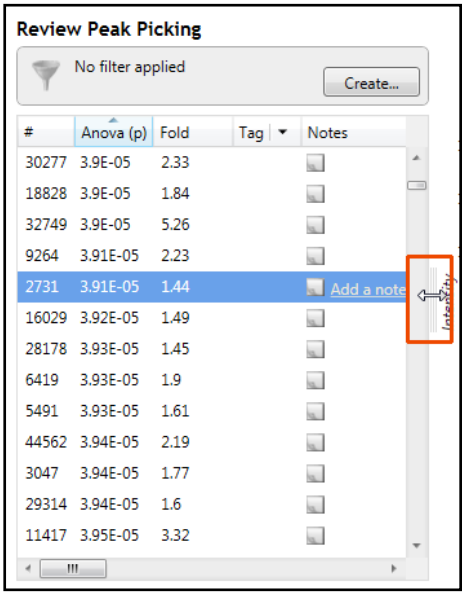
8. **Note**: If you are not satisfied with the editing use the **Undo** button and retry.

9. **Note**: that a tag is automatically added to the edited peptide ion in the table and the peptide ions id. number is changed to the next available one at the end of the list.



The other tools: **split, merge, add and delete** behave in a similar fashion and their use can be combined to achieve the desired results.

Selecting and tagging peptide ions for Peptide Ion Statistics

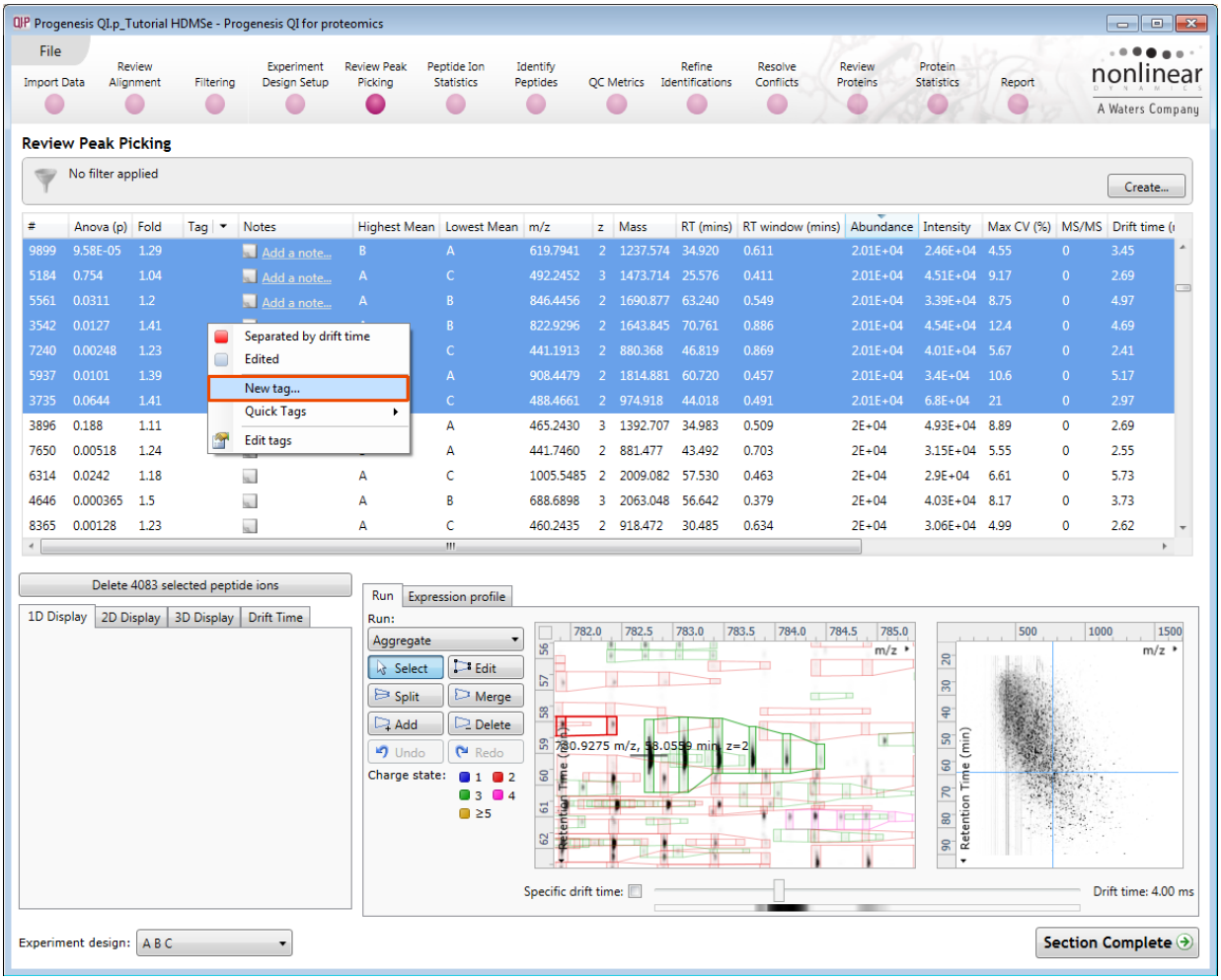


There are a number of ways to ‘refine’ your ‘Ranked List’ of analysed peptide ions before examining them with the Statistical tools in **Peptide Ion Statistics**. These make use of simple ‘Selection’ and ‘Tagging’ tools that can be applied to the various groupings created in Stage 6 (page 28). An example is described below.

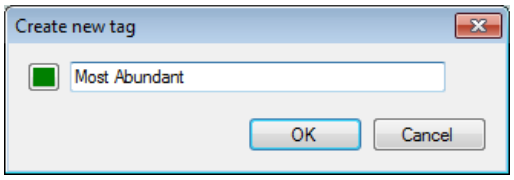
First expand the 'Peptide ions' table to show all the details by double clicking on the ‘Splitter Control’ to the right of the Review Peptide ions table.

Then order on Abundance and select all peptide ions with an Abundance > $2x10^4$.

Right click on the highlighted peptide ions and select 'New Tag'.

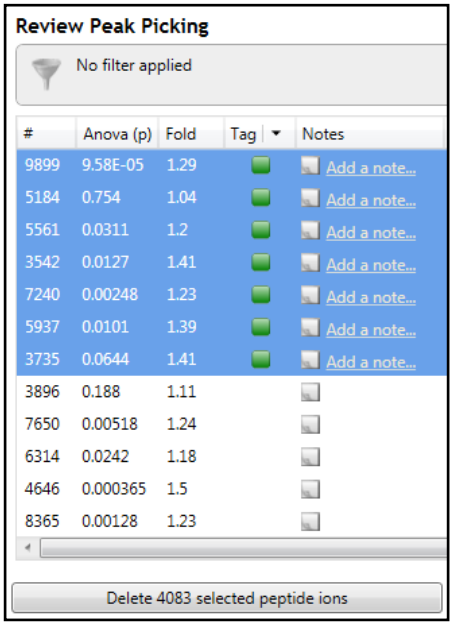


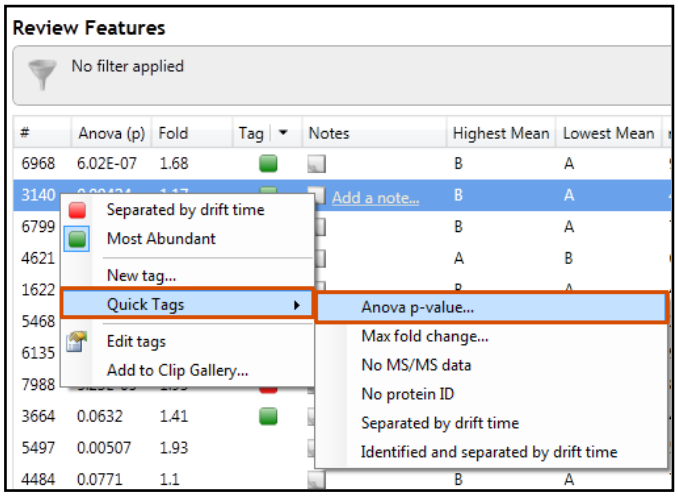
Give the Tag a name. i.e. **'Most Abundant'**.

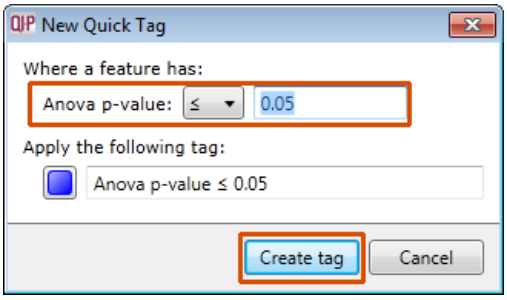


Note: there is already a red tag present that was assigned to those peptide ions that are **Separated by drift time,** which you created in the previous section

On clicking **OK** the Tag is added to the peptide ions highlighted in the table (signified by a coloured square, green in this example).

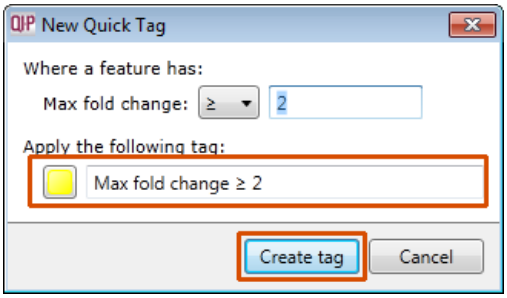


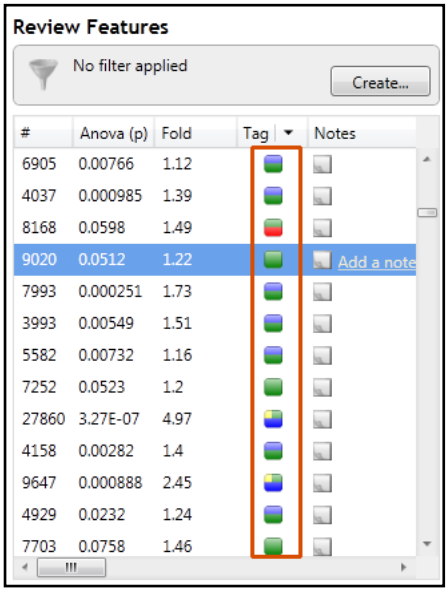




Now right click on any peptide ion in the table and select **Quick Tags** this will offer you a number of standard tag options. Select **Anova p-value....** Then set the threshold as required and adjust default name as required and click **Create Tag**.

Once this tag appears against peptide ions in the table right click on the table again and create another Quick Tag, this time for peptide ions with a **Max fold change ≥ 2**



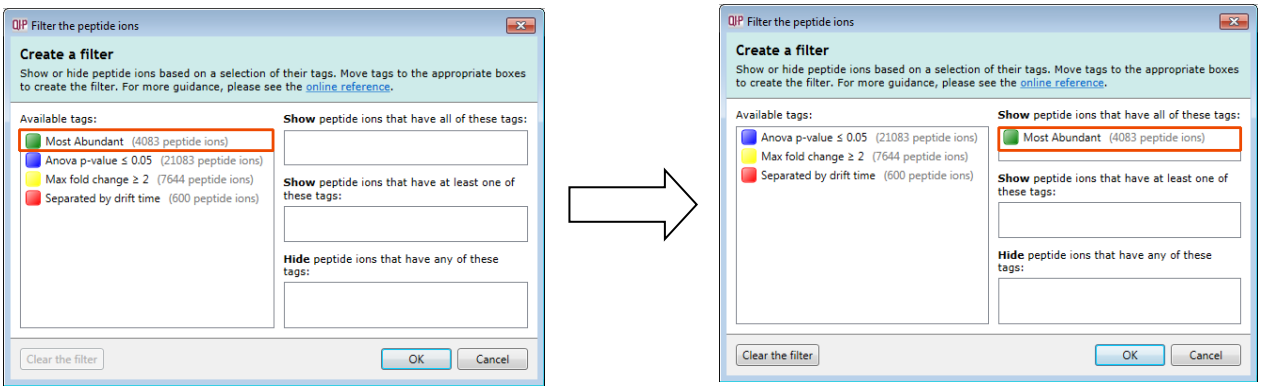


The table now displays peptide ions with multiple tags.

The tags can be used to quickly focus the table on those peptide ions that display similar properties.

For example: to focus the table on displaying those peptide ions that are **Most Abundant** click on **Create** on the filter panel above the table.

Drag the tag on to the panel **Show peptide ions that have all of these tags** and press OK.

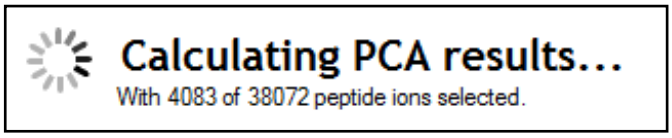


To move to the next stage in the workflow, Peptide Statistics, click **Section Complete.**

Stage 8: Peptide Ion Statistics on Selected Peptide ions

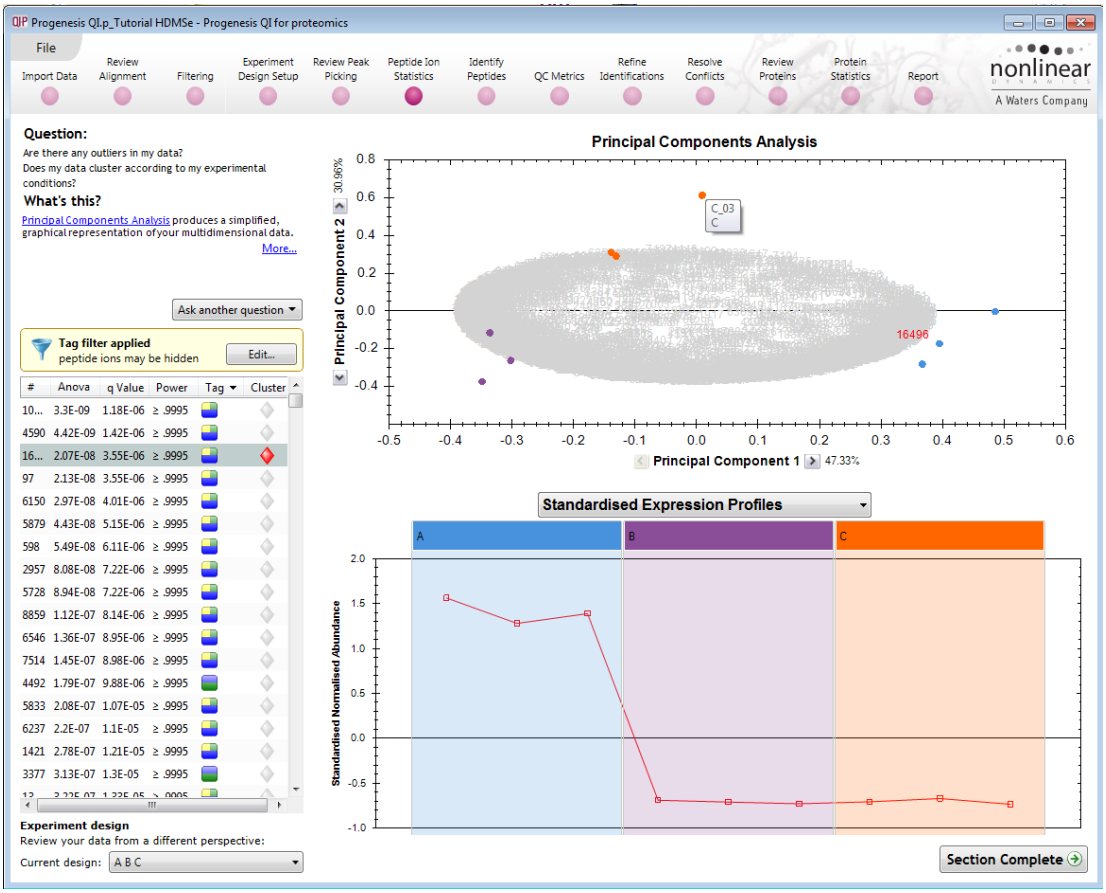
The user guide now describes the functionality of the Multivariate Statistics.

Peptide Statistics opens calculating the Principal Components Analysis (PCA) for the active 'tag' in this case the **Most Abundant** peptide ions.



As an example we will start by examining the behaviour of the **Most Abundant** peptide ions from the previous stage, **Review Peak Picking**.

The statistical analysis of the selected data is presented to you in the form of interactive graphical representation of answers to questions asked of the analysed data.



Note: the LC-MS runs (samples) are displayed as solid coloured circles on the plot. To identify the runs, a tooltip is displayed when the cursor is held over each circle.

Principal Component Analysis (PCA)

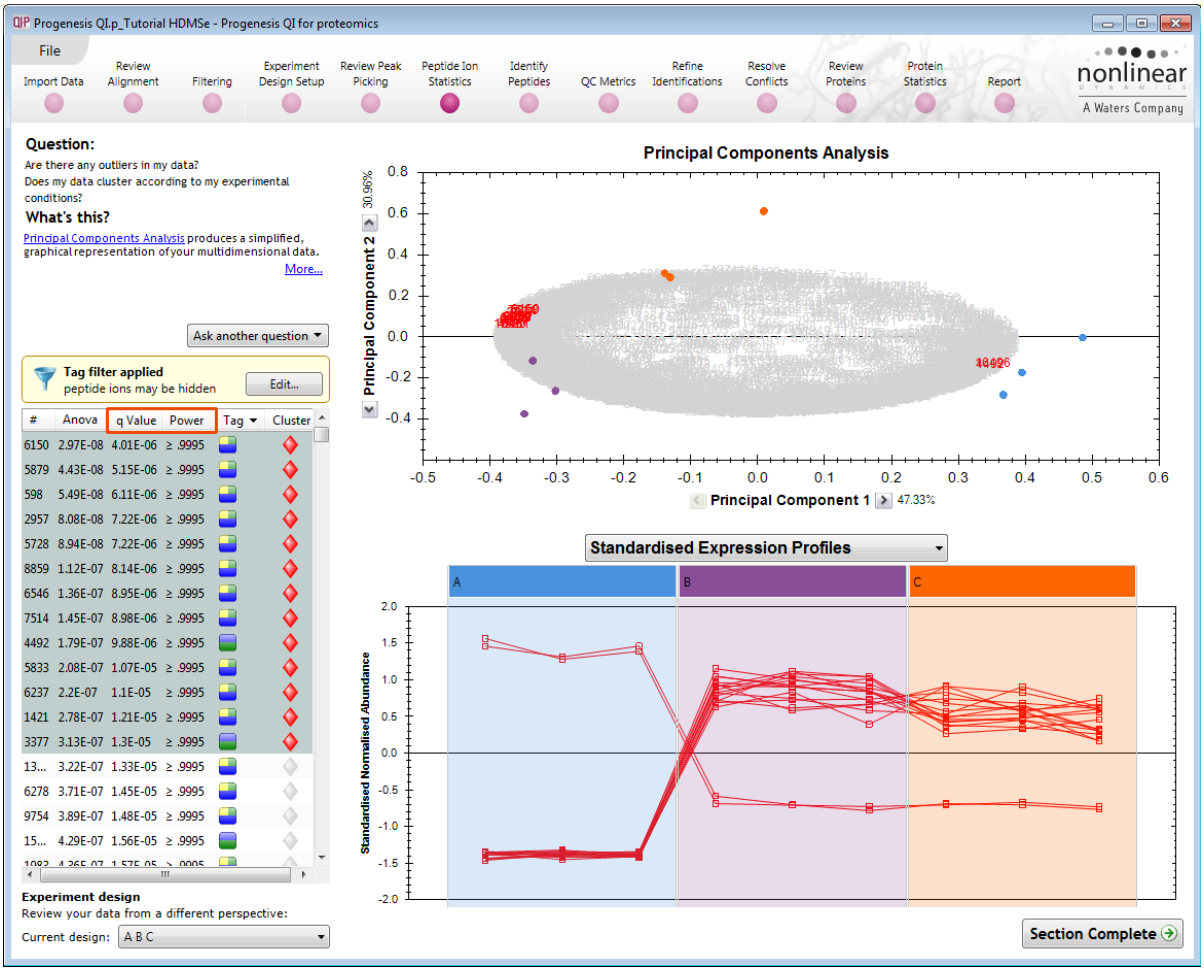
In **Peptide Stats** the first statistically based question asked of the data takes the form of a Quality Control assessment:

Are there any outliers in my data? And does my data cluster according to my experimental conditions?

It answers this question by:

'Using Principal Components Analysis (PCA) to produce a simplified graphical representation of your multidimensional data'.

PCA can be used to determine whether there are any outliers in the data and also look at how well the samples group. The groupings that can be observed on the 2D PCA plot can be compared to your experimental conditions and conclusions can be drawn regarding possible outliers in your data. Selecting peptide ions in the table will highlight the peptide ions on the 'Biplot' and their expression profiles will appear in the lower panel.



Note: the Table in the Stats view contains additional columns:

q value: tells us the expected proportion of false positives if that peptide ion's p-value is chosen as the significance threshold.

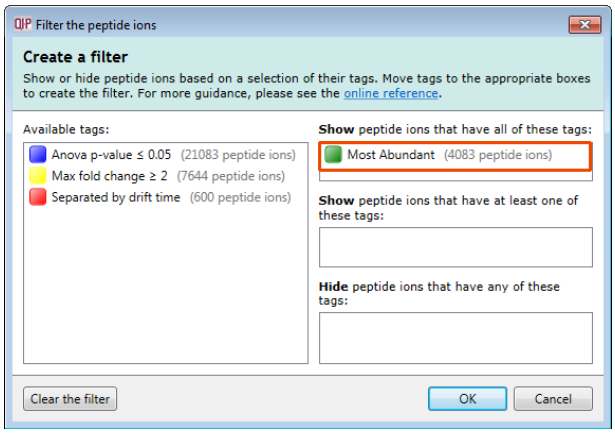
Power: can be defined as the probability of finding a real difference if it exists. 80% or 0.8 is considered an acceptable value for power. The Power Analysis is performed independently for each peptide ion, using the expression variance, sample size and difference between the means.

Also, for a given power of 80% we can determine how many samples are required to ensure we find a difference if it actually exists.

Note: Power analysis is discussed in Appendix 6 (page 90)

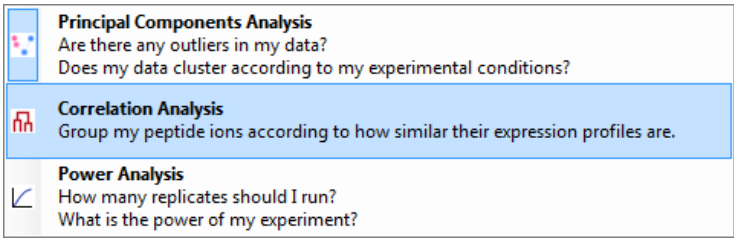
Correlation Analysis

With the tag filter still set to display only the top 4083 **Most Abundant peptide ions**, we are going to explore the Correlation Analysis of these peptide ions.

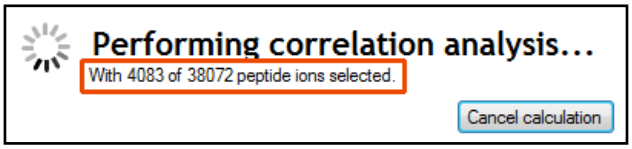


To set up the **Correlation Analysis** using this filtered data set click on **Ask another question** (above the table)

A selection of 3 tools will appear in the form of questions.



Select the second option to explore 'peptide ion correlation based on similarity of expression profiles'



This time the statistically based question(s) being asked is:

'Group my (selected) peptide ions according to how similar their expression profiles are'

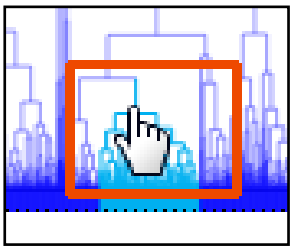
The question is answered by:

'Using Correlation analysis to evaluate the relationships between the (selected) peptide ions' expression profiles'.

The answer is displayed graphically in the form of an interactive dendrogram where the vertical distance, between each peptide ion can be taken as indicative of how similar the expression profiles of each cluster of peptide ions are to each other.

Correlation Analysis enables the grouping of peptide ions together according to how similar their expression profiles are.

For example: To highlight groups of peptide ions demonstrating **different expression profiles** click on a 'node' for a branch of the Dendrogram (as shown above) while holding the mouse button, hold down the **Ctrl** key and then click on another node as shown.



If you have selected 2 nodes then there will be two expression profile graphs

Note: by highlighting a group of peptide ions with similar expression the peptide ions are identified with the same colour of cluster flag in the table. This allows the table to be sorted on cluster and tagged accordingly

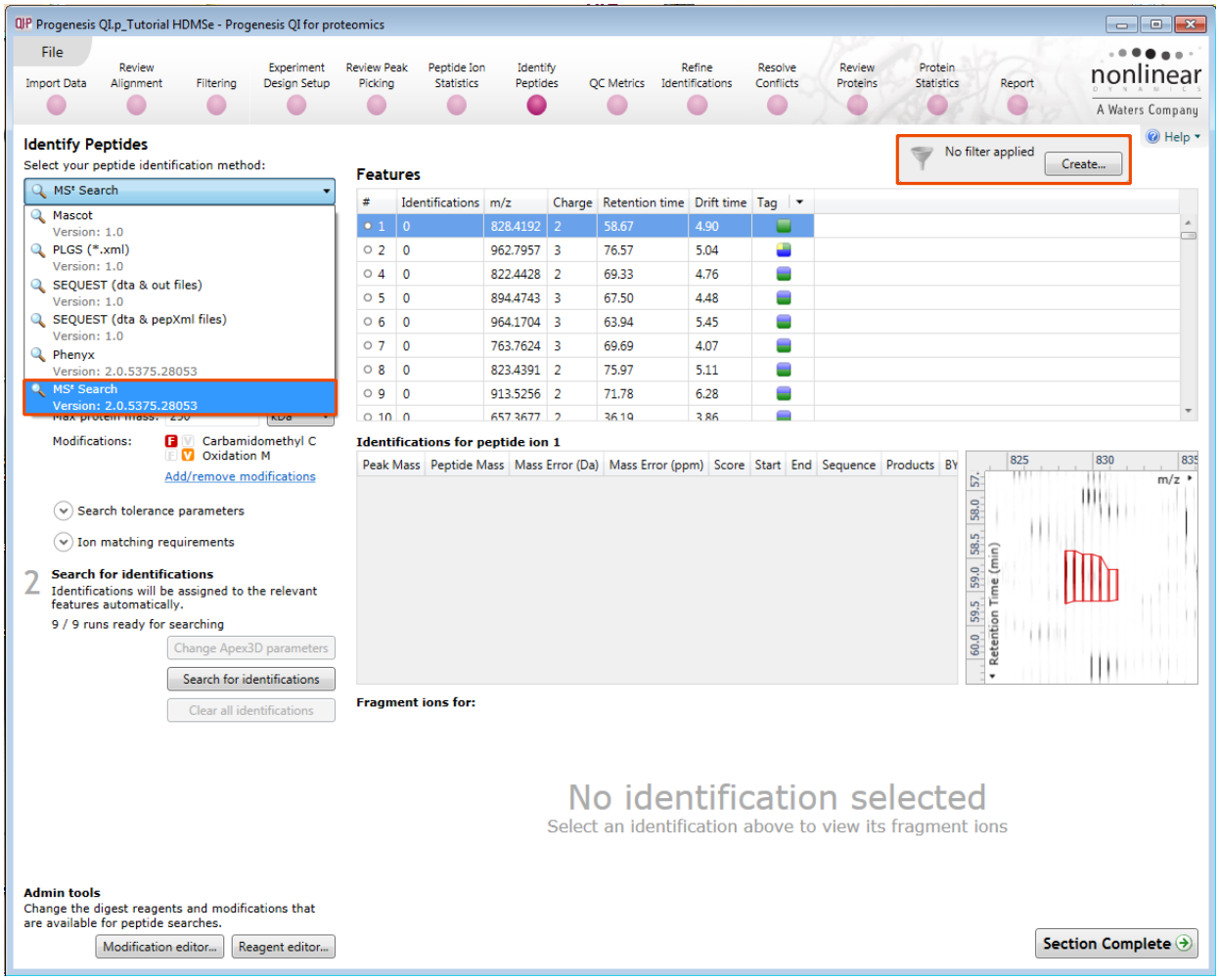
Tip: when reviewing the tags (see above) if you are not applying a new filter then use the **Cancel** button to return to the main view, this prevents unnecessary recalculation of your data.

Before moving to the Identify Peptides stage in the Workflow, first return to the PCA display and clear all tag filters **Clear all Tag filters**.

To move to the next stage in the workflow, **Identify Peptides**, click **Section Complete.**

Stage 9: Identify peptides

Progenesis QI for proteomics is designed to perform peptide identifications either directly or by allowing you to export MS/MS peak lists in formats which can be used to perform peptide searches by various search engines. The resulting identifications can then be imported back into Progenesis QI for proteomics, using a number of different file types, and matched to your detected peptide ions.



For this example we are using the direct method **MS^E Search**.

Note: Following the full automatic processing, described in Stage 2 of this guide, the Identify Peptides page currently displays the full list of the detected peptide ions in your experiment and some of their attributes, including the number of **Identifications** (as shown on page 51)**.** If search results exist these can be cleared by clicking **Clear all identifications,** this will allow you to re-perform the search**.**

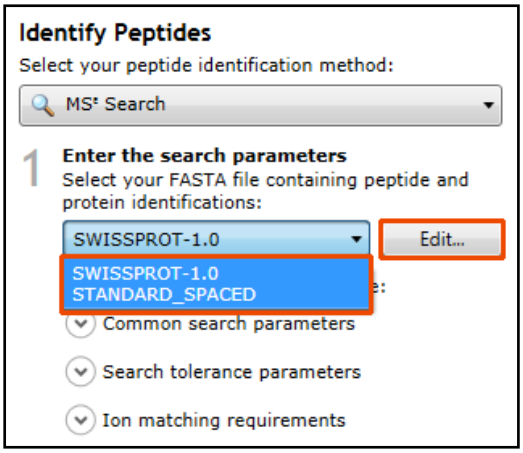
Entering Search Parameters

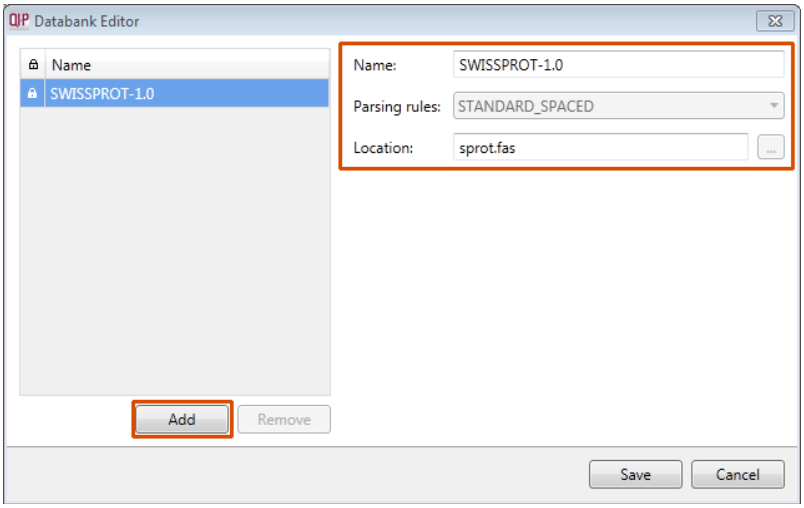
Firstly you need to select the FASTA file containing peptide and protein identifications.

SWISSPROT-1 is provided with the installation of the software.

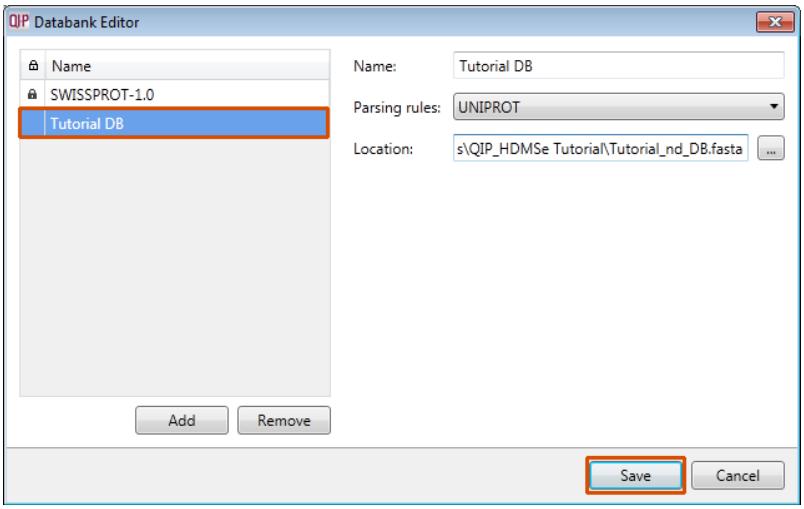
To add new Databanks in the form of FASTA files click on **Edit…** to open the Databank editor

Note: the SWISSPROT-1.0 is locked



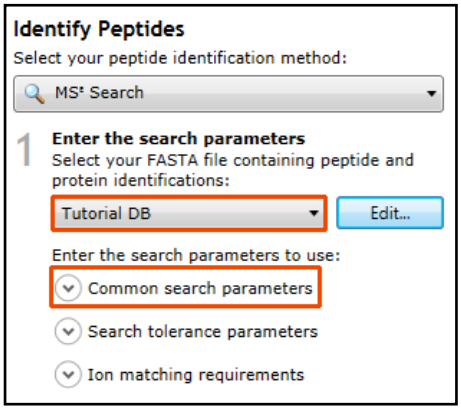


For a new Databank you need to give it name, select the parsing rules and specify the location of the FASTA file, see the example below.



The new Data bank will appear in the left panel now click **Save** to return to the Search parameters.

If your databank is not already displayed then select it from the drop down list.



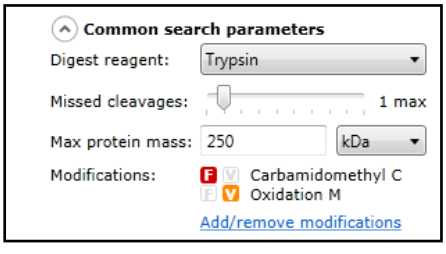
Expand the **Common search parameters**

The default settings are displayed:

Digest reagent: is set as Trypsin**.** Alternative Digest reagents are available from the list and additional ones can be added to the list using the **Reagent editor...**

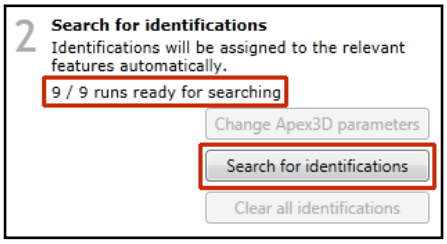
Missed cleavages: is set as 1.

Maximum protein mass: is set at 250kDa

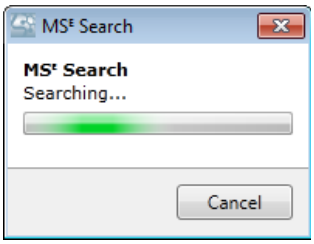


Modifications: are set Carbamidomethyl C (Fixed) and Oxidation M (variable). More modifications are available from the list and additional ones can be added to the list using the **Modification editor…**

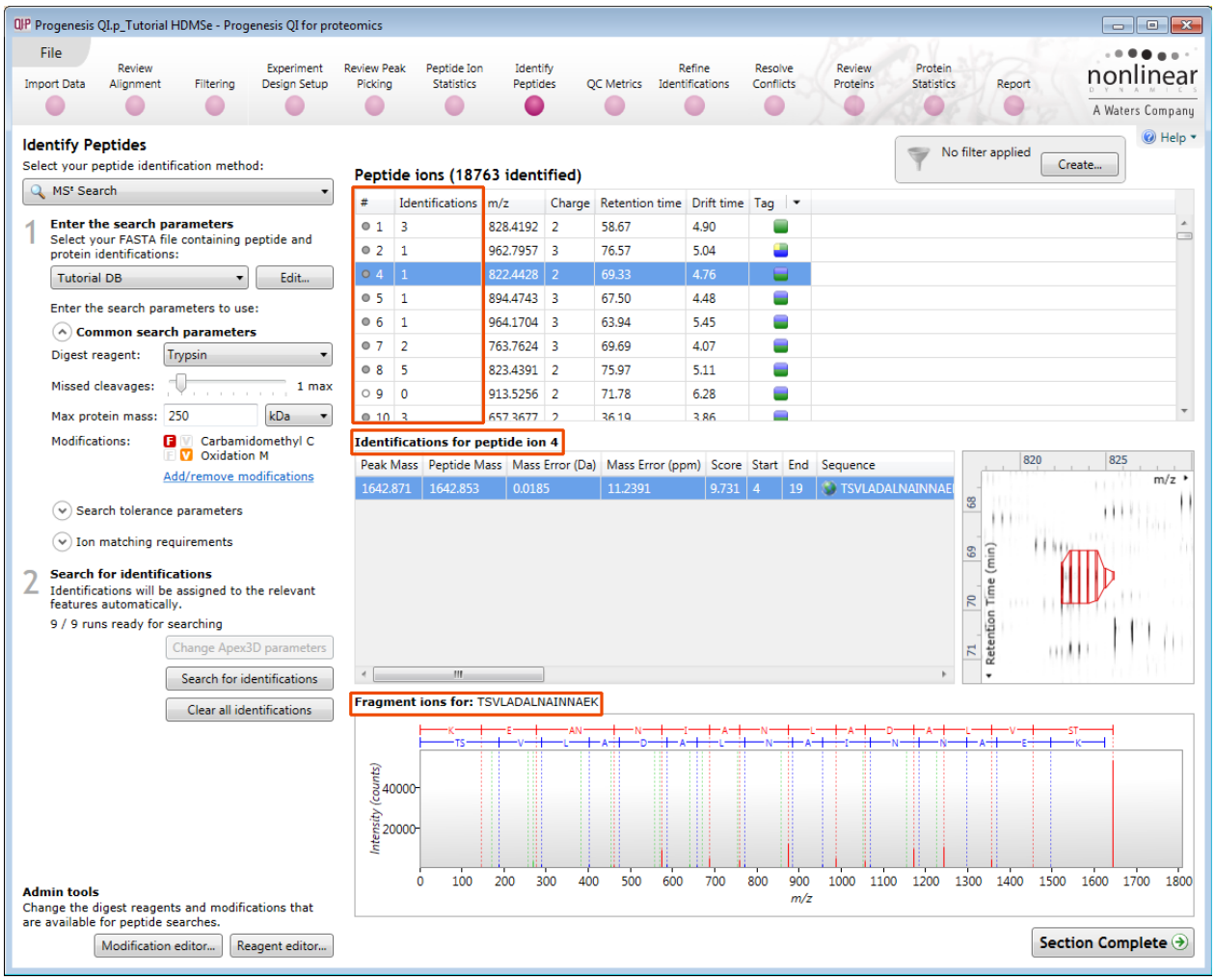
Having selected the Databank and set the parameters, before searching for identifications make sure that all of the runs are available for searching.



Depending on the search parameters the MS^E Search can take some time



Once the Ion Accounting is complete, peptide ions with identifications are identified with a solid grey symbol and the number of identifications appears in the next column.



Details for the current peptide ion identifications are displayed in the table below and the Fragment ions for the current identification are displayed in the bottom panel.

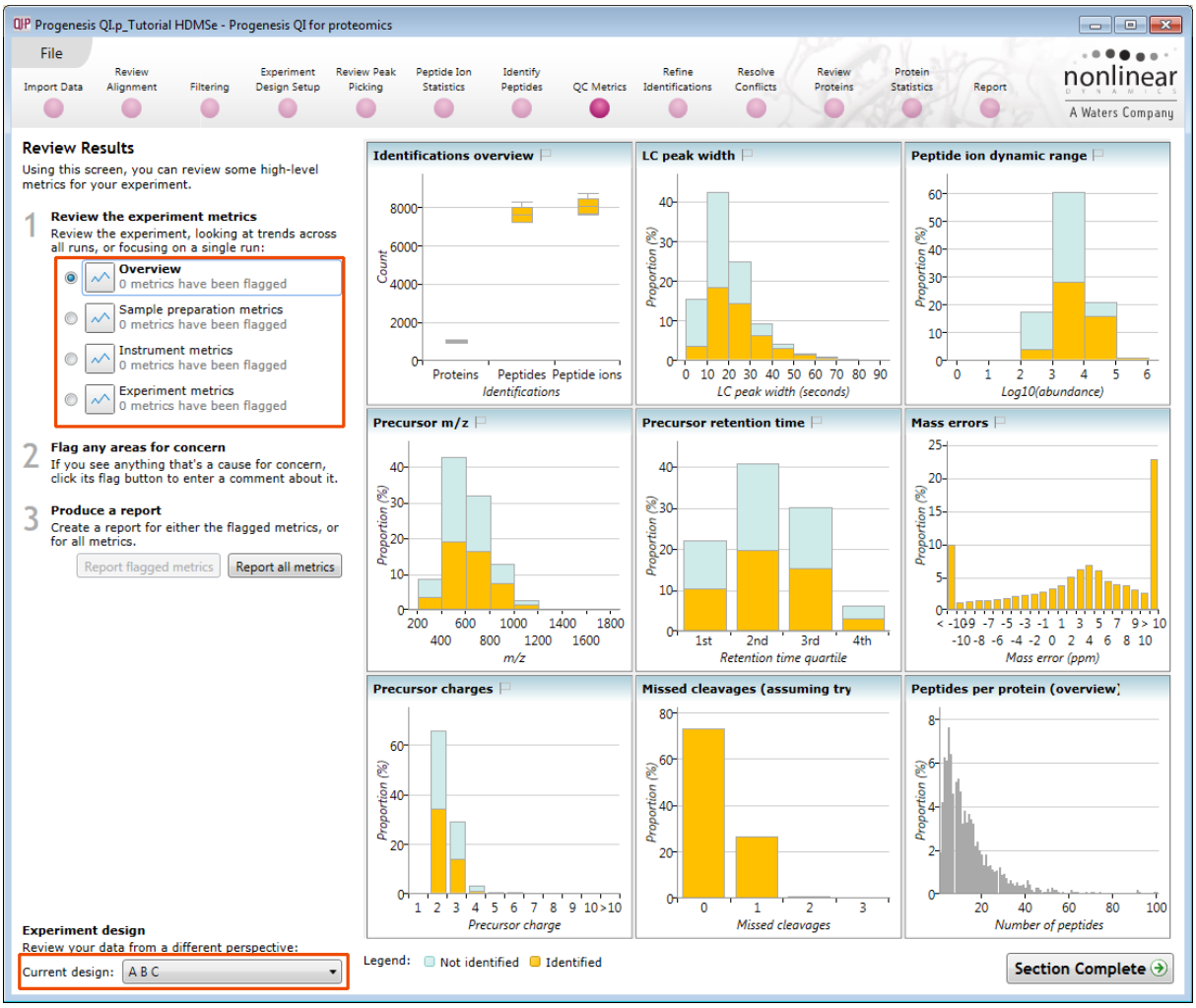
Note: if you want to perform the search with a new set of parameters then first select **Clear all identifications**

Having performed the process of peptide identification you can review the overall quality of your analysis by using the various batch-based quality metrics in the next section, by clicking on QC Metrics on the workflow.

Stage 10: QC Metrics

Progenesis QI for proteomics includes a number of batch-based QC metrics which you can examine to increase confidence in, or identify issues with, your data. These views are presented at the “QC Metrics” page. By setting up experimental groupings that correspond to your batch metadata, you are able to investigate whether there are any systematic effects associated with your processing. This gives you confidence that your conclusions are not affected by technical biases within the course of the experiment.

The charts are updated with changes in your active experimental design, allowing you to examine the effects of all stages of your processing if you wish through using different batch divisions.



Metrics

There are two sets of metrics. Firstly, the overview page provides nine at-a-glance summary measurements which cover the experiment as a whole. Then, the remainder of the pages provide more detailed information and bring in the batch-by-batch detail, some of which builds on the overview charts:

- Sample preparation metrics – highlighting issues or problems with the preparation of your samples:
  Missed cleavages, modifications and abundance dynamic range
- Instrument metrics – highlighting whether your chromatography column and mass spectrometer are configured and performing correctly:
  Mass accuracy, abundance dynamic range (again), precursor charges, MS1 scan rates
- Experiment metrics – concerning the identified proteins and peptides in your experiment, allowing you to pick out any outlying runs or conditions:
  Proteins, peptides, peptides per protein, % of peptide ions identified, proteins per condition

More detailed information on the QC metrics is available on the FAQ pages.

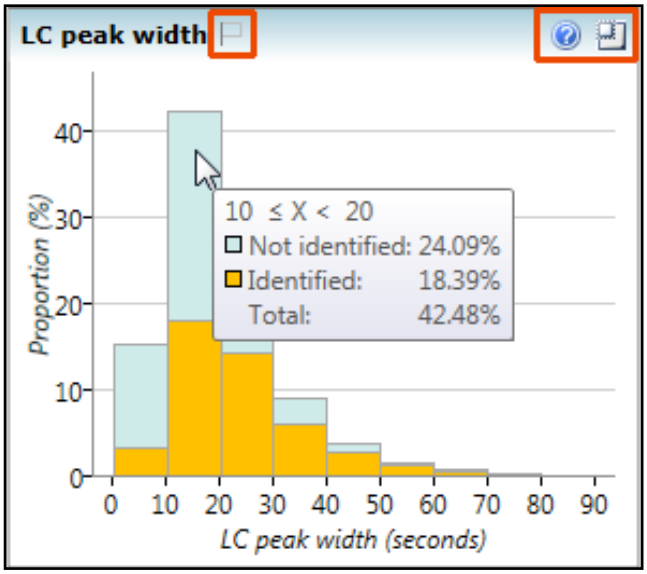
Interpretation and use

It's worth noting that there is no simple 'right' or 'wrong' answer as to whether your data are of high enough quality.

Note: QC measurements are designed to flag up potential issues for investigation, hence there must be an assignable cause to the variation observed to render QC practical.

Furthermore, the QC metrics will be rebuilt if you delete and re-do any stages leading up to them (for example, clearing identifications and re-searching with changed identification parameters, or altering the alignment). Hence, if you archive your experiment before and after your changes, or export the charts before and after, you can see the effects of your changes upon the quality measurements – a very useful method for assessing your interventions.

Note: that the overview metrics can each be expanded by clicking on the top-right icon in the sub-window. Additionally, hovering over a column will bring up a tooltip containing quantitative information on the results.

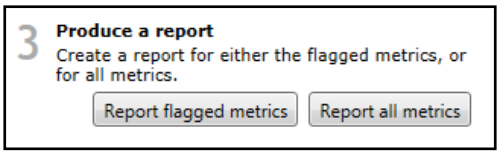


You can also tag metrics of interest or concern with a comment for your records, which is saved with the experiment. To do this, click the empty flag icon in the tab header for the given metric (or in the tile header on the overview screen):

This will pop up a dialog allowing you to enter a comment describing why this metric has been flagged.

Reporting

You can export a report for your own records, or to pass on to another member of your team:



Report all metrics:

Generates a report containing all metrics shown in the application (overview, sample preparation metrics, instrument metrics and experiment metrics). This may be useful for documentation purposes, or to verify the quality of your experiment.

Report flagged metrics:

Generates a report of only metrics you have flagged, along with the message you provided. This may be useful for giving to a technician or other team member, to highlight areas of the experiment that need improvement.

In order to review, and refine the quality of the **Peptide Search** results click on the next stage in the workflow, **Refine Identifications**.

Stage 10: Refine Identifications

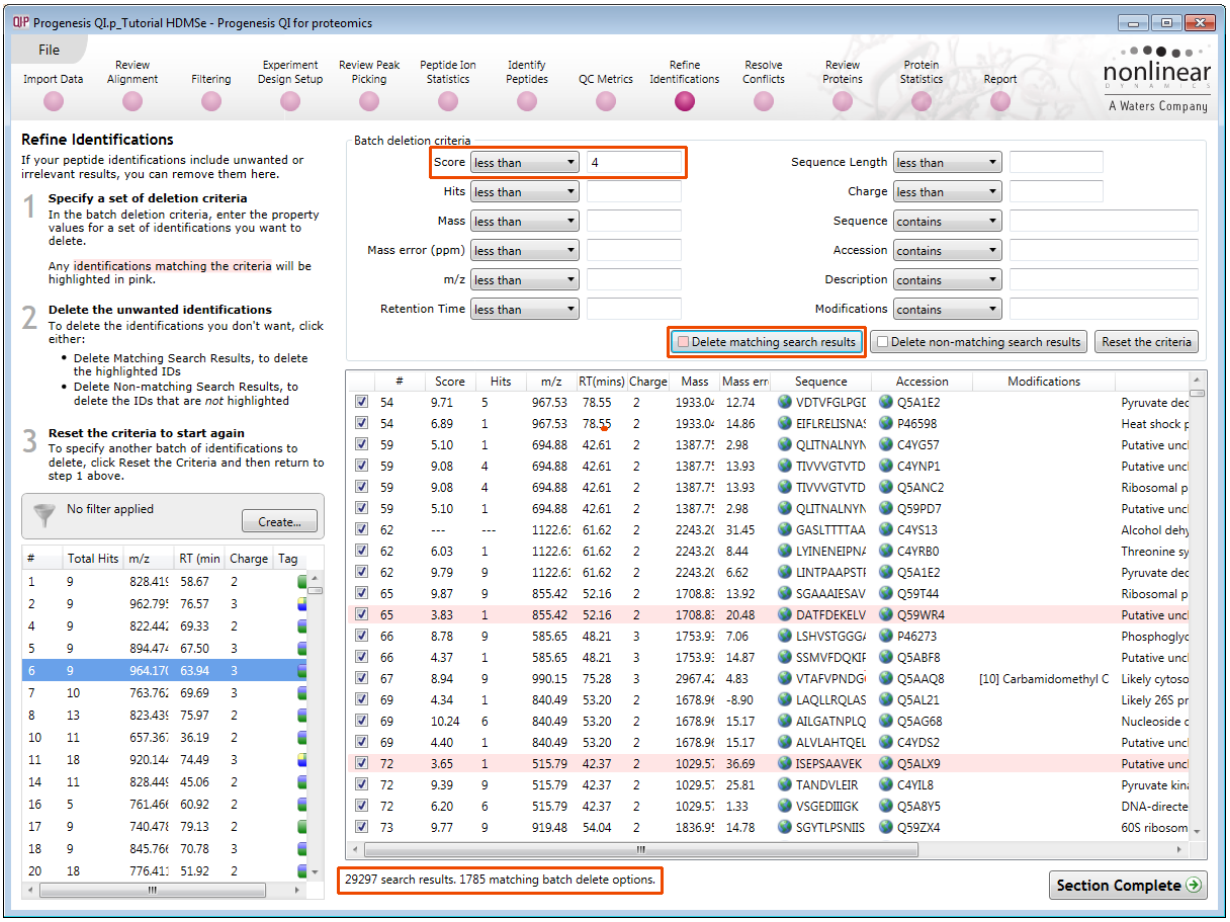
In this example we are going to apply a number of filters to 'refine' the quality of the Databank search.

Note: before removing any identifications, make sure there are **no** tag filters applied at the Identify peptides stage.

As an **example** the following section describes how sequential filtering of the Peptide results can be performed using the following thresholds described below:

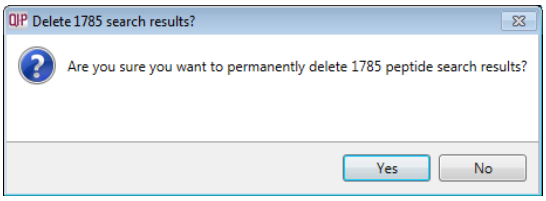
- Remove identifications with a Score less than 4
- Remove identifications where less than 2 hits were returned
- Remove all identifications where the Protein Description **Contains** the following: 'Putative', 'Probable', 'Like', 'Potential' and 'Predicted'

To perform these filters, on the Batch detection options panel, set the Score to less than 4, then **Delete matching search results**.

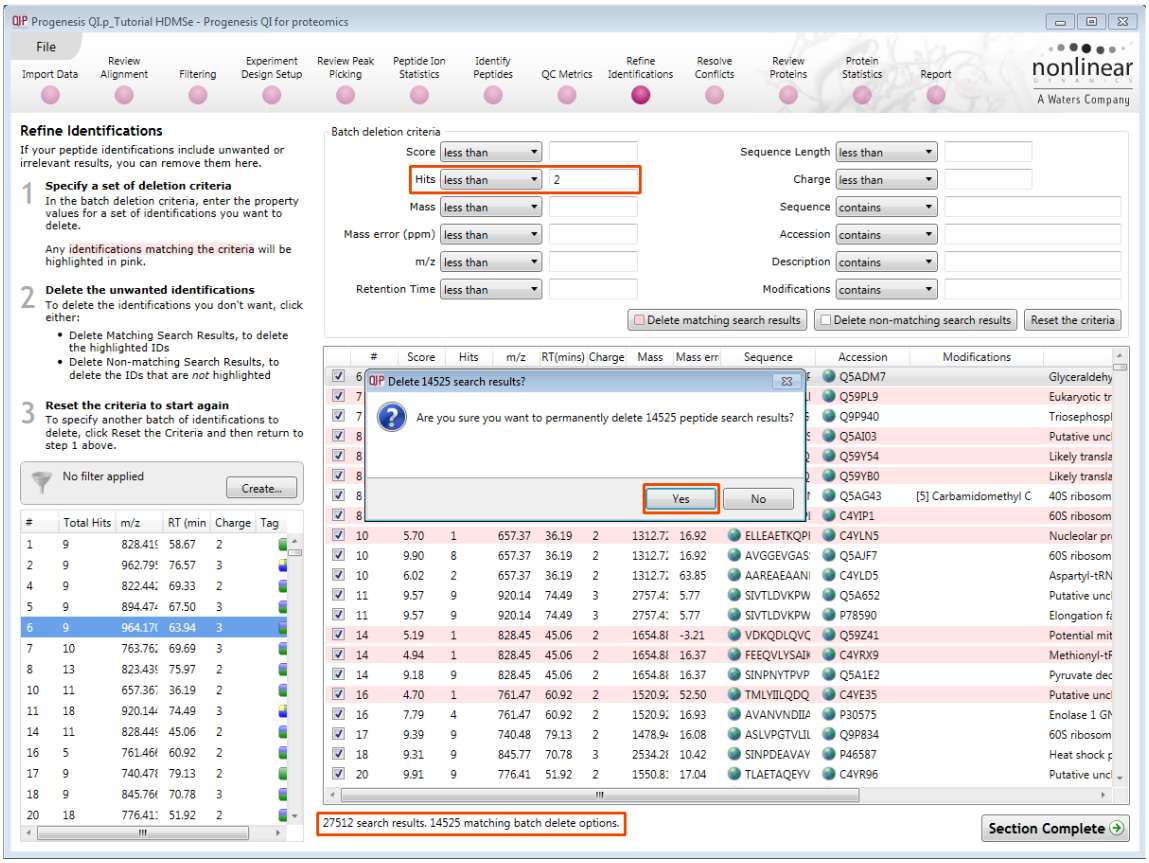


Note: the search results matching the filter criteria turn pink and the total is displayed at the bottom of the table (1785 matching out of 29297)

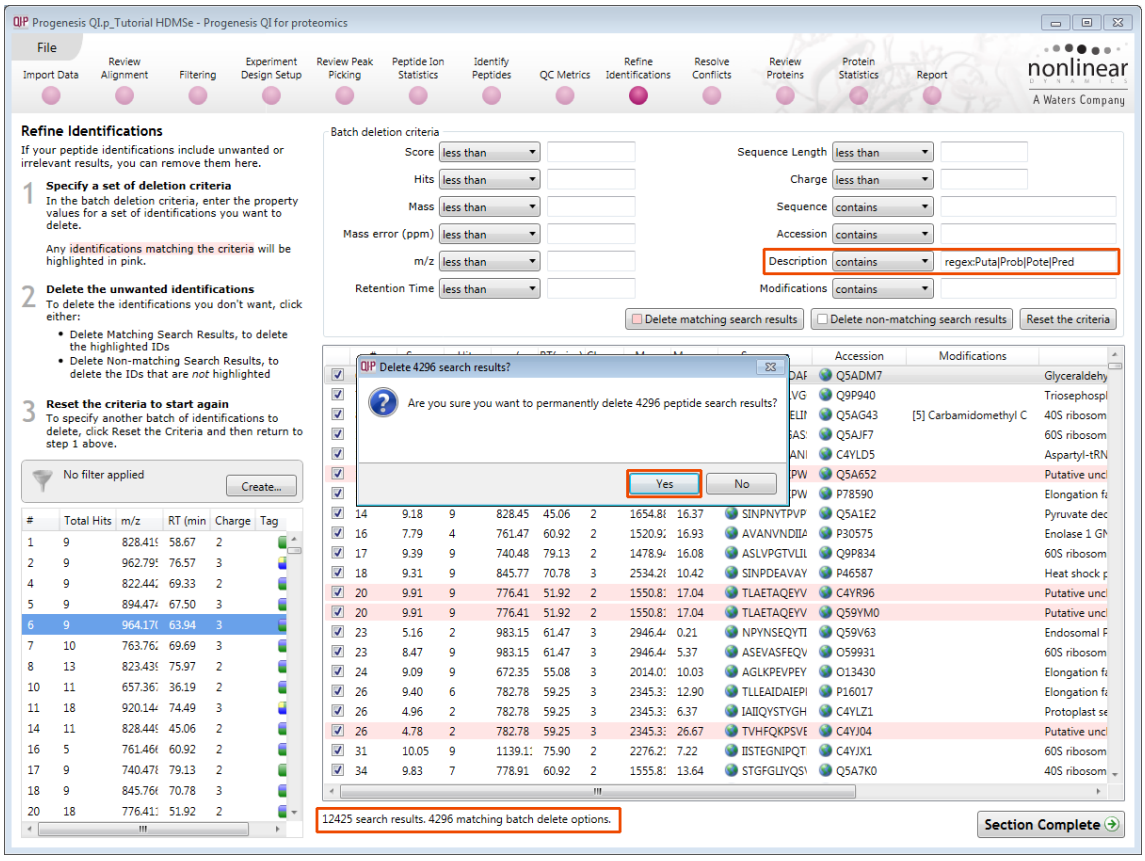
Note: a dialog warns you of what you are about to delete



Now click **Reset the criteria** to clear the previous threshold then apply the next filter 'Hits: less than 2'.



Now in the Description first enter '**Like**' and delete matching search results. Then enter the 'regular expression': **regex: Puta|Prob|Pote|Pred** and delete matching search results.



Having applied all the filters there will be **8129** search results remaining

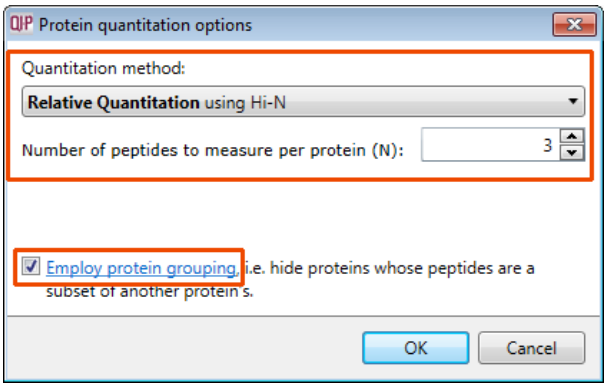
To validate the Peptide search results at the protein level select the next stage in the workflow by clicking on **Resolve Conflicts**.

Stage 11: Resolve Conflicts

This stage allows you to examine the behaviour of the identified peptides and choose to resolve any conflicts for the various peptide assignments at the protein level.

The number of conflicts you have to resolve will depend on the scope and stringency of the filters you apply at the **Refine Identifications** stage.

Note: the default **Protein options** for protein grouping and Protein quantitation are set as shown



This means that if you choose **not** to resolve the conflicts then proteins, to be considered for quantitation, require at least one unique peptide (number in brackets after peptide count).

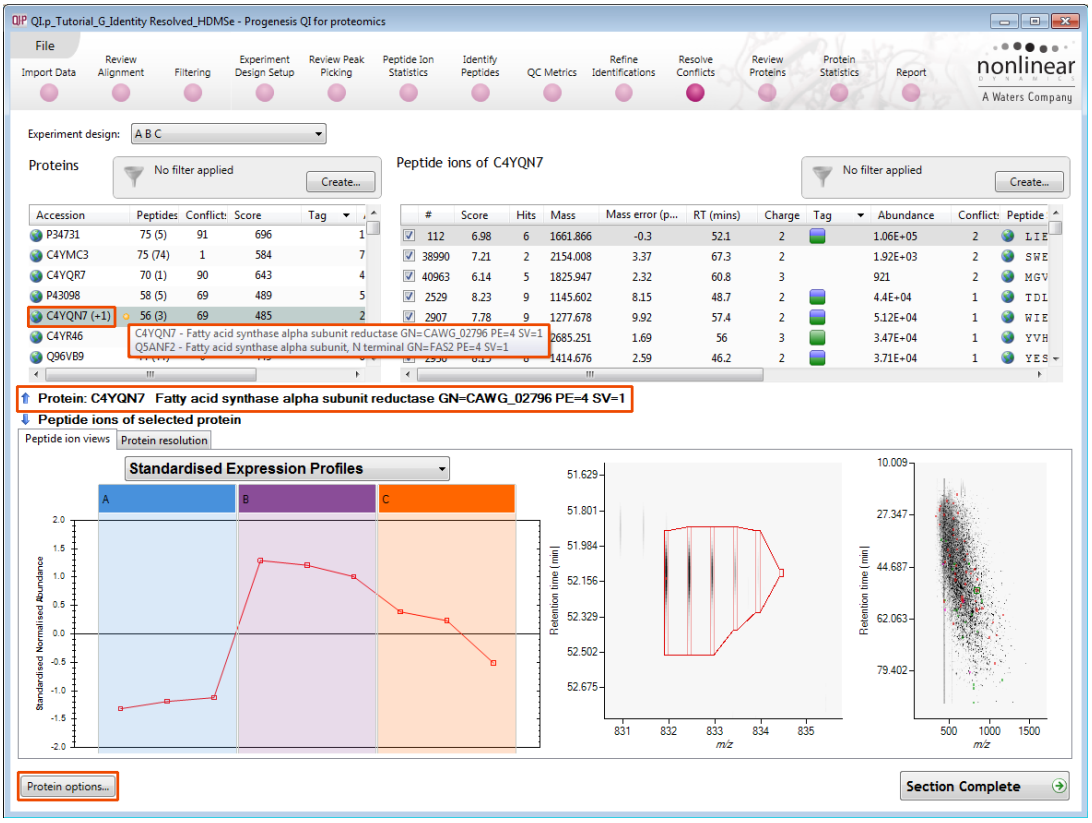
For more details on Protein Grouping page 60 and Protein Quantitation options go to page 62.

If you choose **not** to resolve conflicts then click **Section Complete** to go to **Review Proteins**.

To proceed with resolving conflicts there are some simple rules that you can apply using this stage in the workflow.

With **Group similar proteins** selected the additional members are indicated by a bracketed number located after the Accession number. As an example, when the cursor is held over the accession number the group members appear in a tool tip.

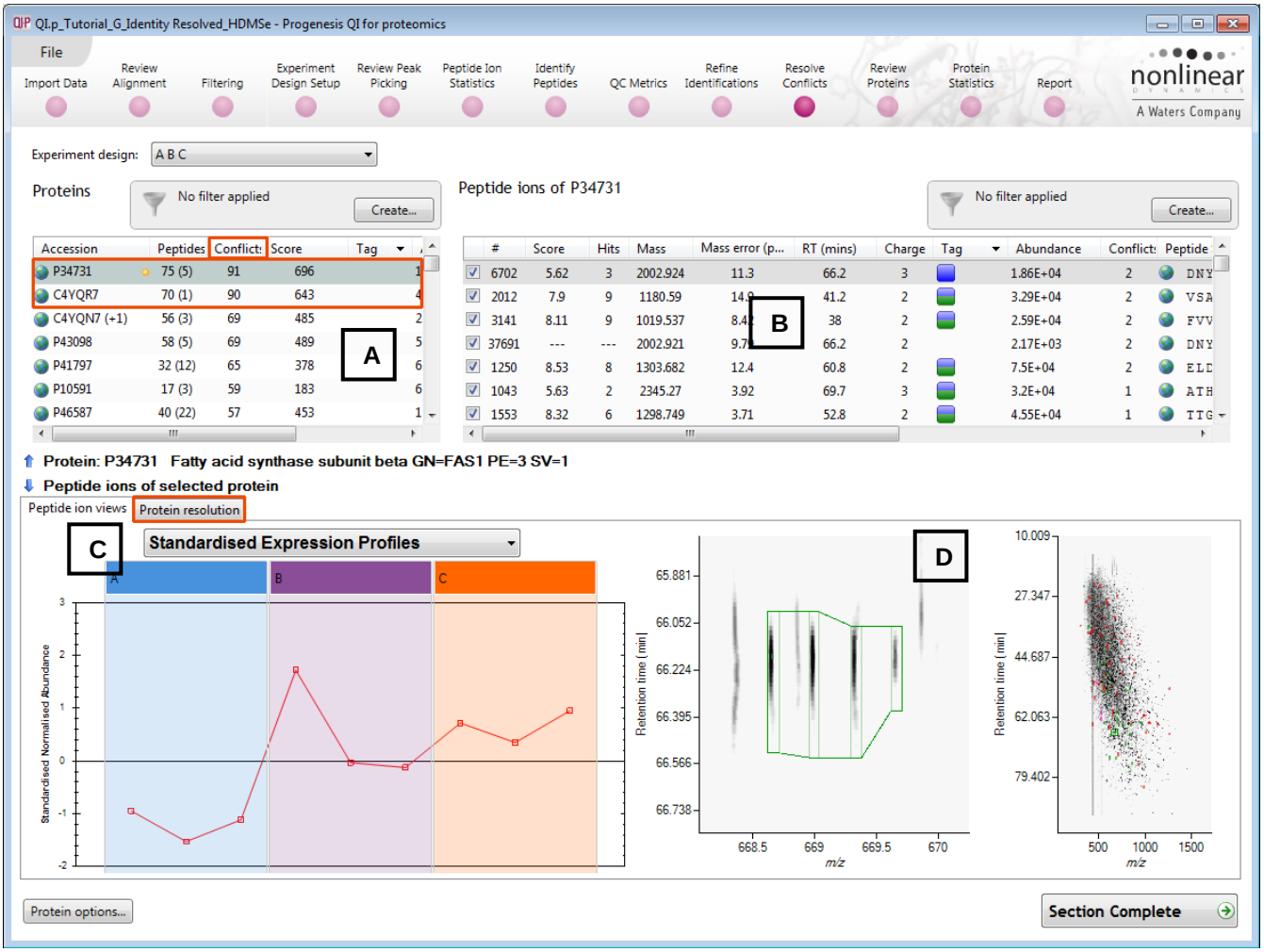
The bracketed number in the **Peptides** column indicates the number of unique peptides used for quantitation.



The **Resolve Conflicts** stage provides a number of interrelated graphical and tabular views to assist you in the validation of the peptides that have been assigned to proteins and also to review the relevance of the data returned from the search.

Open **Resolve Conflicts** and order the data in the Proteins table A on the basis of **Conflicts**.

Note: the look of the tables (with regards to ordering) in the following section may vary slightly.

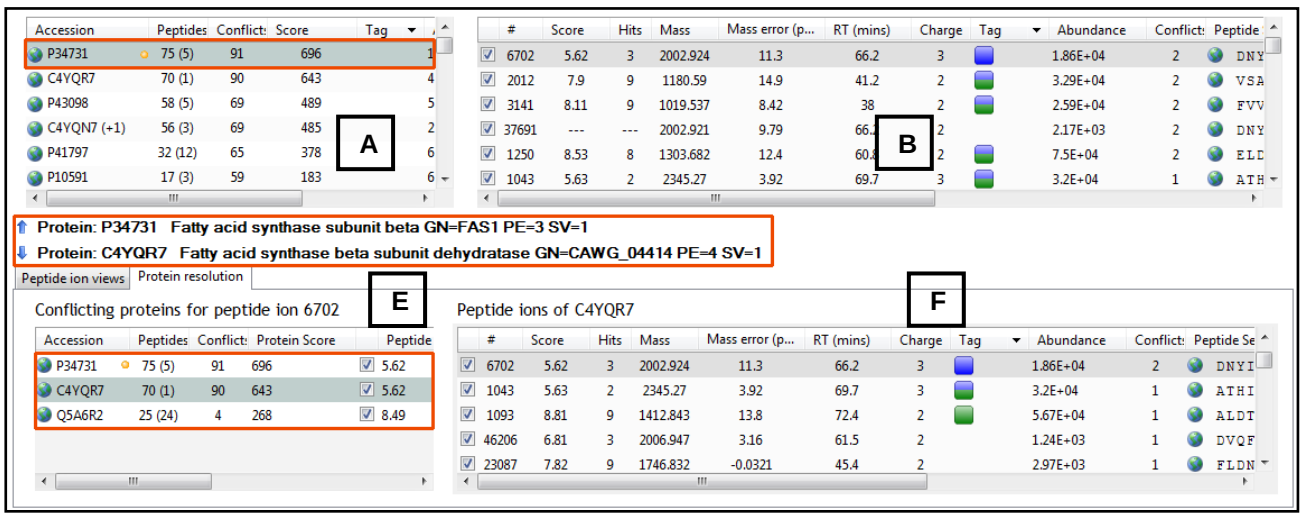


Select the first protein in list A (in this case it has 91 conflicts) the panel to the right B lists the peptides for this protein and the conflicting protein for each peptide.

Panel C shows the expression profile(s) for the peptide(s) selected in list B

Panel D shows the details for the selected peptide.

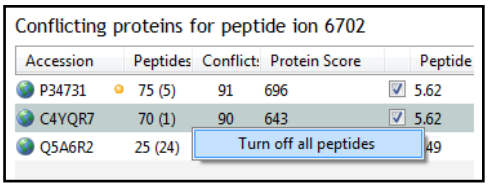
Now click on the **Protein Resolution** in Panel C to display the proteins that are conflicting.



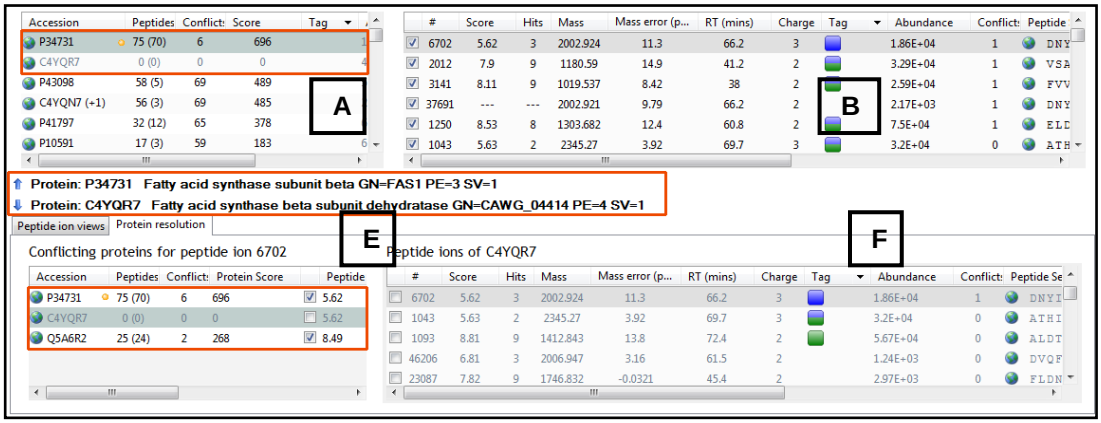
The lower left panel E displays the Conflicting proteins for the peptide ion highlighted in panel B this includes the current protein in panel A as indicated by the orange ball to the right of the accession.

The Accession and description for the 2 proteins highlighted in Panels A and E are shown in the middle margin. As most of the peptide ions are conflicting between the 2 closely related proteins one simple way to resolve these conflicts is to favour the protein with the higher score and greater number of non-conflicting peptides.

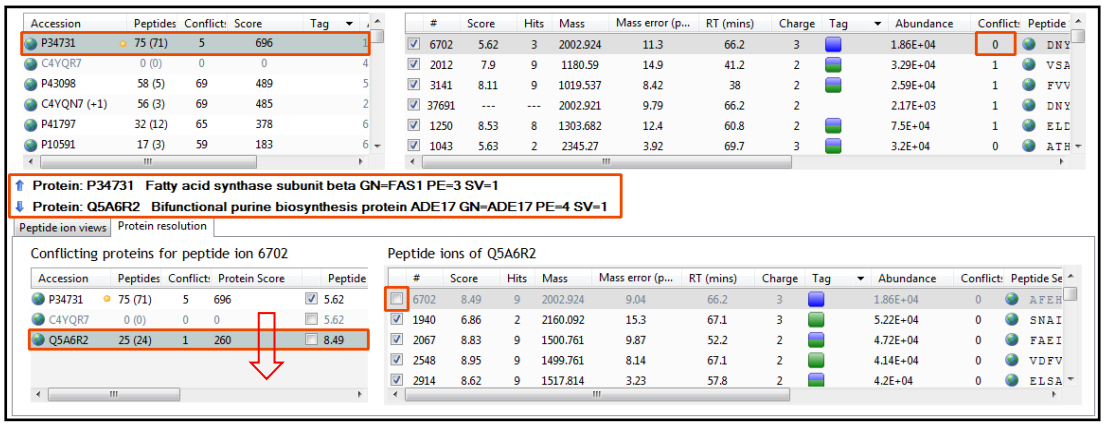
One way to do this is to right click on the lower scoring protein in panel E which only has one unique peptide and turn off all its peptides



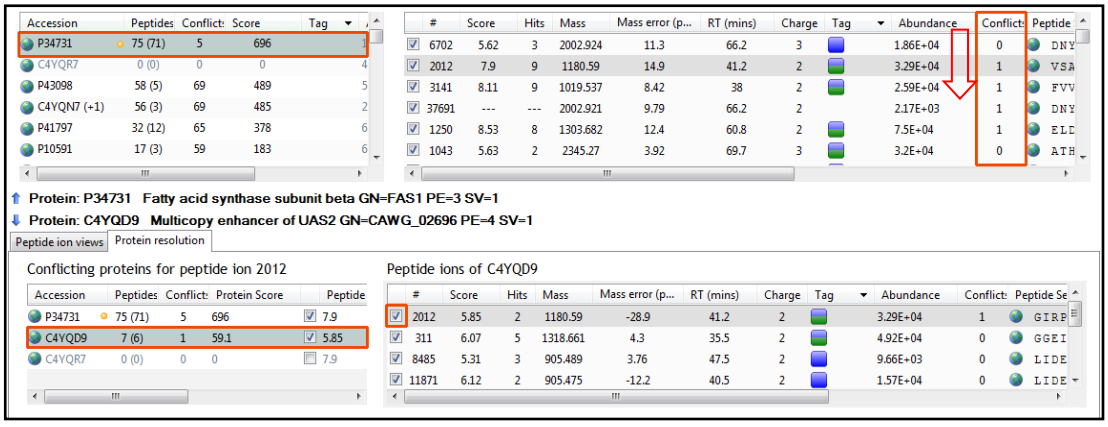
All the peptides are now switched off in panel B and all the entries for the lower scoring protein are set to zero. The higher scoring protein now has 70 non-conflicting peptides and only 6 conflicts



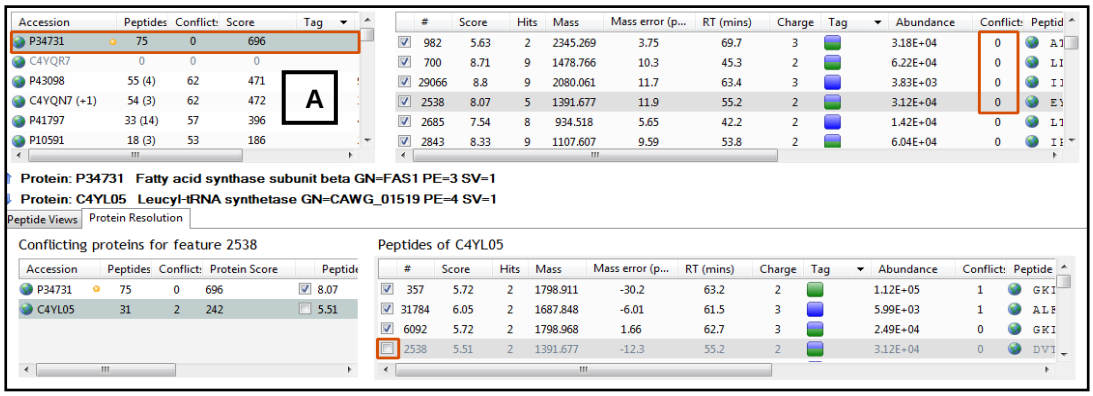
To resolve these remaining conflicts first order the conflicts in panel B and select the top one (which may still be selected) Panel B will display the peptides for this protein and the number of conflicts for each peptide. Panel E will also update to show the conflicting protein.



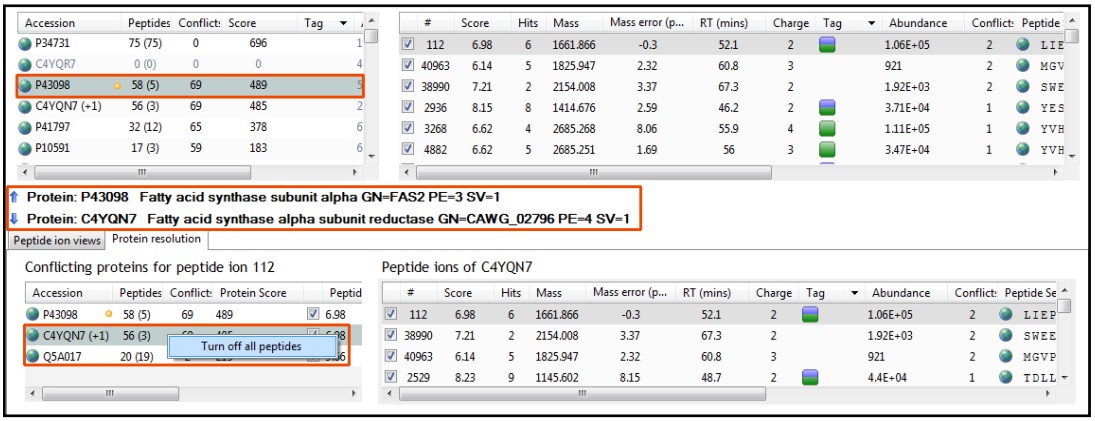
Again, favouring the protein with the higher score, but this time resolve the conflict by switching off (or un-assigning) the peptide in panel F for the protein with the lower score. By doing this the other 3 panels update to show the change in conflicts.



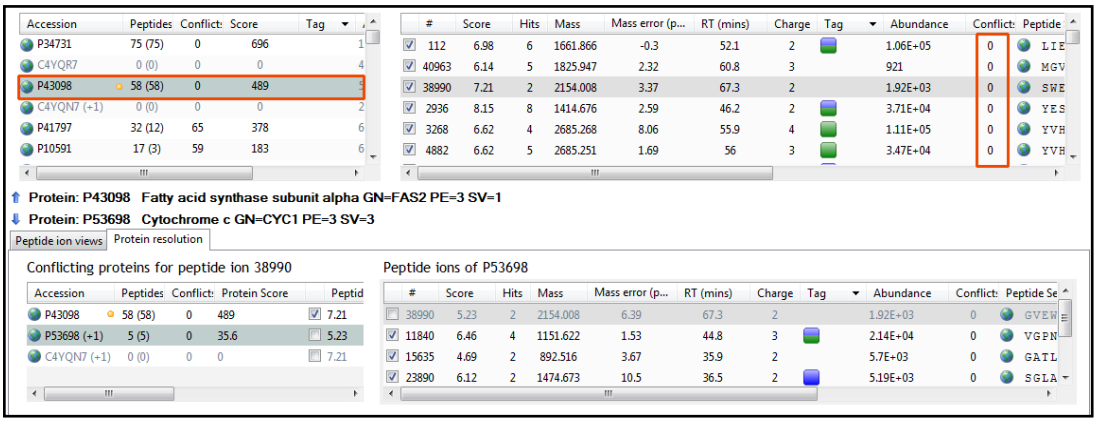
Repeat this process until there are no conflicts remaining for the current protein in Panel A.



Now repeat using a similar approach for the next protein in Panel A, here the situation is similar.

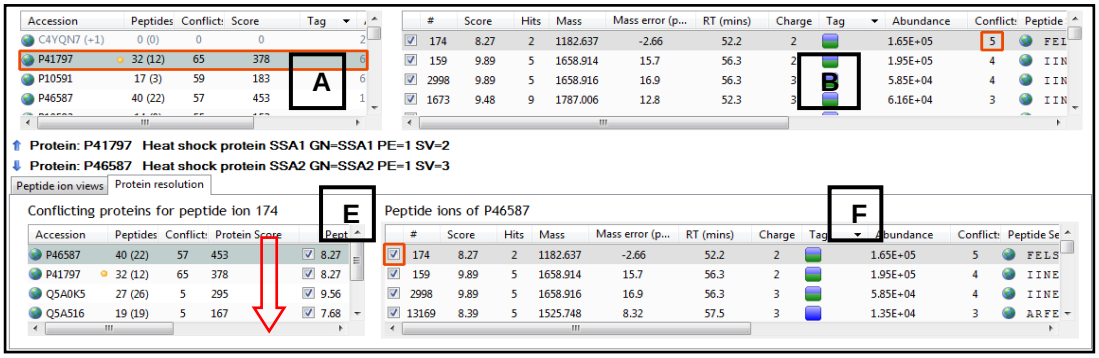


Resolution of conflicts for this protein

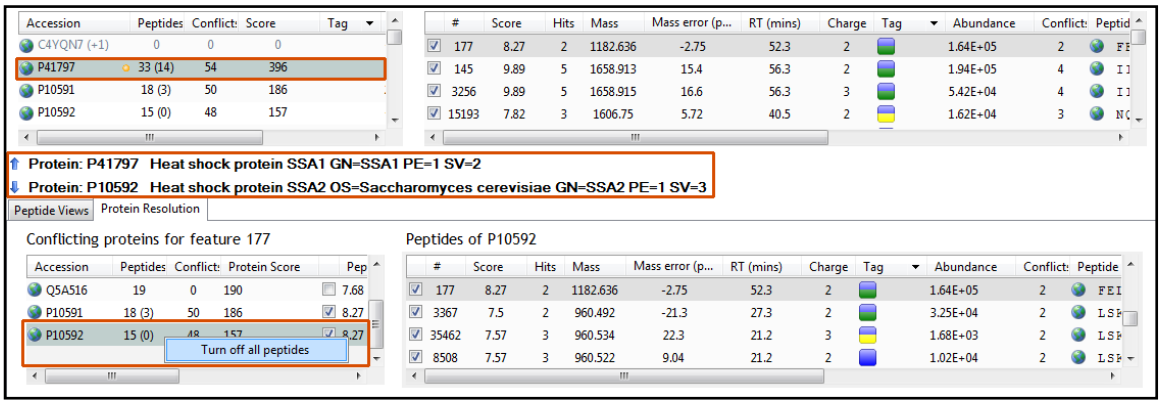


Adopting a similar approach to the next protein favouring the protein with the highest score as each conflict is examined.

In this case the first peptide for protein (P41797) has 5 conflicting proteins therefore move down the 5 conflicting proteins in panel E resolving the conflict in favour of this protein before moving on to the next peptide (which has 4 conflicts) in Panel B.



Using this approach, resolve this protein's conflicts.

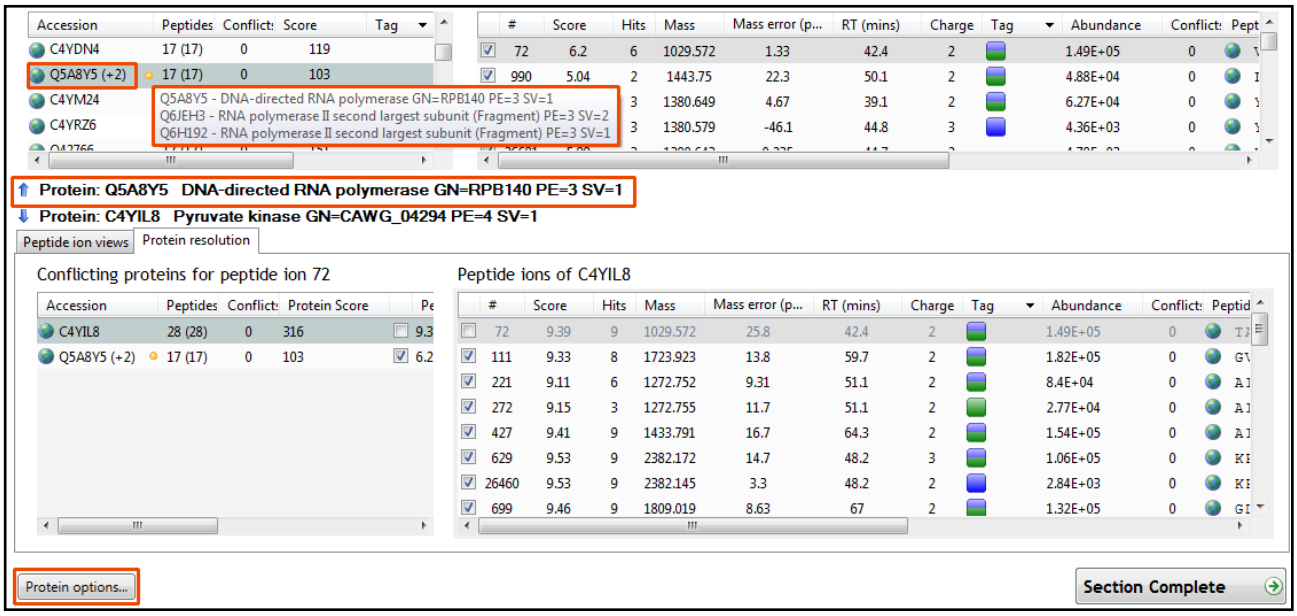


Note: where a protein has no or very few unique peptides and shares a large number of conflicts with a similarly named protein from a different species then right click and turn off all peptides

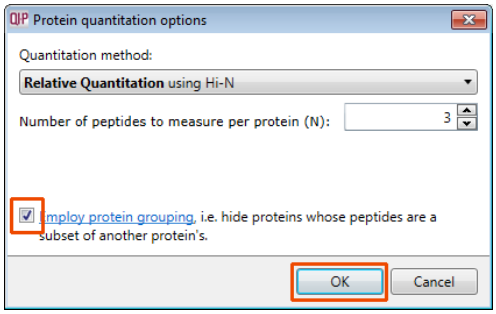
Protein Grouping

Where proteins are identified containing the same peptides then they are effectively indistinguishable aside from score. Also where one protein contains only peptides that represent a subset of another protein's peptides, the protein with fewer peptides can be subsumed into that with the greater number.

With protein grouping switched on (default setting) protein groups and the additional members are indicated by a bracketed number located after the Accession number. Taking **DNA-directed RNA polymerase** as an example, when the cursor is *held over the accession number the group members (3 in total) appear in a tool tip.*

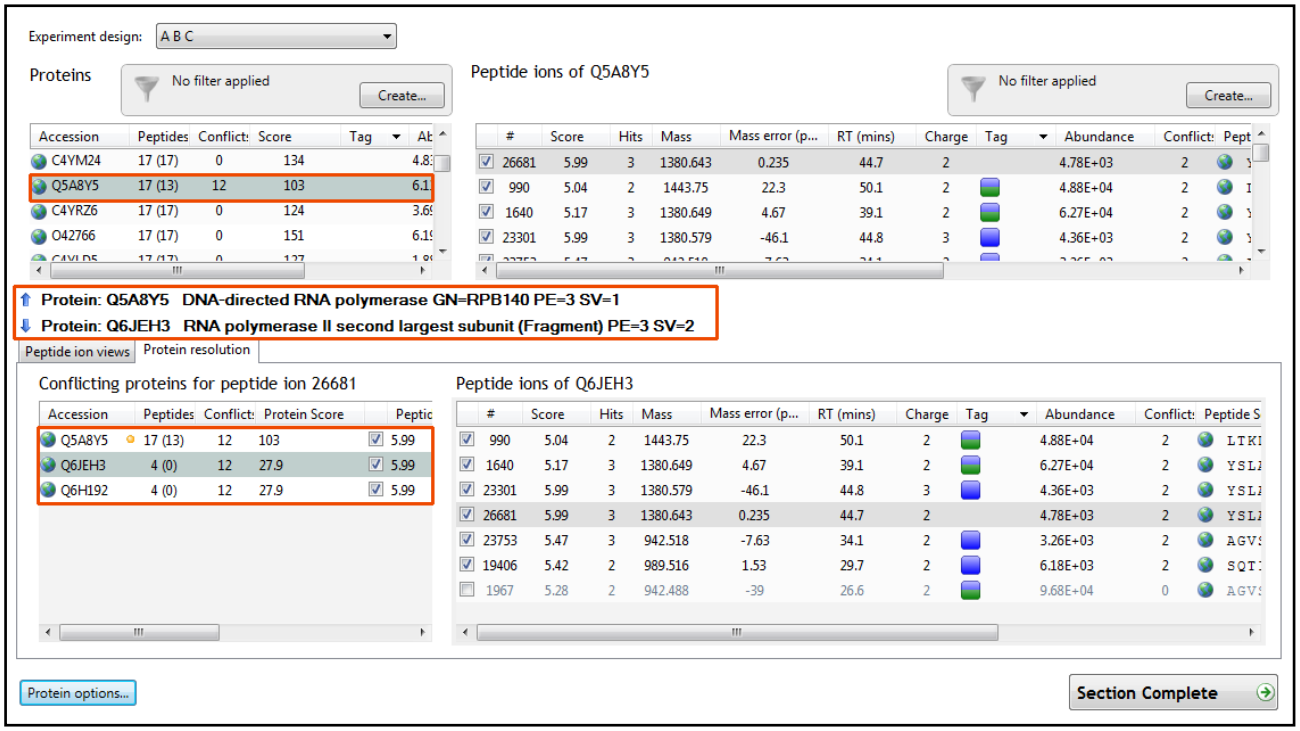


Having performed the conflict resolution with **Employ protein grouping** and **Relative Quantitation using Hi-N** now switch off the protein grouping.



As grouping is switched off the grouped proteins appear with conflicts to the other group members and the number of unique peptides appear in brackets after the peptide number.

Note: the two other group members have **no unique** peptides (brackets after the peptides field) as they are all present in **DNA-directed RNA polymerase** protein hence the reason for grouping. As a result all the conflicts are internal to the group.



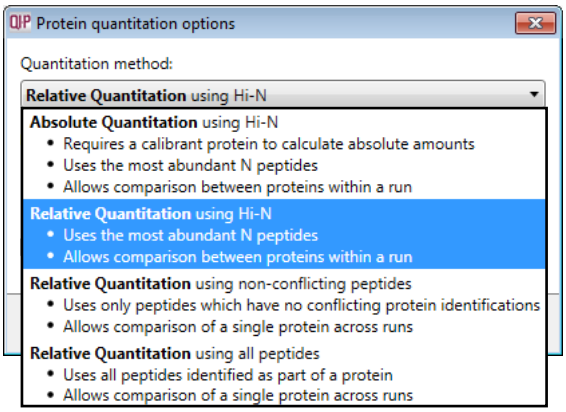
Now set the Protein Options back to **Employ protein grouping**

Protein Quantitation options

There are 4 options with which to control how the Protein Quantification is performed by Progenesis QI for proteomics.

The default option that will be applied is **Relative Quantitation using Hi-N**.

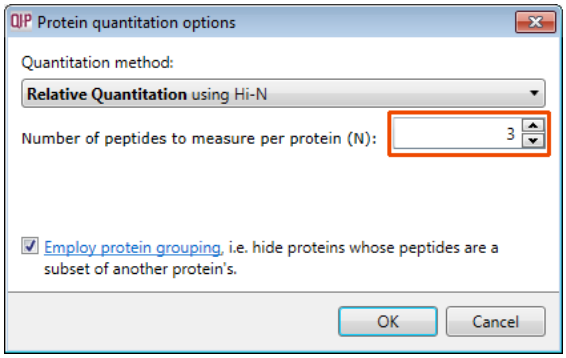
If you have selected one of the other options during the setup of the Auto Processing of your data (Stage 1) then this option will be applied.



Hi-N in Progenesis QI for proteomics is an implementation of Hi-3 as described by Silva *et al.* [References]. After peptide and protein identification, the abundance of each peptide is calculated from all its constituent peptide ions.

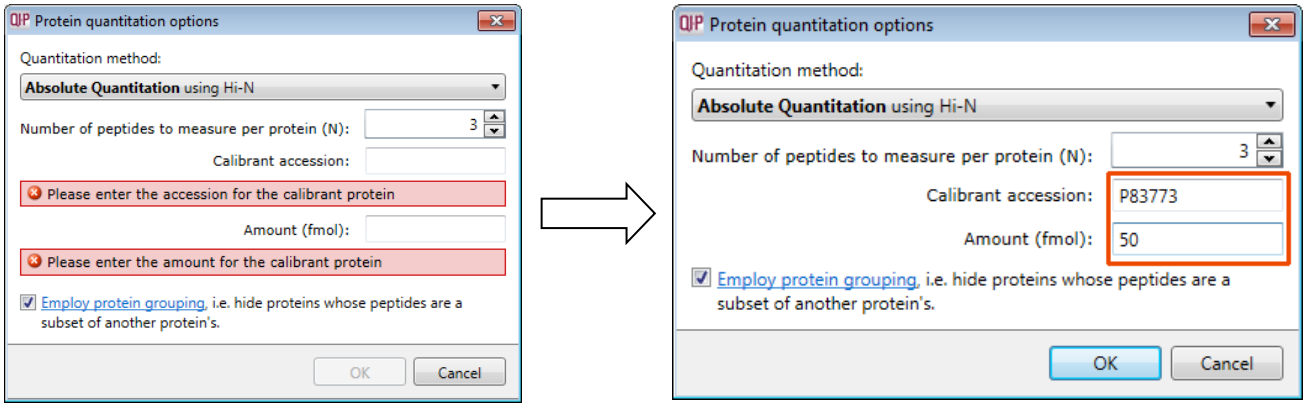
For each protein, the N most abundant peptides (N being set according to the user selection) have their **abundances averaged** to provide a reading for the protein signal.

The ranking of peptide abundance is based on the integrated value across all the runs, allowed by the accurate alignment and lack of missing values. This gives added confidence in the peptide selection, taking all runs into account to make the ranking robust.

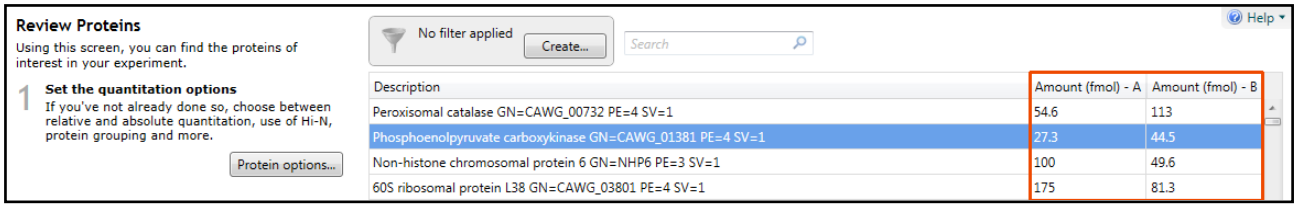


The averaged abundance readings not only make possible the **relative quantitation** of the same protein across all runs to be determined but also allow, with the inclusion of a known amount of a calibrant protein in each run, this to be converted to an **absolute** reading for protein amount.

To generate values for **absolute quantitation** enter the accession number and amount for the calibrant.



The absolute amounts, based on the calibrant used, are reported at the Review Proteins stage as additional columns (one for each condition) following the protein description.



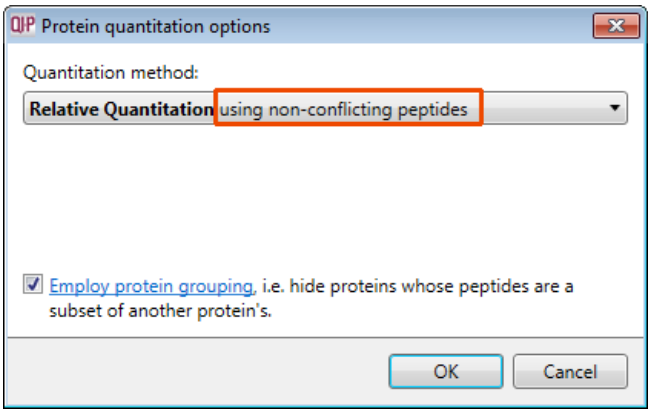
The ranking of peptide abundance is based on the integrated value across all the runs, allowed by the accurate alignment and lack of missing values. This gives added confidence in the peptide selection, taking all runs into account to make the ranking robust. These Hi-N methods allow the relative and absolute comparison of proteins within the same run.

Note: When there are **peptide conflicts** (a peptide is shared between two proteins, for example) it is important to assign the signal correctly for absolute quantitation. To do this, Progenesis QI for proteomics carries out a two-step process. Firstly, Hi-N is carried out only on the N most abundant **unique** (non-conflicting) peptides of the proteins concerned. This provides a ratio estimate for the two proteins based only on unshared peptides. The abundance of any shared peptides is then divided and allotted in this ratio between the two proteins, and the full Hi-N calculation is then applied using the divided values for conflicted peptides.

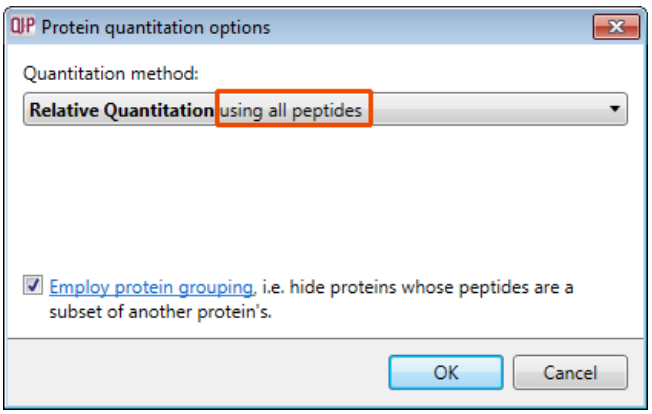
Naturally, if the conflicted peptides are not among the N most abundant in either protein initially, then this will not cause any difference in the result.

Relative Quantitation can also be performed comparing a single protein across all the runs using only the unique or non-conflicting peptides. Select the third method from the drop down

Using non conflicting peptides



The relative Quantitation can also be performed using all peptides.

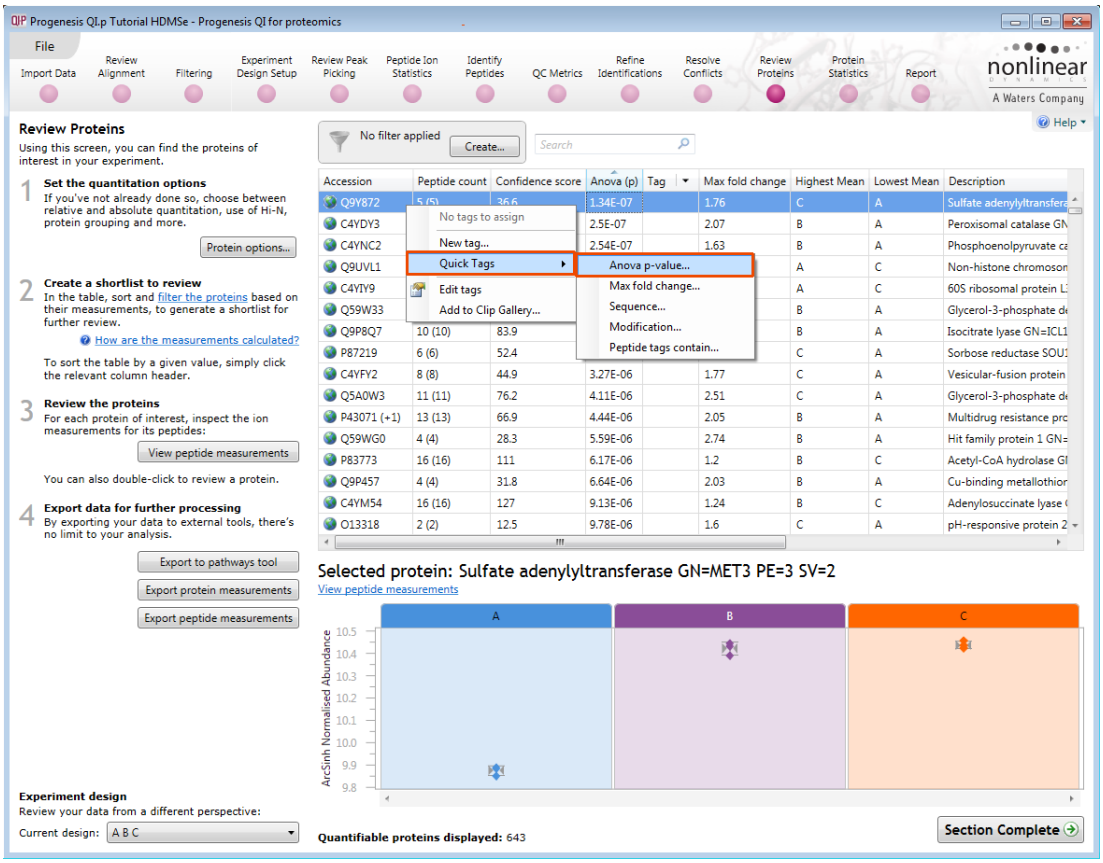


Note: if you have performed conflict resolution then there will be no difference between these methods.

Now move to the Review Proteins section by clicking on **Review Proteins** icon on the workflow at the top of the screen.

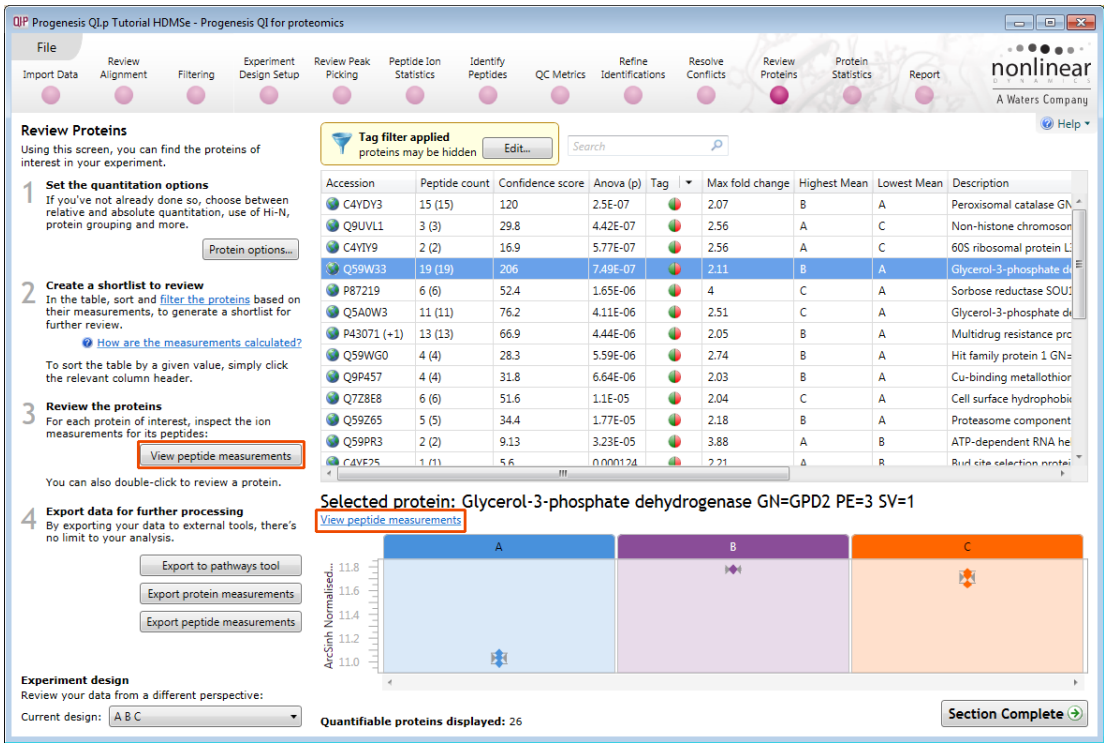
Stage 12: Review Proteins

The **Review Proteins** stage opens displaying details for all proteins. You can now create tags at the level of the proteins. Right click on the table and create Quick Tags for proteins with an Anova p value ≤ 0.05 and Max Fold change ≥ 2.



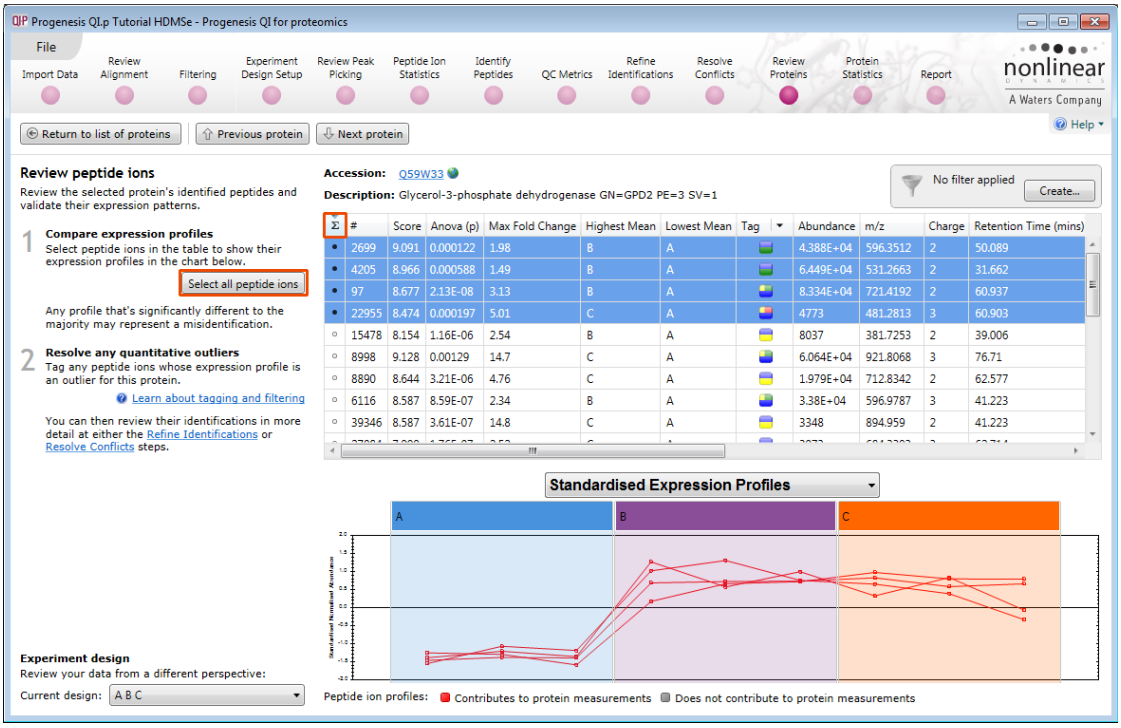
As an example let us explore Glycerol-3-phosphate dehydrogenase.

First filter the table to show only proteins showing a significant 2 fold or greater change.

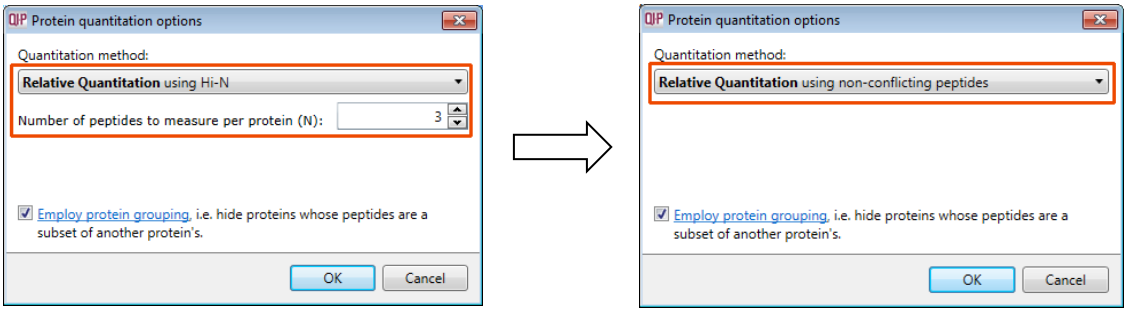


The table indicates that this protein is most highly expressed in Condition B, 2.11 fold over the lowest condition (C).

To view the corresponding peptide measurements for the current protein either double click on the protein in the table or use the **View peptide measurements** beside table.

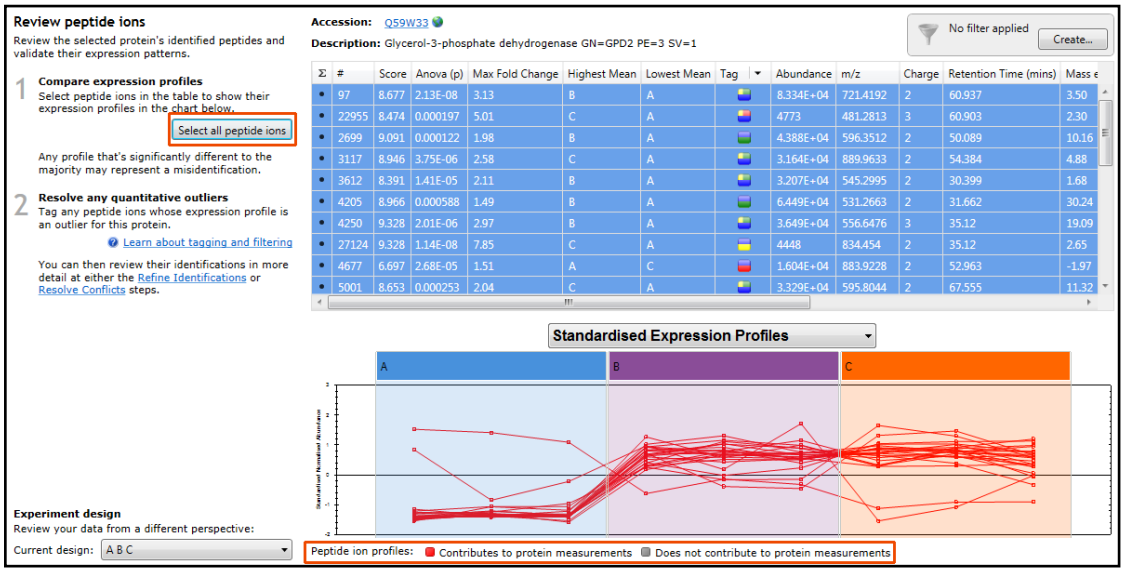


The solid icon in **Σ** column indicates that the peptide contributes to protein measurements.

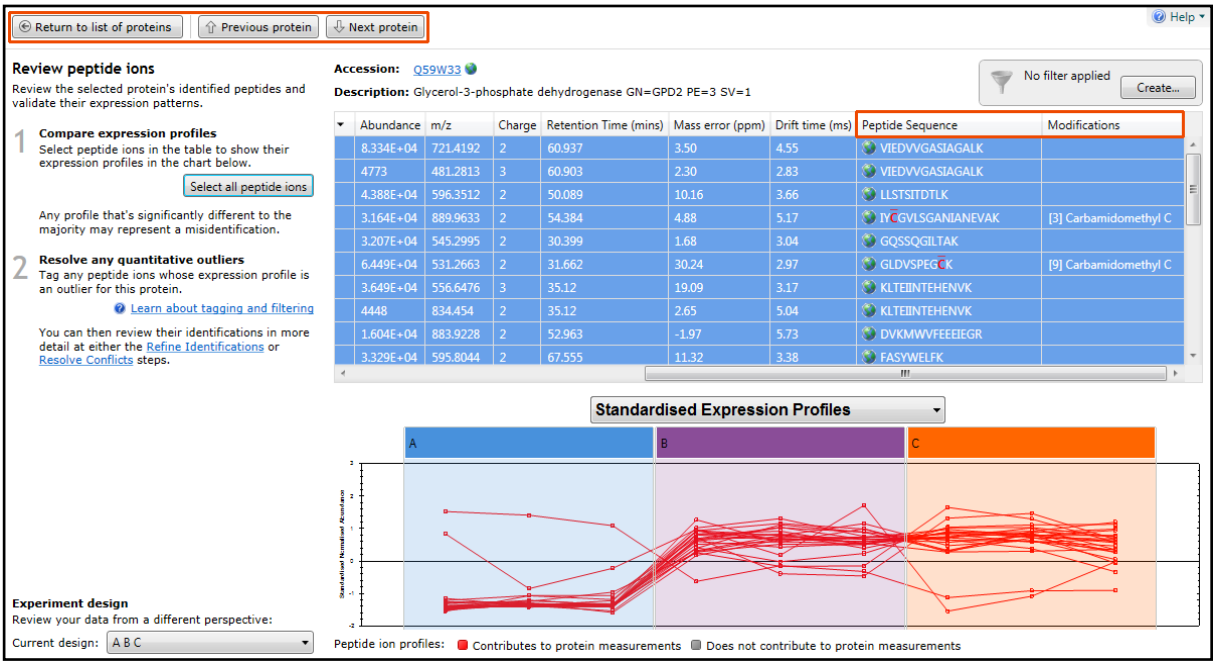


Note: the default quantitation method is based on the Relative Quantitation using Hi-N, in this case n=3 so the top 3 most abundant peptides are used to determine the relative abundance for each protein. Where there are multiple charge states the combined abundance of the charge states is used.

Changing to **Relative Quantitation using non conflicting** peptides will utilise all non conflicting peptides in the abundance calculation. (as shown below)

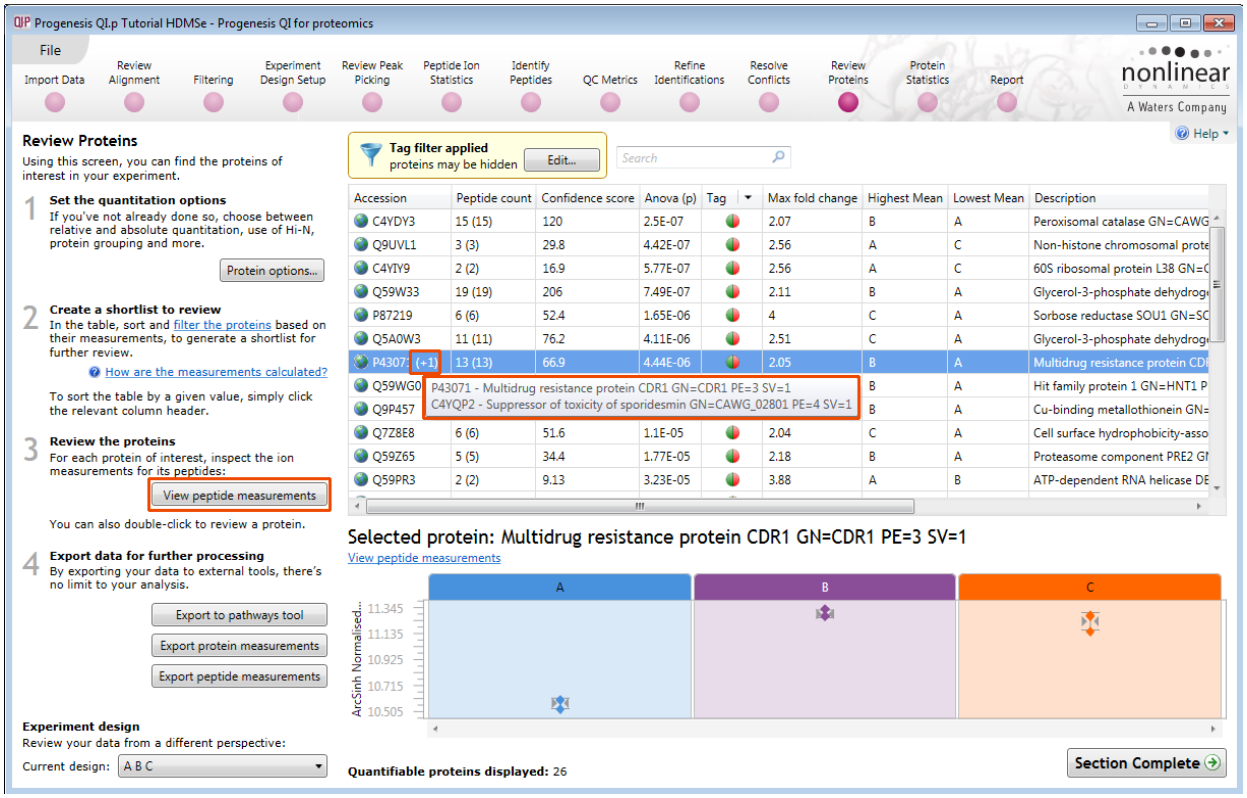


Scroll to the right on the table for the sequence and modification details. Use the navigation buttons on the top left either to review each protein's peptides by stepping through the list or return to the protein list and select a specific protein to review.



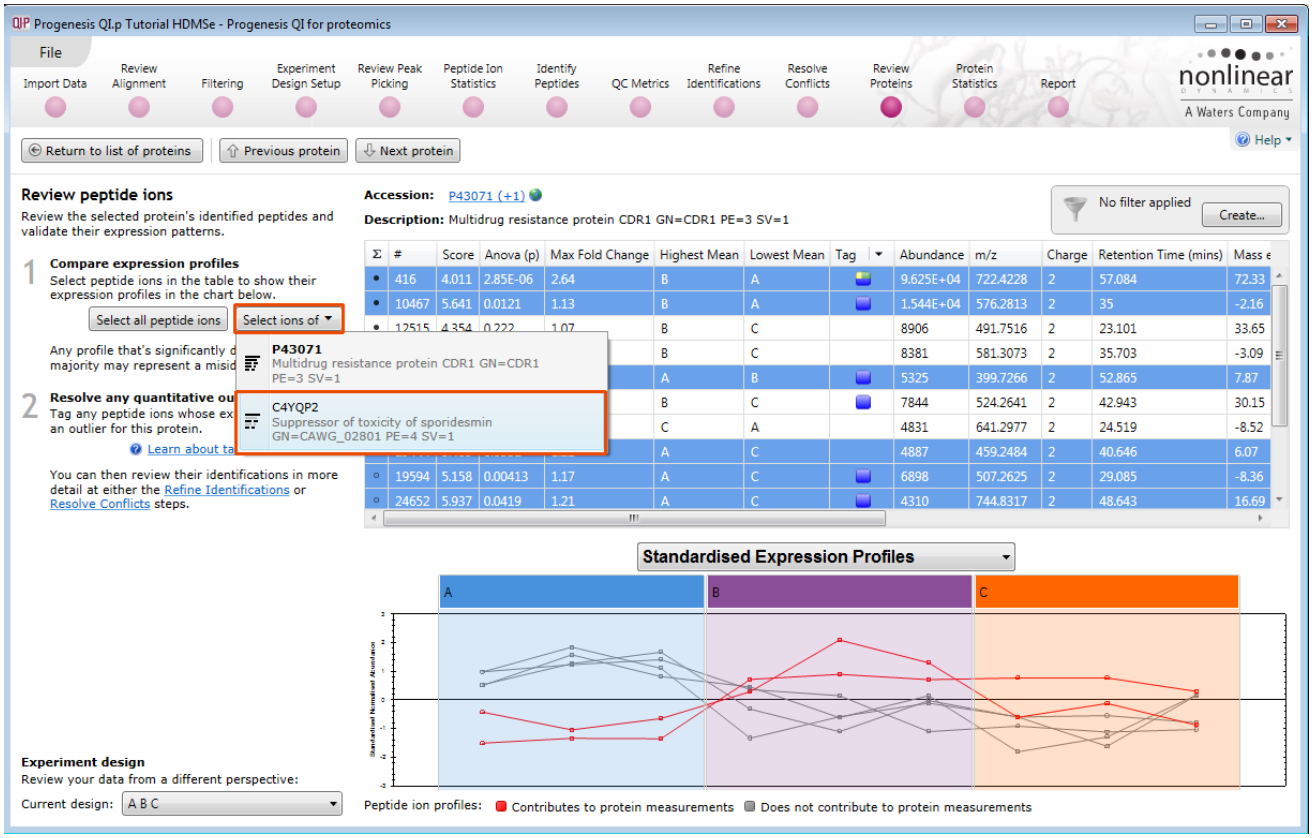
To explore protein grouping return to the Review Proteins list and set the Protein Options back to the default setting (Relative Quantitation using Hi-N).

With protein grouping switched on (default setting) protein groups and the additional members are indicated by a bracketed number located after the Accession number. The additional members can be seen by holding the cursor over the accession number.

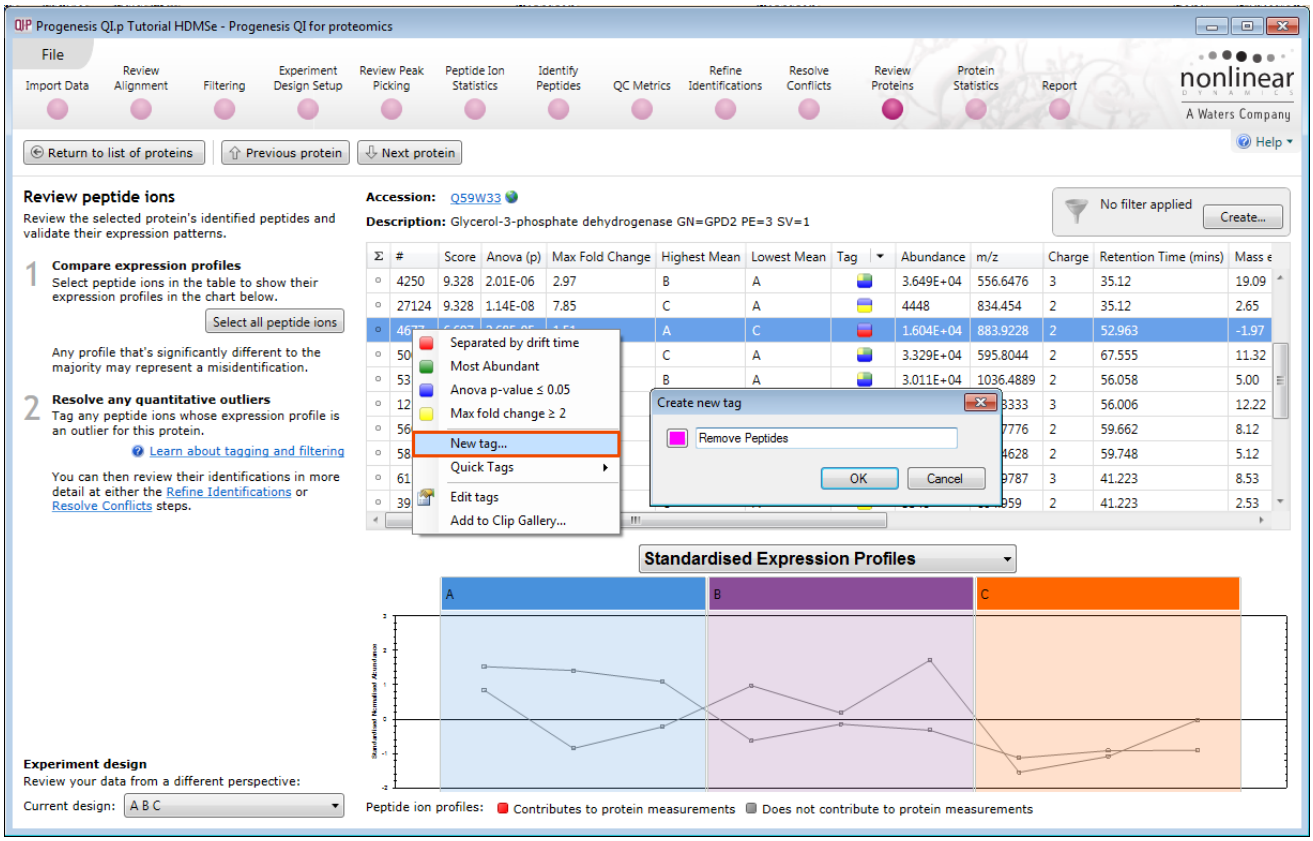


To view members peptide ions click on **View peptide measurements**.

Click on **Select ions of** to show the list of additional group members. Then as you click on one of the member proteins its peptide ions are highlighted in the table.

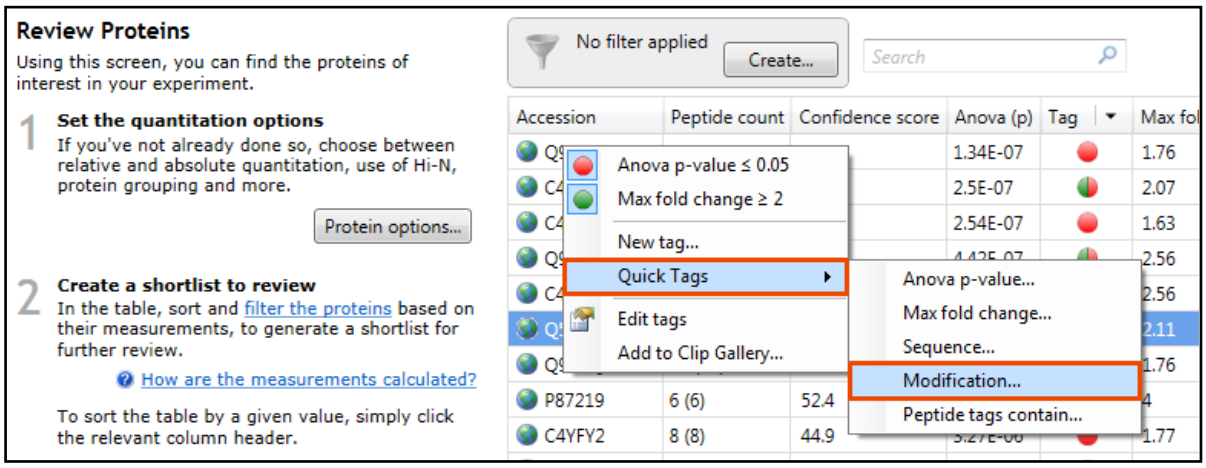


You can tag all the peptide ions for a protein or edit and tag accordingly for 'atypical behaviour'. Then remove these peptides at the **Refine identification** stage in the Workflow.

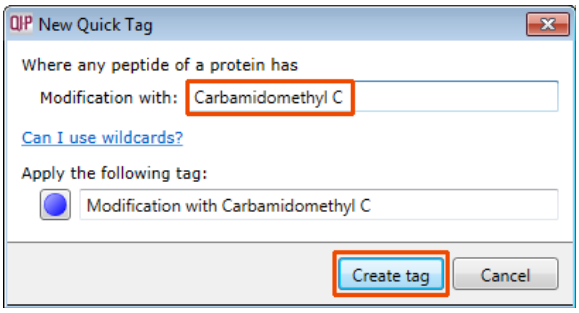


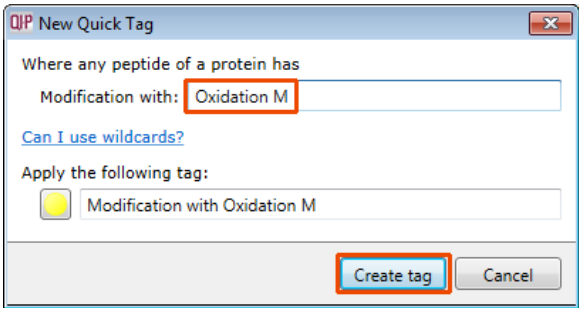
Modified proteins can be located by specifically searching for proteins containing modified peptides.

Click the **Return to list of proteins** button to return to the proteins list and right click on it and select **Modification** from the list of **Quick Tags.**

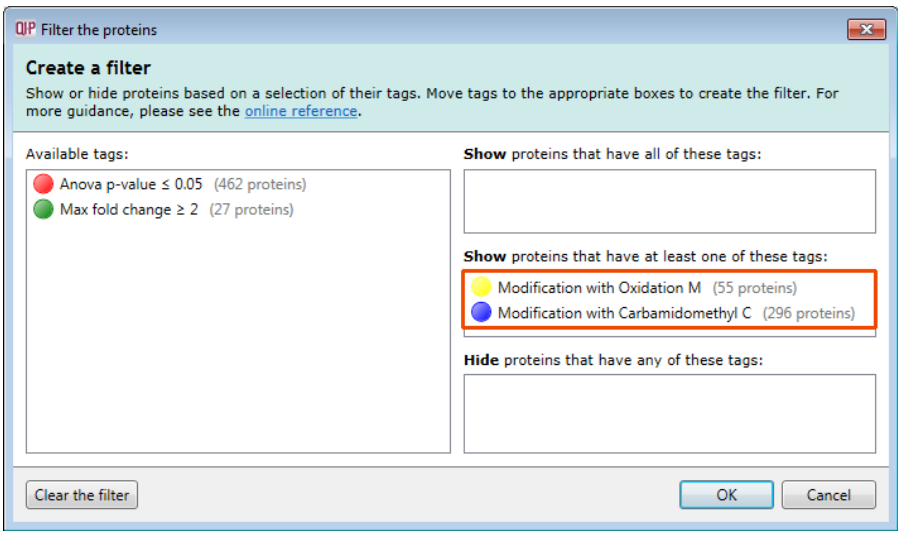


To find those proteins containing peptides with Carbamidomethylated cysteine and Oxidated methionine residues create Quick tags for each modification as shown below.

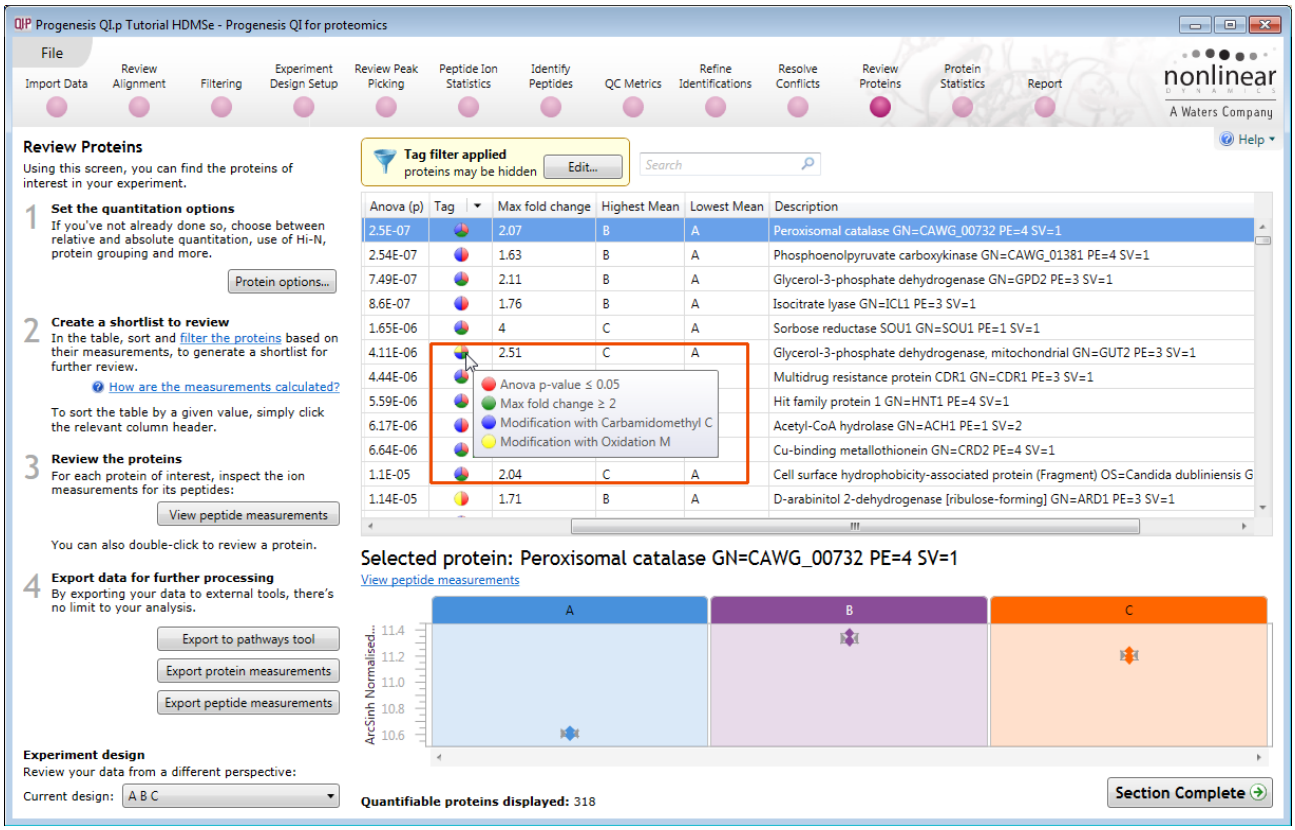




To reduce the table to displaying only these proteins with modified peptides (on cysteines and or methionines) use a tag filter to focus on these proteins by placing the appropriate tags in the **Show proteins that have at least one of these tags:**



The proteins table will now only display those proteins containing the modified peptides.



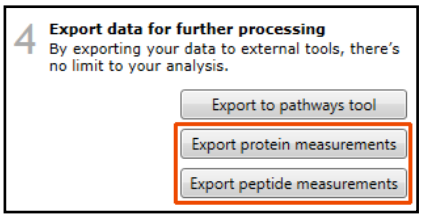
Note: hold the cursor over the tags for a description of the proteins current tags.

Note: the **Sequence** Quick tag can be used to locate Proteins containing peptides with specific motifs.

Stage 14: Exporting Protein Data

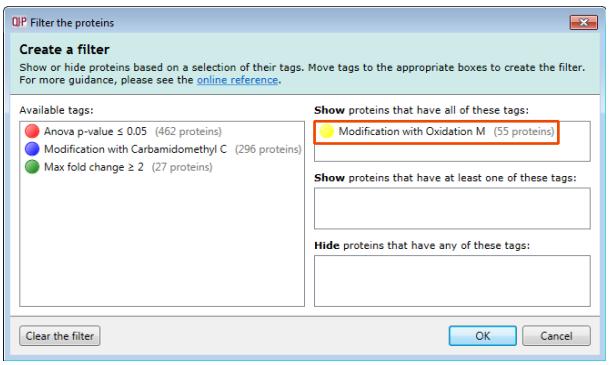
Protein data can be exported in a csv file format. You can either export the **Protein and/or peptide measurements** using the options in the File Menu or use the buttons under Step 4 both available at the **Review Proteins** stage**.**

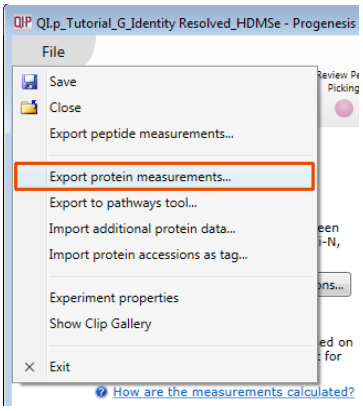
As an example of Data export use the Tag filtered set from the previous section. Where you are only going to export measurements for those proteins that a have Oxidised Methionine residues.



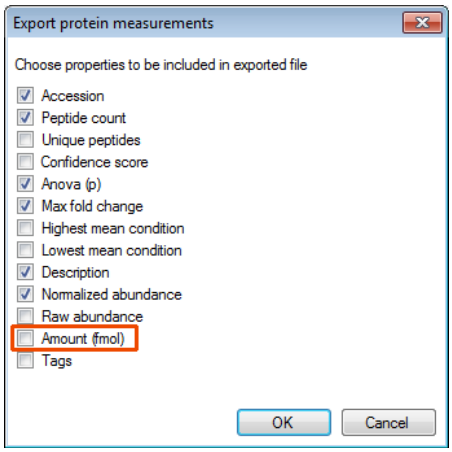
First set the tag filter as shown below. Then select **Export Protein Measurements**.

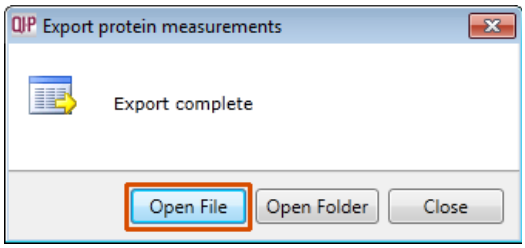
The Export Protein Measurements dialog opens. Select the required fields and click OK. **Save** the file and then open the exported data file using the dialog that opens.





Note: if you have performed Absolute Quantification then the 'Amount' field will be available.



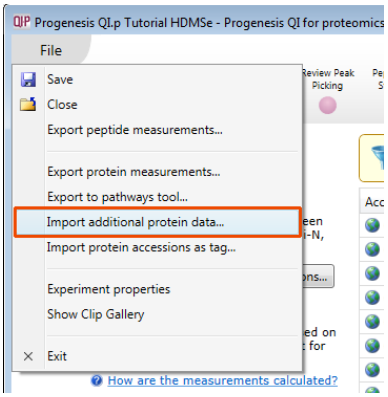


Excel will open displaying the exported protein measurements

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1						Normalized abundance								
2						A			B			C		
3	Accession	Peptide c	Anova (p)	Max fold change	Description	A_01	A_02	A_03	B_01	B_02	B_03	C_01	C_02	C_03
4	Q5A0W3	11	4.10E-06	2.507552239	Glycerol-3-phosphate dehydrogenase	15.12836	17.8059	17.42256	33.67985	37.23808	37.10354	43.16004	42.30314	40.80919
5	P43066	5	1.14E-05	1.707978538	D-arabinitol 2-dehydrogenase [ribulo	30.01781	31.6882	29.63293	51.33972	52.68843	51.97678	44.99624	45.58278	40.76944
6	Q5ADM7	22	0.004402752	1.792122914	Glyceraldehyde 3-phosphate dehydrc	267.1653	354.06	332.0898	624.3207	489.5577	477.9149	617.7853	601.1242	489.5483
7	Q59WJ0	18	0.007745492	1.225181506	60S ribosomal protein L8-B GN=RPL82	308.0216	337.7146	319.4018	325.3763	331.0352	326.1782	282.0446	278.2706	241.68
8	P46586;Q5F0I1	7	0.008317701	1.098727897	ATP phosphoribosyltransferase GN=H	53.02957	52.28622	50.36581	56.76164	58.39307	55.89701	54.10511	51.59631	53.62655
9	C4YMB9	29	0.009689284	1.140862508	C-1-tetrahydrofolate synthase, mitocl	166.7958	175.6785	163.3459	186.2693	184.9605	179.6682	166.7231	162.8926	153.2628
10	Q5AG68	10	0.009917467	1.297286139	Nucleoside diphosphate kinase GN=Y	241.9186	236.6929	232.7635	323.7893	308.8791	290.1886	282.4256	282.7152	236.6925
11	Q5AJD0	16	0.013316685	1.158584277	ATP-dependent RNA helicase DBP5 GI	55.51844	54.55947	50.61161	44.59633	47.80023	46.29815	48.98838	49.0226	45.45075
12	Q59LU0	19	0.013858782	1.165314395	ATP-dependent RNA helicase DBP2 GI	36.04126	37.33473	37.93417	46.36719	42.70991	40.63424	40.97525	41.85573	40.03178
13	C4YRA2	15	0.014291036	1.211200756	Eukaryotic translation initiation factor	36.96303	37.20492	36.82498	48.78505	42.72804	42.92163	41.79669	44.32693	39.57481
14	Q5A900	12	0.019494899	1.239238666	40S ribosomal protein S2 GN=RPS21 PI	221.6514	235.5562	222.803	194.9576	179.9049	173.8701	200.3961	197.6079	166.4748
15	Q59PL9	22	0.02750144	1.169704475	Eukaryotic translation initiation factor	102.4817	111.1875	105.8497	105.297	102.1849	100.3602	96.71522	93.58582	82.86103
16	Q59S75	9	0.02927799	1.171272074	Ribosomal protein L1 GN=RPL10A PE=	213.5715	238.9348	211.4338	253.7061	264.8743	244.0113	225.9009	223.1294	202.0495

Note: where there are multiple group members the other accession numbers are also exported.

At each stage in the Work flow there are a number of Export and Import options available from the **File** Menu. This includes the option to **Import Additional Protein Data** which can be used to increase the Protein meta data and also be used to sort the existing tabular data.



Exporting Protein Data to Pathways Tool(s)

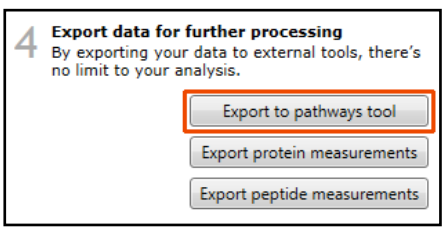
Using Progenesis, you can export protein lists to pathway analysis tools to help you understand your data in a wider biological context.

Currently Progenesis for proteomics supports the export to:

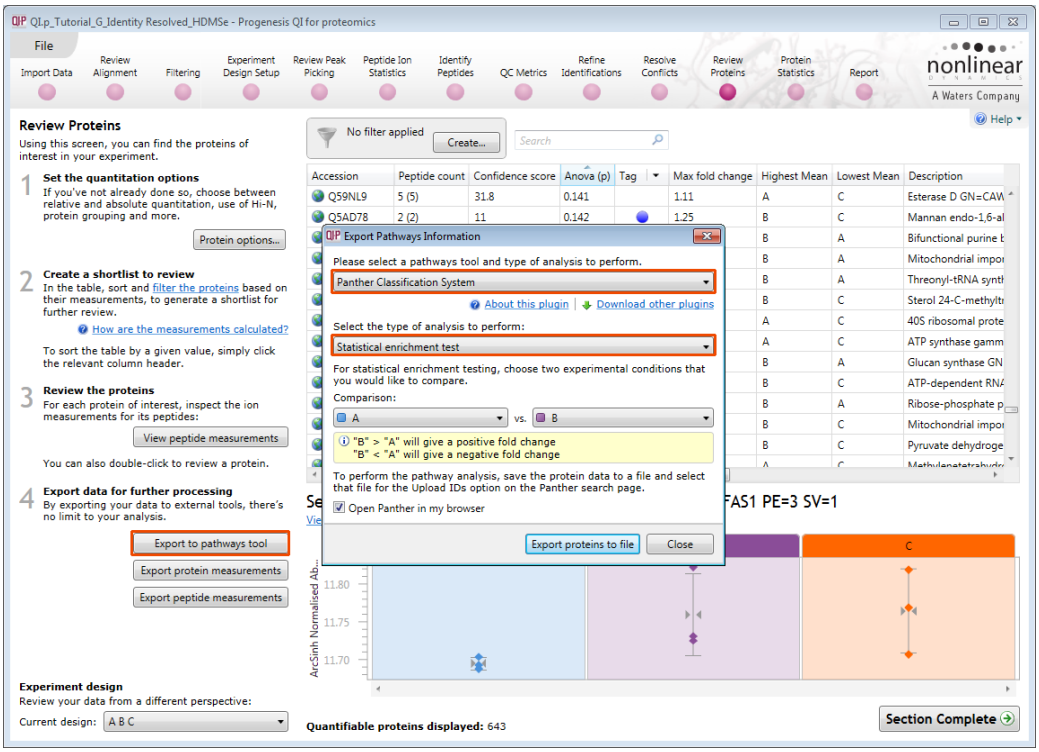
IMPaLA: which aggregates and queries many other pathway analysis tools including KEGG and Reactome)

PANTHER: classifies proteins for high-throughput analysis.

Note: Plugins for these tools are provided as standard.

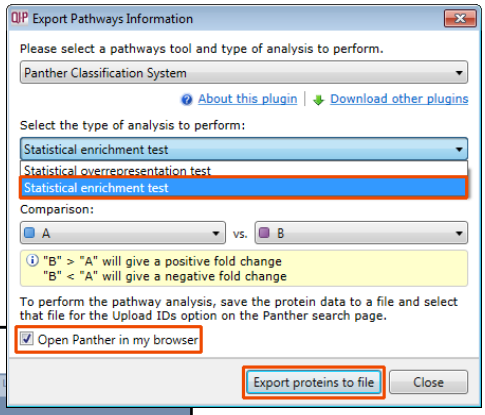


First select the protein data to export to the pathways tool using tag filtering to 'focus' the set to export. Then click **Export to pathways tool**, and select the tool followed by the test to be performed

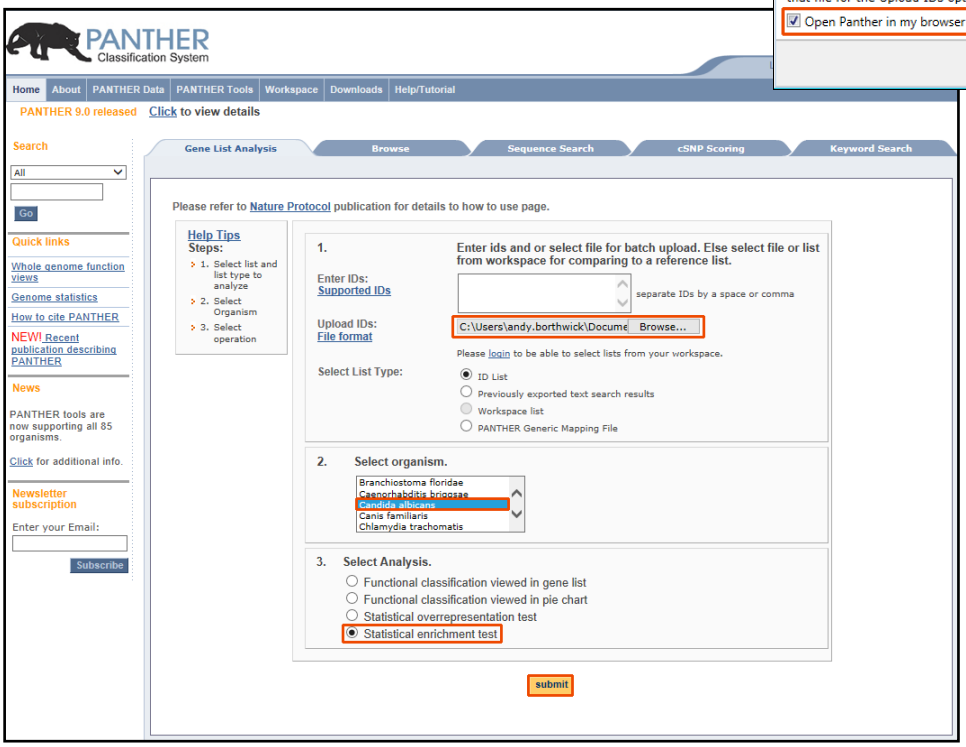


Select either **Statistical over-representation test** or **Statistical enrichment test.**

Make sure the **Open Panther in my browser** is ticked and then click **Export proteins to file.** Save file with appropriate name.

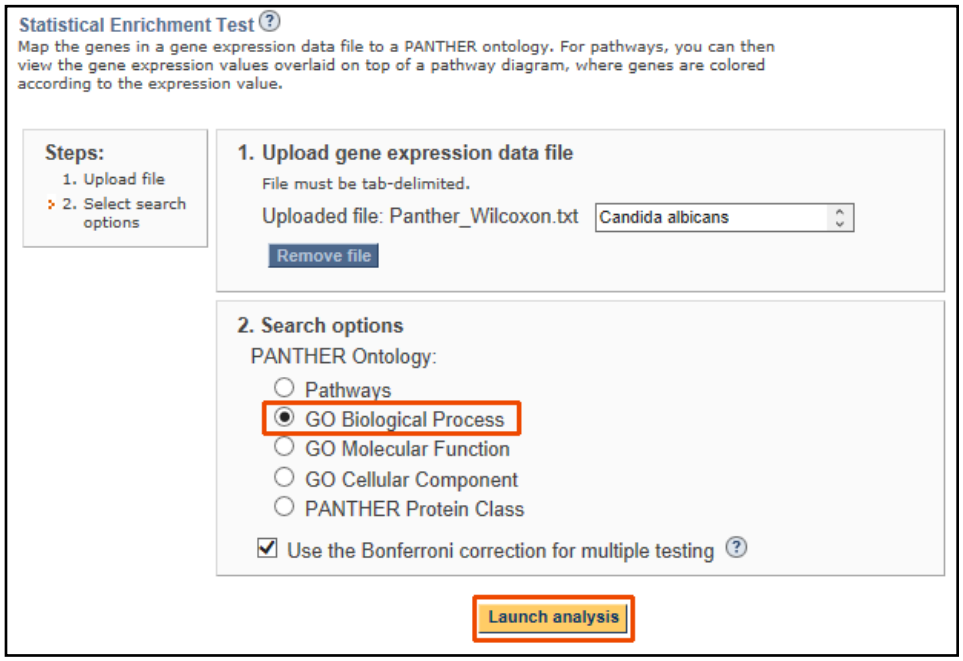


When Panther opens locate the file, select the appropriate organism (Candida albicans) and tick Statistical enrichment test.

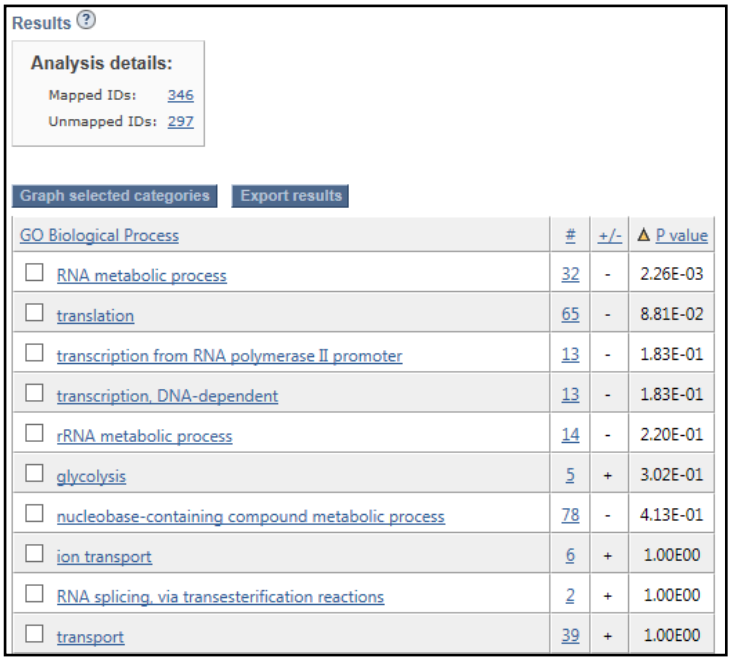


Click **submit**.

Select a search option, in this example GO Biological process.



Analysis results list is returned:



GO Biological Process	#	+/-	P value
RNA metabolic process	32	-	2.26E-03
translation	65	-	8.81E-02
transcription from RNA polymerase II promoter	13	-	1.83E-01
transcription, DNA-dependent	13	-	1.83E-01
rRNA metabolic process	14	-	2.20E-01
glycolysis	5	+	3.02E-01
nucleobase-containing compound metabolic process	78	-	4.13E-01
ion transport	6	+	1.00E00
RNA splicing, via transesterification reactions	2	+	1.00E00
transport	39	+	1.00E00

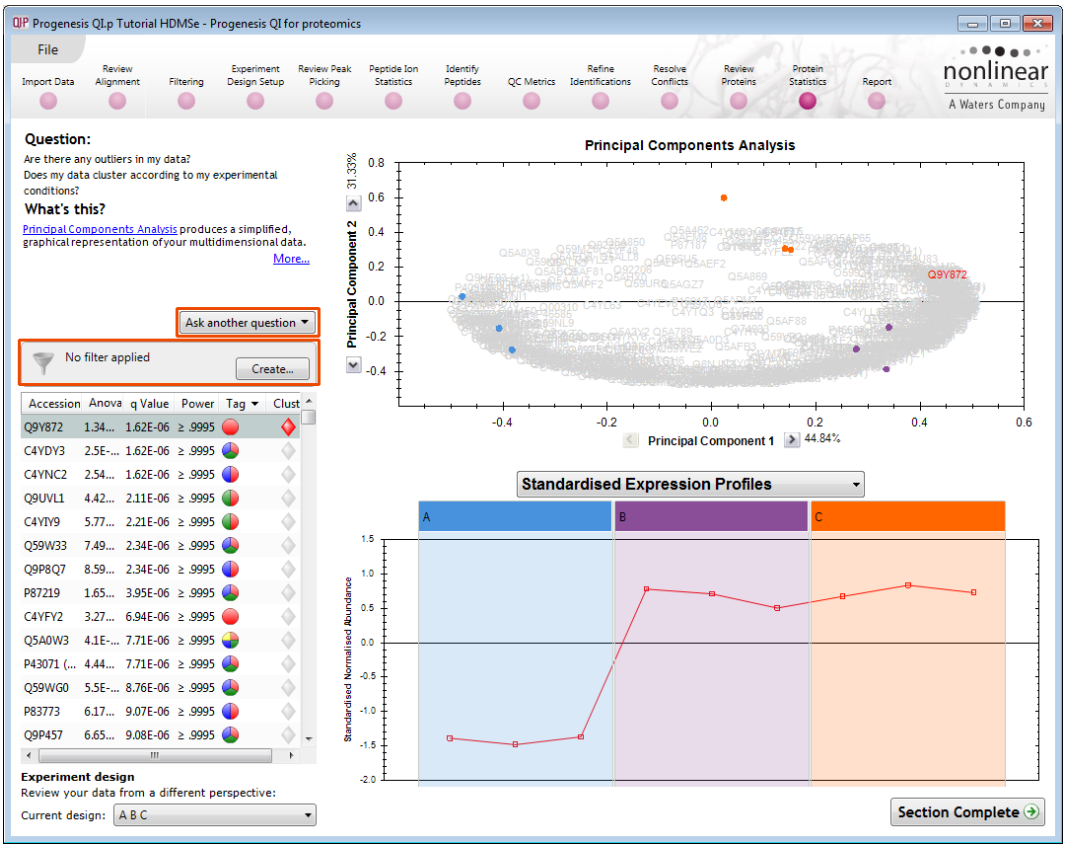
Click to explore returned analysis.

Note: the success of a Pathways analysis is dependent on the organism under study being available to search using **Impala** and **Panther.**

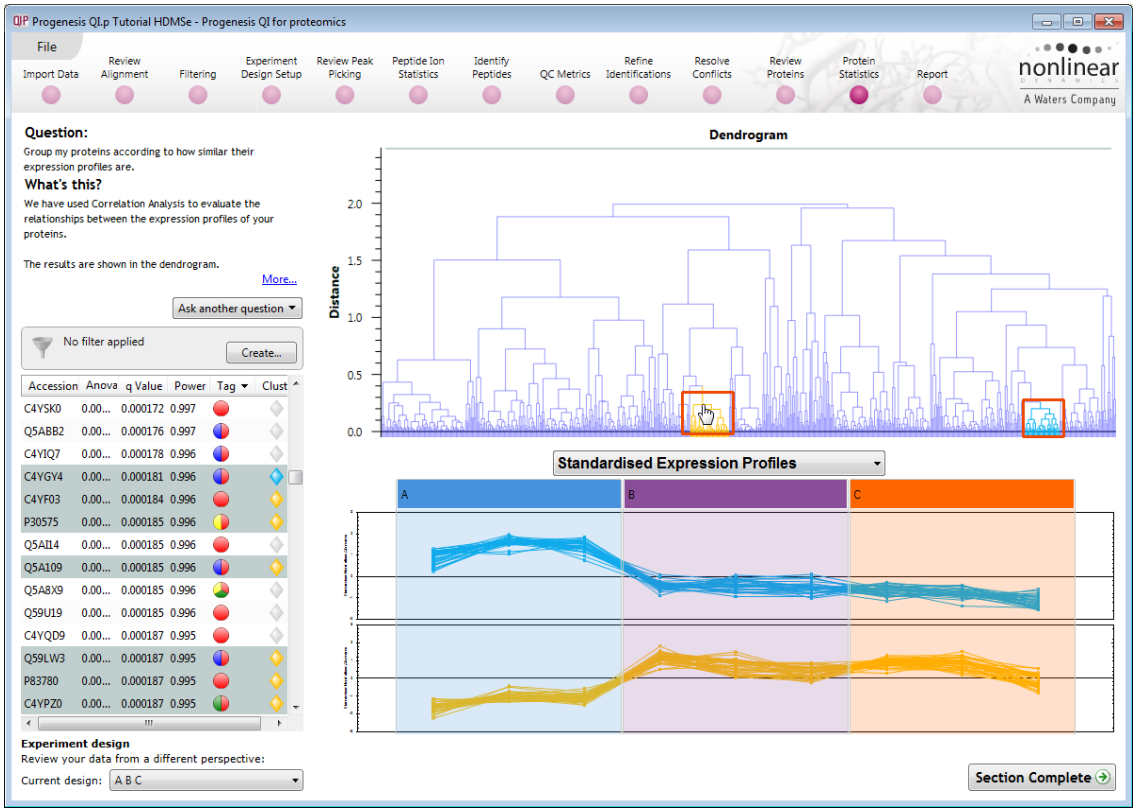
Now move to the **Protein Statistics** section by clicking section complete.

Stage 15: Protein Statistics

Protein Statistics opens with a Principal Components Analysis (PCA) for all the proteins displayed.



The Multivariate Stats can now be applied to all or subsets of proteins as determined by the current Tag filters. Allowing you to identify similar paterns of expression using the Correlation Analysis. Click on 2 of the branches (holding the **Ctrl** key down) to see differing patterns of expression.

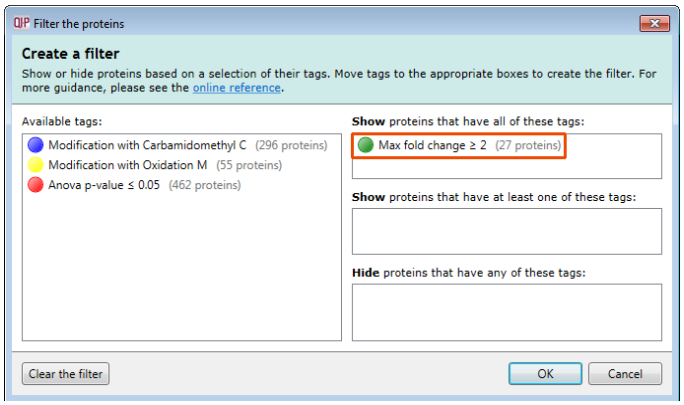


Now move to the **Report** section to report on Proteins and /or peptides.

Stage 16: Reporting

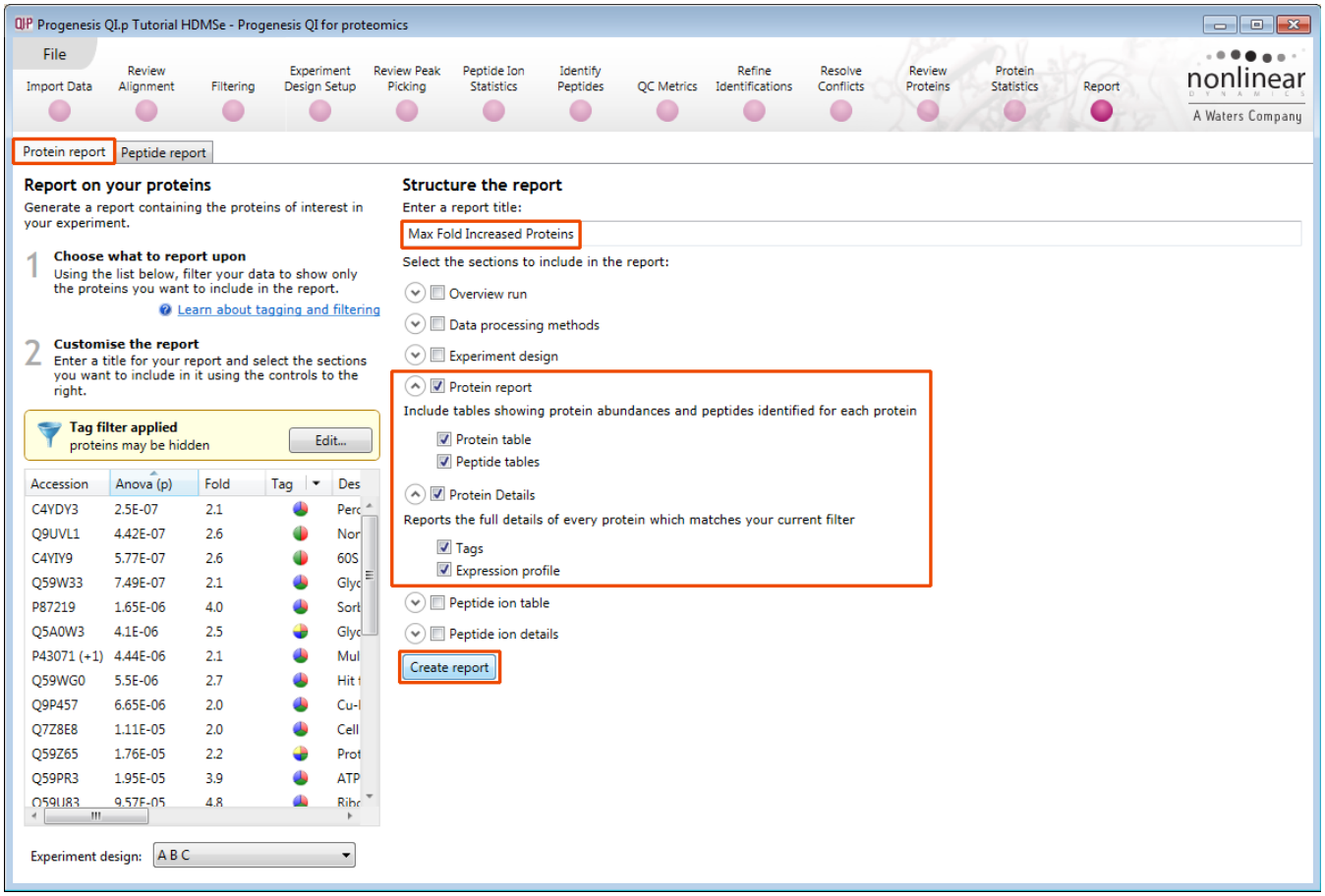
The **Report Design** stage allows you to select what views you want to include in a report based on the list of **currently selected proteins.**

Note: this facility is used to generate Html reports on a limited selection of Proteins in your data. Creating a report on all the data in your experiment can take a long time



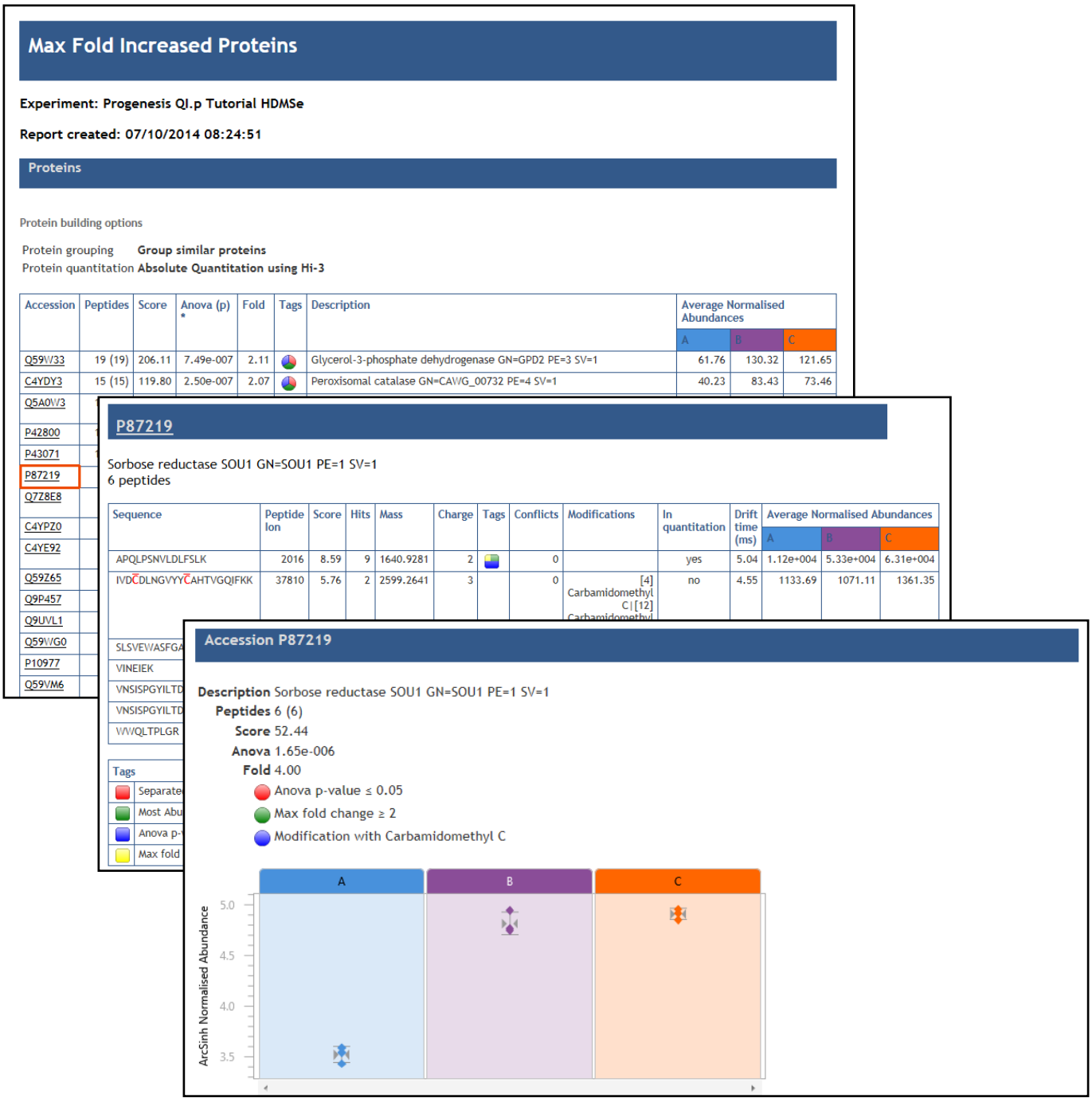
As an example we will create a report for **only** the proteins showing a Max Fold change of greater than 2.

1. First reduce the proteins to report on by selecting the '**Max fold change ≥ 2**' tag. In this example it reduces the number of proteins in the table to 27.
2. Expand the various Report Design options (by default they are all selected)
3. Un-tick as shown below
4. Click **Create Report**



This opens a dialog to allow you to save the report, after which it will be opened in the form of a web page.

Click on the **Accession No**. in the proteins section of the Report and this will take you to the Assigned peptides for this protein



Having closed the report it can be reopened by double clicking on the saved html file.

Note: you can also copy and paste all or selected sections of the report to Excel and/or Word.

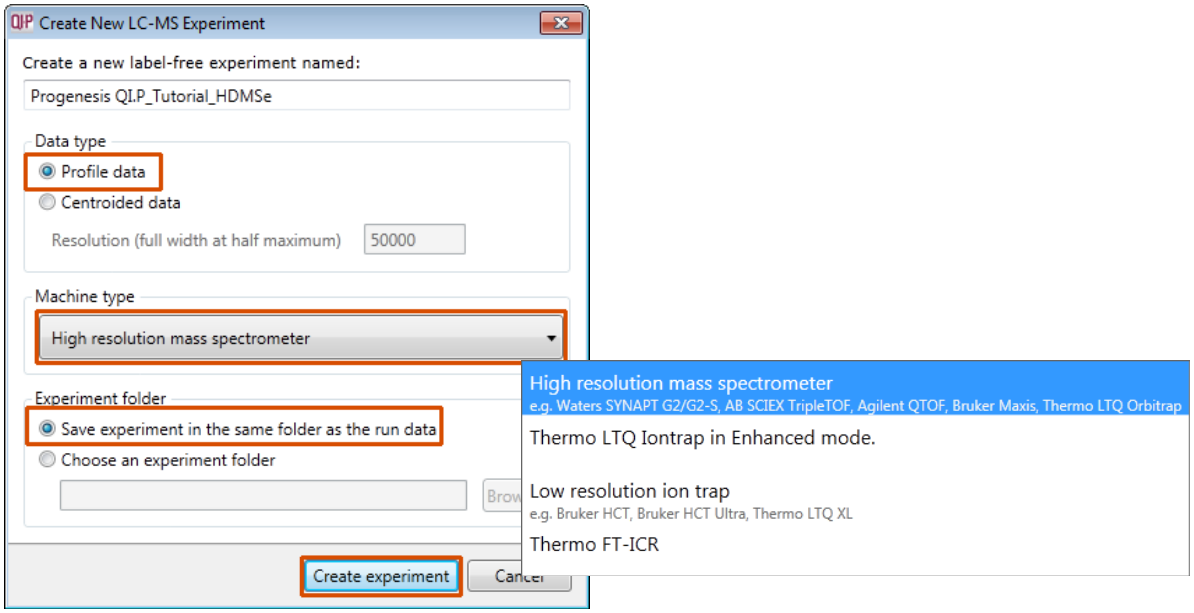
Note: there are separate panels for reporting on Proteins and Peptides.

Appendix 1: Stage 1 Data Import and QC review of LC-MS data set

You can use your own data files, either by directly loading the raw files (Waters, Thermo, Bruker, ABSciex and Agilent) or, for other Vendors, convert them to mzXML or mzML format first.

To create a new experiment with your files: open Progenesis QI for proteomics and click **New,** bottom left of the **Experiments** page and give your experiment a name. Then select data type, the default is 'Profile data'.

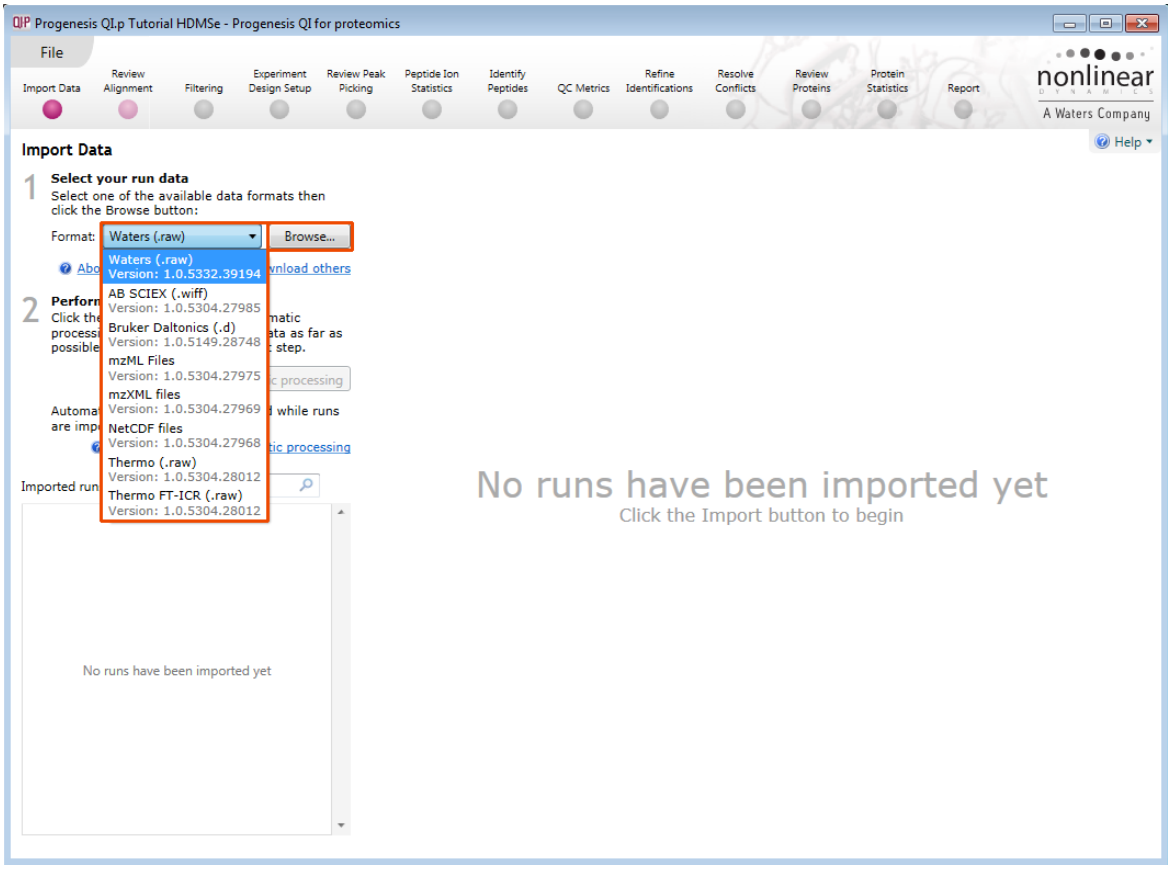
Note: if you have converted or captured the data as centroided then select Centroided data and enter the Resolution for the MS machine used.



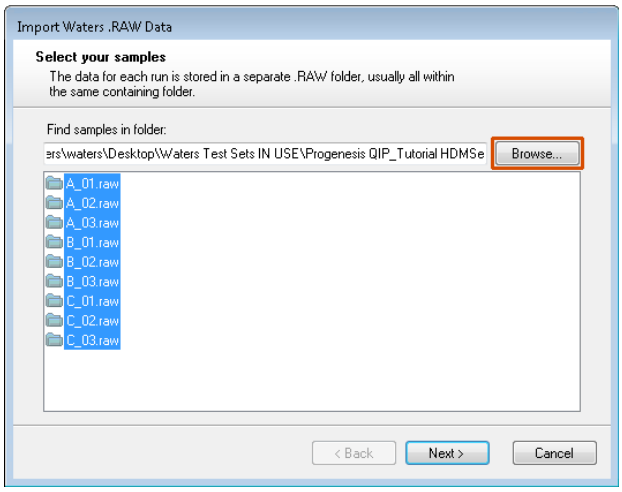
Click **Create experiment** to open the LC-MS Data Import stage of the workflow.

Select the 'Import Data file format', in this example they are Waters/SYNAPT data

Then locate your data files using **Import...**

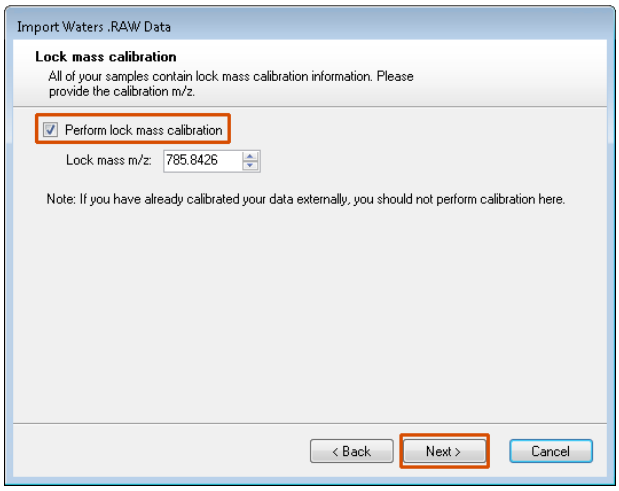


Locate and select all the .RAW folders (A_01 to C_03).



On importing, the lock mass calibration is read and presented on this dialog

You can, if required, alter the lock mass calibration at this step.



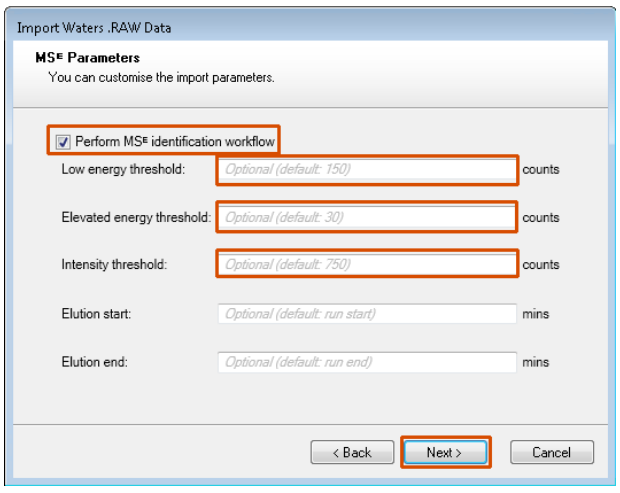
You can set the MS^E parameters depending on how your data was acquired.

For HDMSe the settings are 150, 30 and 750

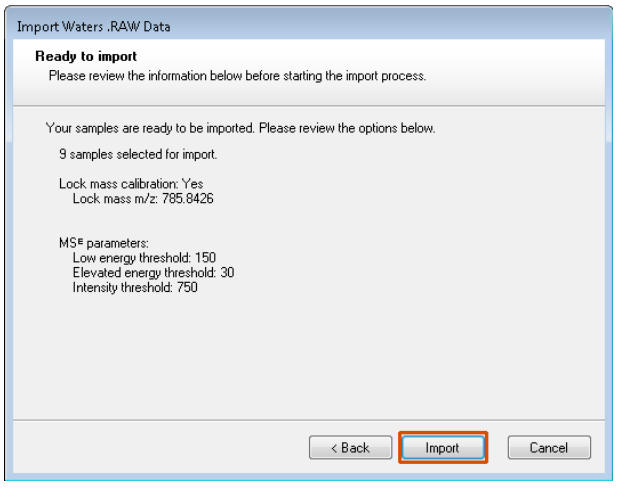
For MSe the settings are 135, 30 and 500.

The default is set to HDMSe.

Note: There is an option to switch off the MS^E identification workflow



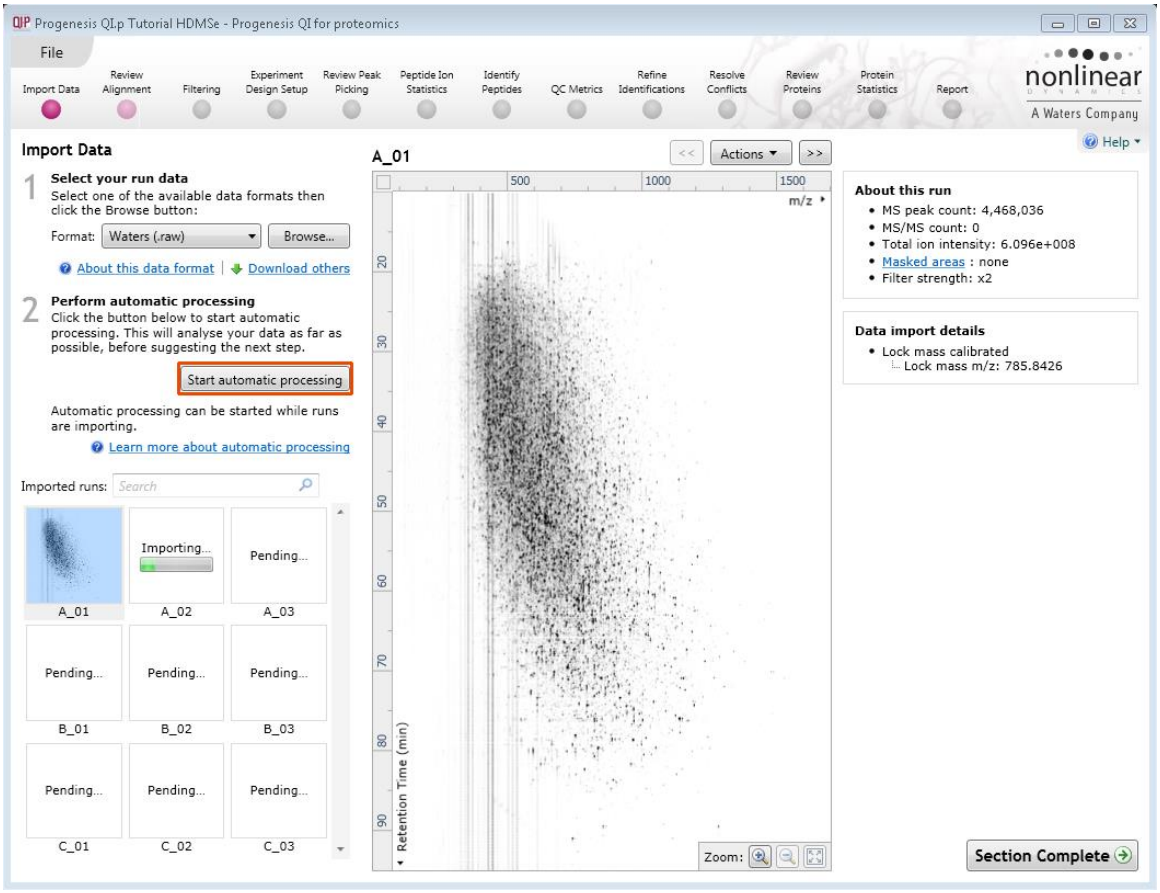
A summary of the loading parameters is provided before you click Import



On loading the selected runs your data set will be automatically examined and the size of each file will be reduced by a 'data modelling routine', which reduces the data by several orders of magnitude but still retains all the relevant quantitation and positional information.

Note: For a large number of files this may take some time.

Note: as the loading process starts you can also start the automatic processing before the loading has completed.

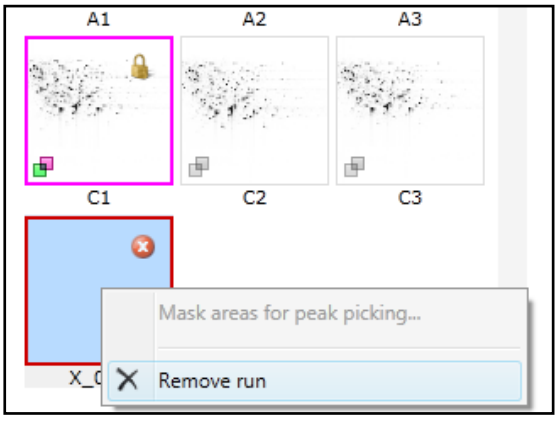


For details of setting up the steps in the automatic processing wizard return to Stage 2A page 7.

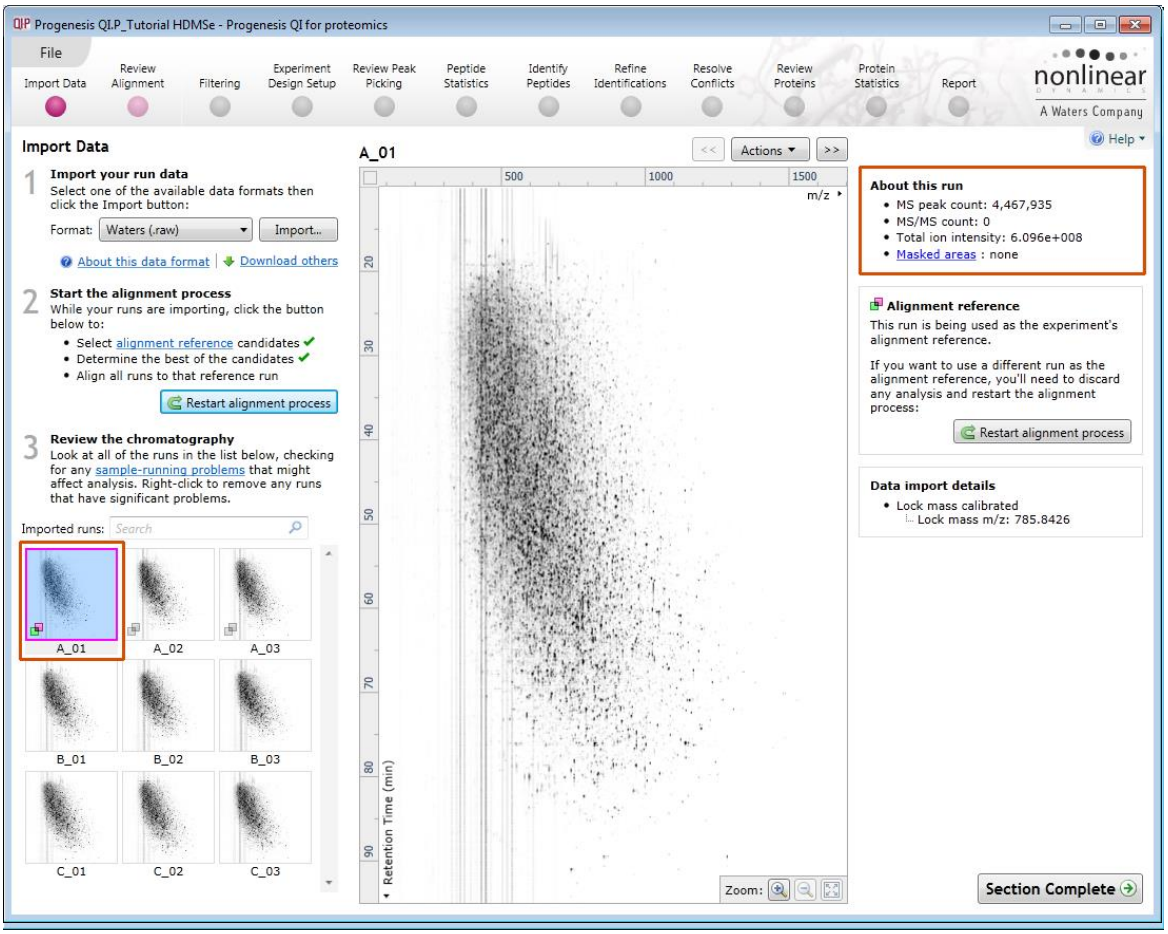
Review Chromatography

Each data file appears as a 2D representation of the run. At this stage you will be warned if any of the data files have been 'centroided' during the data acquisition and conversion process.

You can delete run(s) by left clicking on the run in the list.



At the Import Data stage you can examine the quality of the imported runs using the 2D representation of the runs



Note: details of the current run appear on the top right of the view.

Once you have reviewed the imported runs click on **Review Alignment** or **Section Complete** to move forward to the Review Alignment Stage.

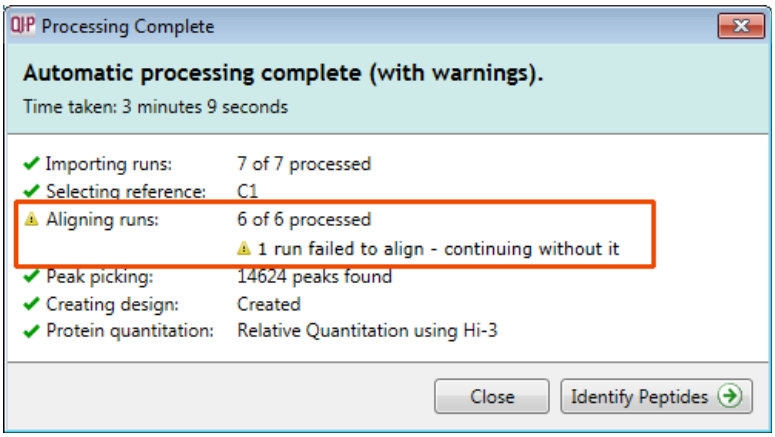
Note: you will be offered the automatic alignment if you have not performed it automatically already.

Now move to the next stage in the workflow (page 7 in this user guide) by clicking **Section Complete.**

Appendix 2: Stage 1 Processing failures

If a stage fails to complete successfully or only partially completes, the automatic processing dialog will warn you of the problem. This may or may not allow the automatic processing to complete.

For example, a run that fails to automatically align will trigger a warning, although analysis will continue; however, the automatic processing dialog will prompt you to 'drop-off' at the **Review Alignment** stage on completion to investigate the problem.

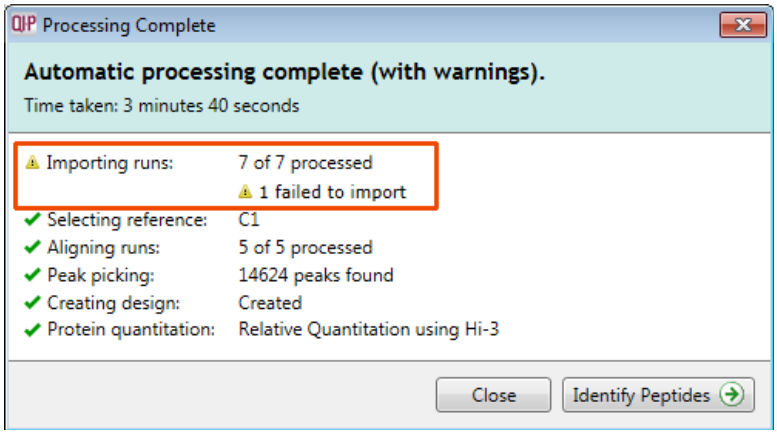


Note: in this example the run that failed to align will not contribute to the peak picking and will be excluded at the alignment stage (a cross appears in the include column).

You can either remove the run from the experiment at the **Import Data** or add it back in at the **Review Alignment** stage once the alignment of the run has been corrected.

As another example, runs that import successfully but with warnings at the **Import Data** stage will cause a flag in the readout to notify you of the potential quality issue.

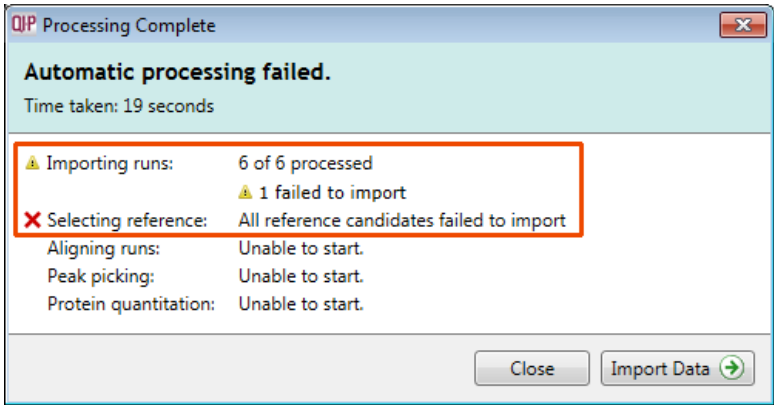
If some runs in a data set fail to import (but not all), the automatic processing will continue informing you that one or more runs have failed to import.



In this case you can remove the runs at Import Data and if appropriate replace them with additional runs.

Note: adding additional runs will then be aligned and peak picking should be re-done to include data from the added runs in the generation of the aggregate.

An example of a problem that would halt the automatic processing would be the failure to successfully import all the potential reference candidates, (for example: while importing, you specified the selection of the alignment reference to be made from several runs before they were fully imported and set the processing underway, and they later failed to import owing to problems with the runs).



In this case, the processing dialog would halt and prompt you to select another reference.

Appendix 3: Licensing runs (Stage 3)

When setting up a **New experiment** if you are evaluating Progenesis QI for proteomics with unlicensed runs then the licensing page will open after **Import Data section**.

If you already have a programmed dongle attached to your machine then the License Runs page will not appear.

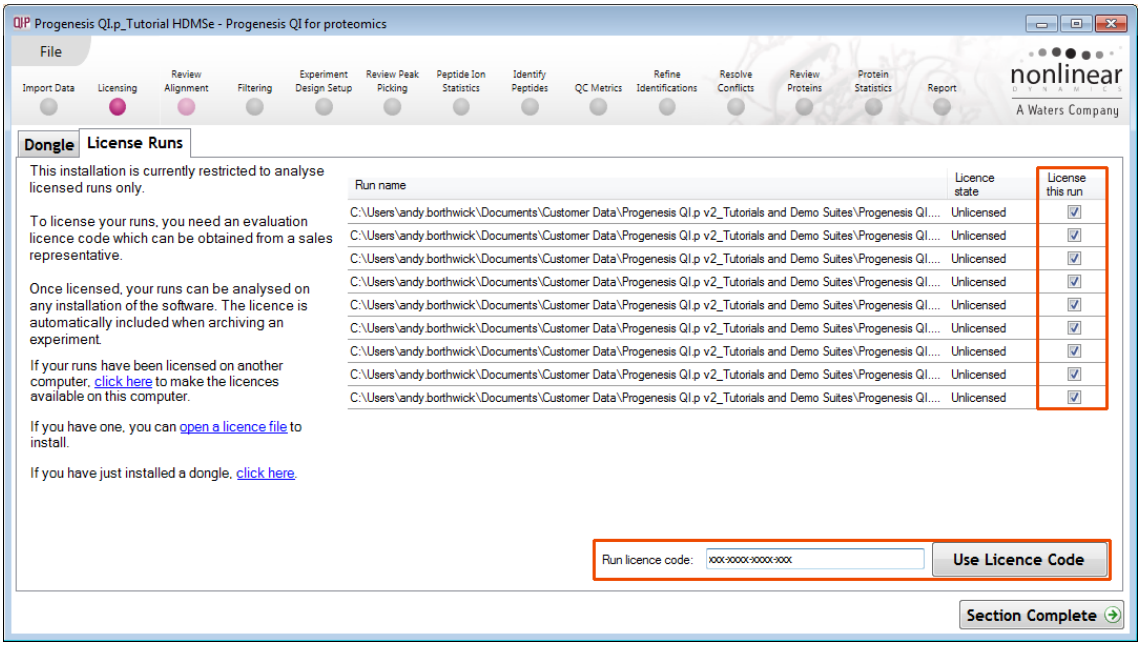
To use this page to License your Runs you must first either obtain an **'Evaluation' Licence Code from a Sales Person or purchase a licence code directly**.

Each code will allow you to license a set number of runs.

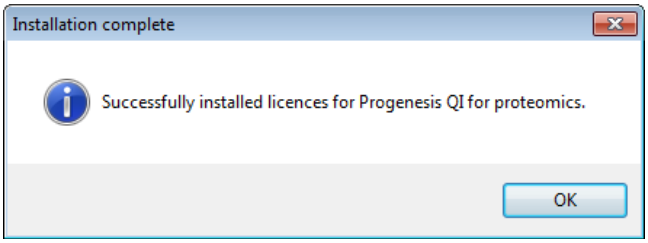
The runs in your experiment will be listed as shown below.

To activate license(s) for the selected runs enter the code in the space provided and click **Use Licence code**.

Note: you will need an internet connection to use this method.



A message confirming successful installation of your licences will appear.

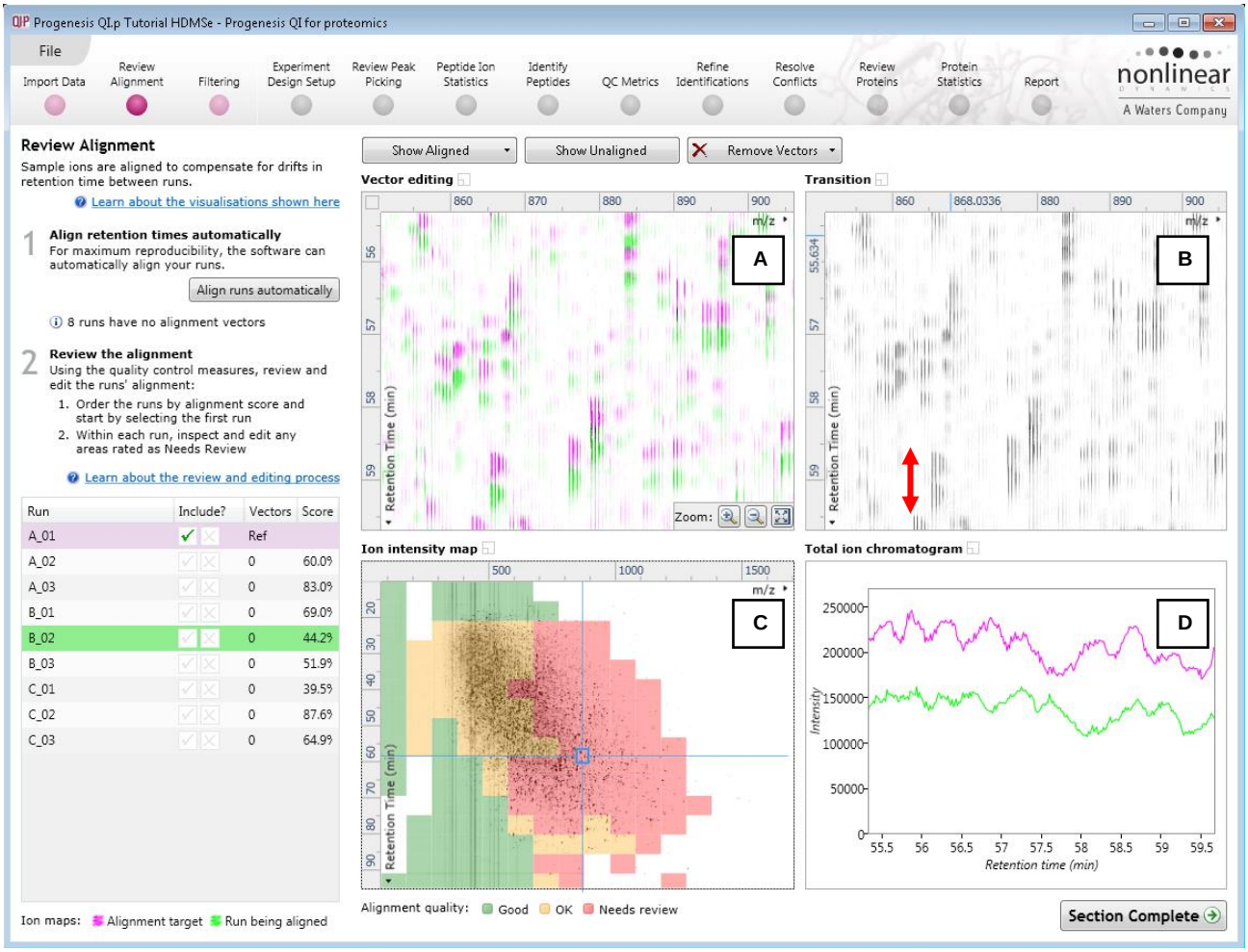


Click **OK**, the view will update and Alignment, the next stage in the workflow, will open with the licensed files.

Appendix 4: Manual assistance of Alignment

Approach to alignment

To place manual alignment vectors on a run (B_02 in this example):



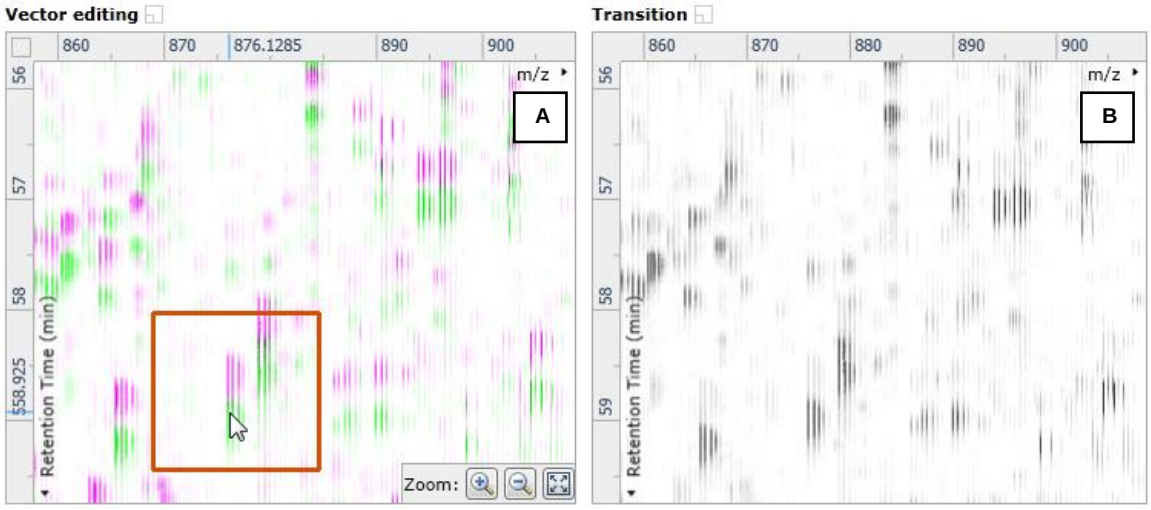
1. Click on Run B_02 in the **Runs** panel, this will be highlighted in green and the reference run (A_01) will be highlighted in magenta.
2. You will need to place approximately 5 - 10 **alignment vectors** evenly distributed from top to bottom of the whole run (RT range).
3. First drag out an area on the **Ion Intensity Map** (C), this will reset the other 3 windows to display the same 'zoomed' area

Note: the peptide ions moving back and forwards between the 2 runs in the **Transition** window (B) indicates the misalignment of the two runs.

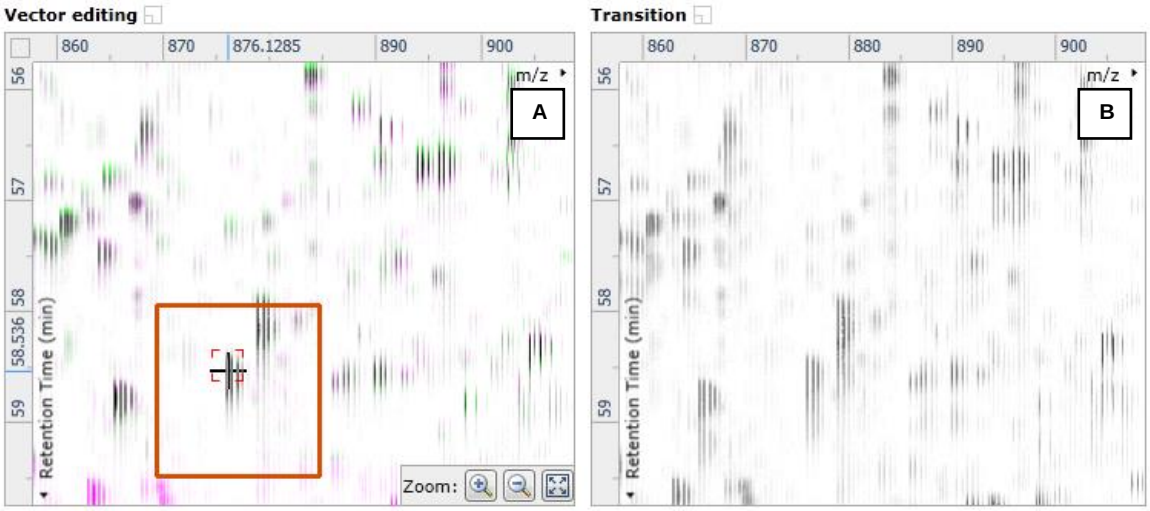
Note: the Ion Intensity Map gives you a colour metric, visually scoring the current alignment and an overall score is placed next to the Vectors column in the table. With each additional vector added this score will update to reflect the 'changing' overall quality of the alignment. The colour coding on the Ion intensity Map will also update with each additional vector.

Note: The **Total Ion Chromatograms** window (D) also reflects the misalignment of the 2 runs for the current Retention Time range (vertical dimension of the current Focus grid in the **Ion Intensity Map** window).

4. Click and hold on a green peptide ion in Window A as shown below.

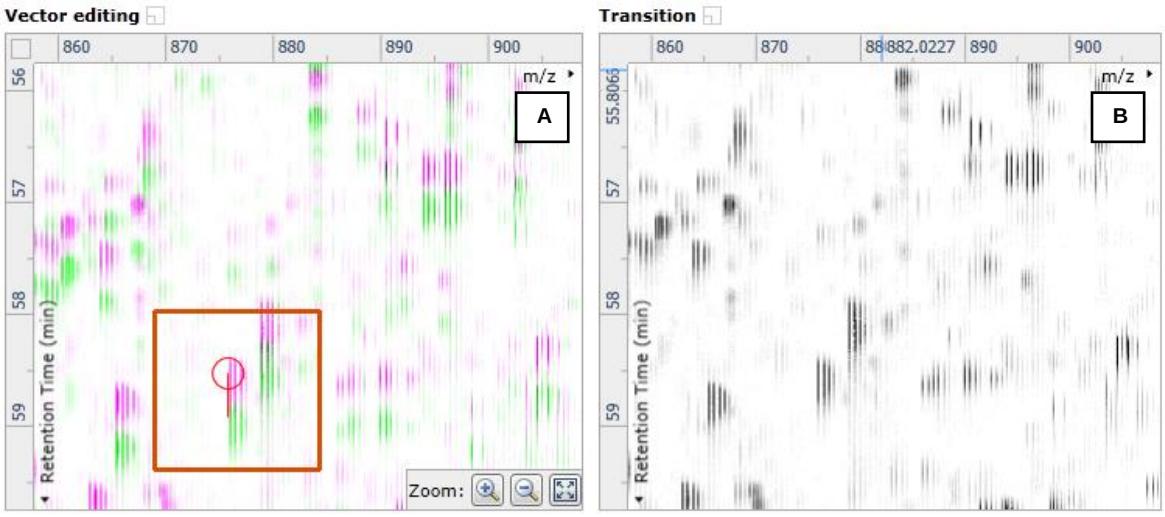


5. As you are holding down the left mouse button (depending on the severity of the misalignment), the alignment vector will automatically find the correct lock. If not, drag the green peptide ion over the corresponding magenta peptide ion of the reference run. The red box will appear as shown below indicating that a positional lock has been found for the overlapping peptide ions.



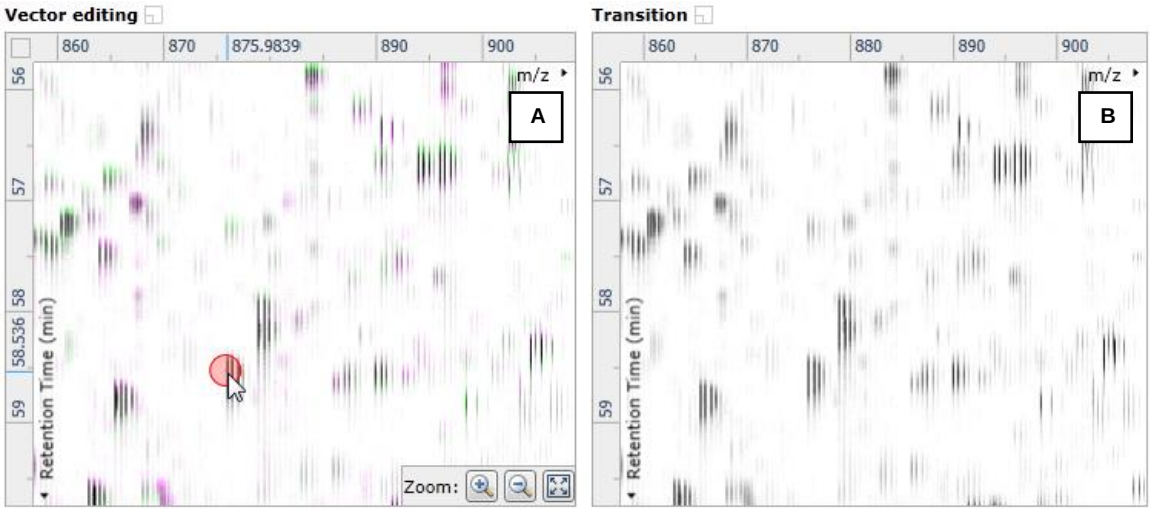
***Tip*:** while holding down the mouse button hold down the **Alt** key. This will allow smooth movement of the cursor as the **Alt** key allows you to override the 'automatic alignment' performed as you depress the mouse button.

6. On releasing the left mouse button the view will 'bounce' back and a red vector, starting in the green peptide ion and finishing in the magenta peptide ion will appear.

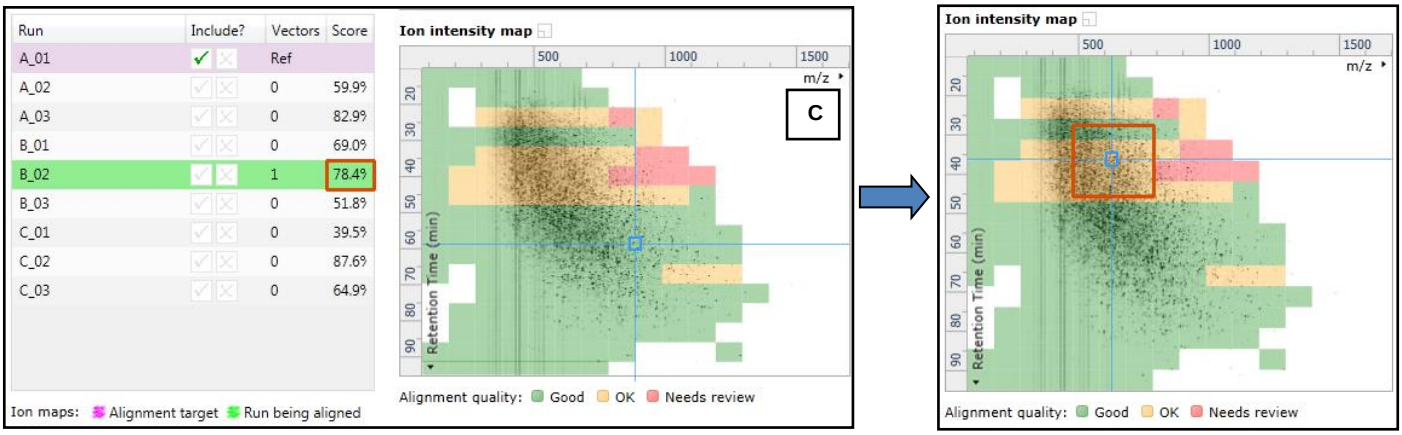


Note: an incorrectly placed vector is removed by right clicking on it in the **Vector Editing** window and selecting delete vector.

7. Now click **Show Aligned** on the top tool bar to see the effect of adding a single vector.



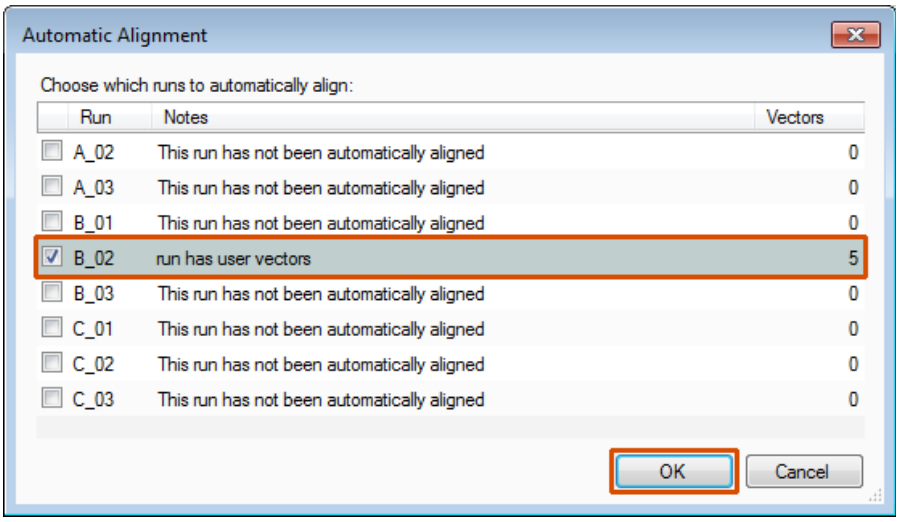
8. With the placement of a single manual vector the increase in the proportion of the **Ion Intensity Map** (C) showing green is reflected in the improved alignment score in the table. Now click in the Ion Intensity Map to relocate the focus in order to place the next manual vector.



9. Adding an additional vector will improve the alignment further as shown below.

10. The shift in the Retention Time (RT) is as a result of incorrect running of the chromatography. In many of these cases if the Automatic Alignment fails to generate a good alignment then removing all the alignment vectors for this run and placing a single manual vector to act as a 'seed' for the Automatic Alignment algorithm maybe all that is required to generate a good alignment.
11. In the case of the example shown above placing a small number of vectors from the top to the bottom of the run is sufficient to markedly improve the alignment.

12. At this point you would redo the automatic alignment of this image by selecting automatic alignment. **Note**: if you are focusing only on the alignment of one run, then un-tick the other runs in the alignment dialog.

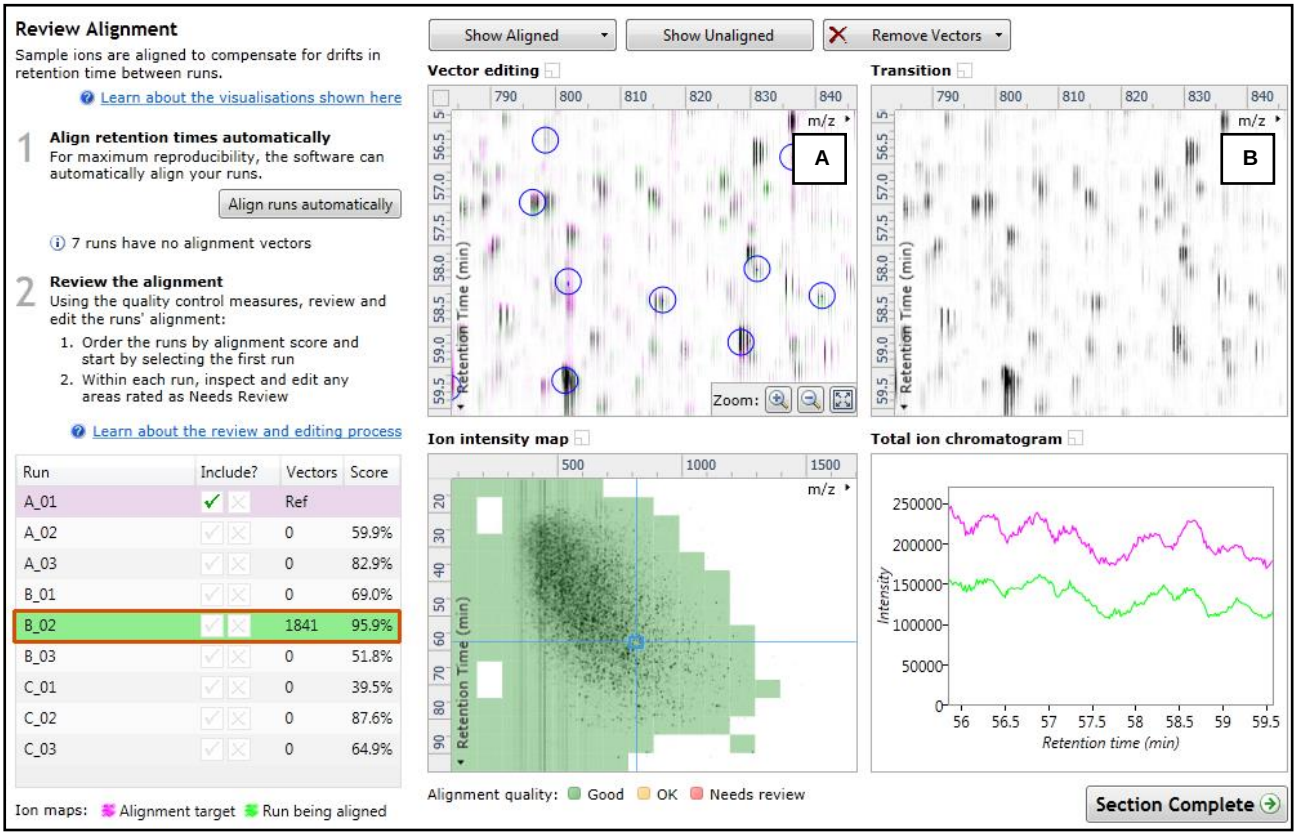


13. On pressing OK the Automatic Alignment will run for the selected run. On completion the table and views will update to display the automatically generated vectors (shown in blue).

14. Repeat this process for all the runs to be aligned.

The number of manual vectors that you add at this stage is dependent on the misalignment between the current run and the Reference run.

Note: In many cases only using the Automatic vector wizard will achieve the alignment.



Tip: a normal alignment strategy would be: to run the automatic alignment first for all runs, then order the alignments based on score. For low scoring alignments remove all the vectors and place 1 to 5 manual vectors to increase the score then perform automatic alignment. Then review the improved alignment score.

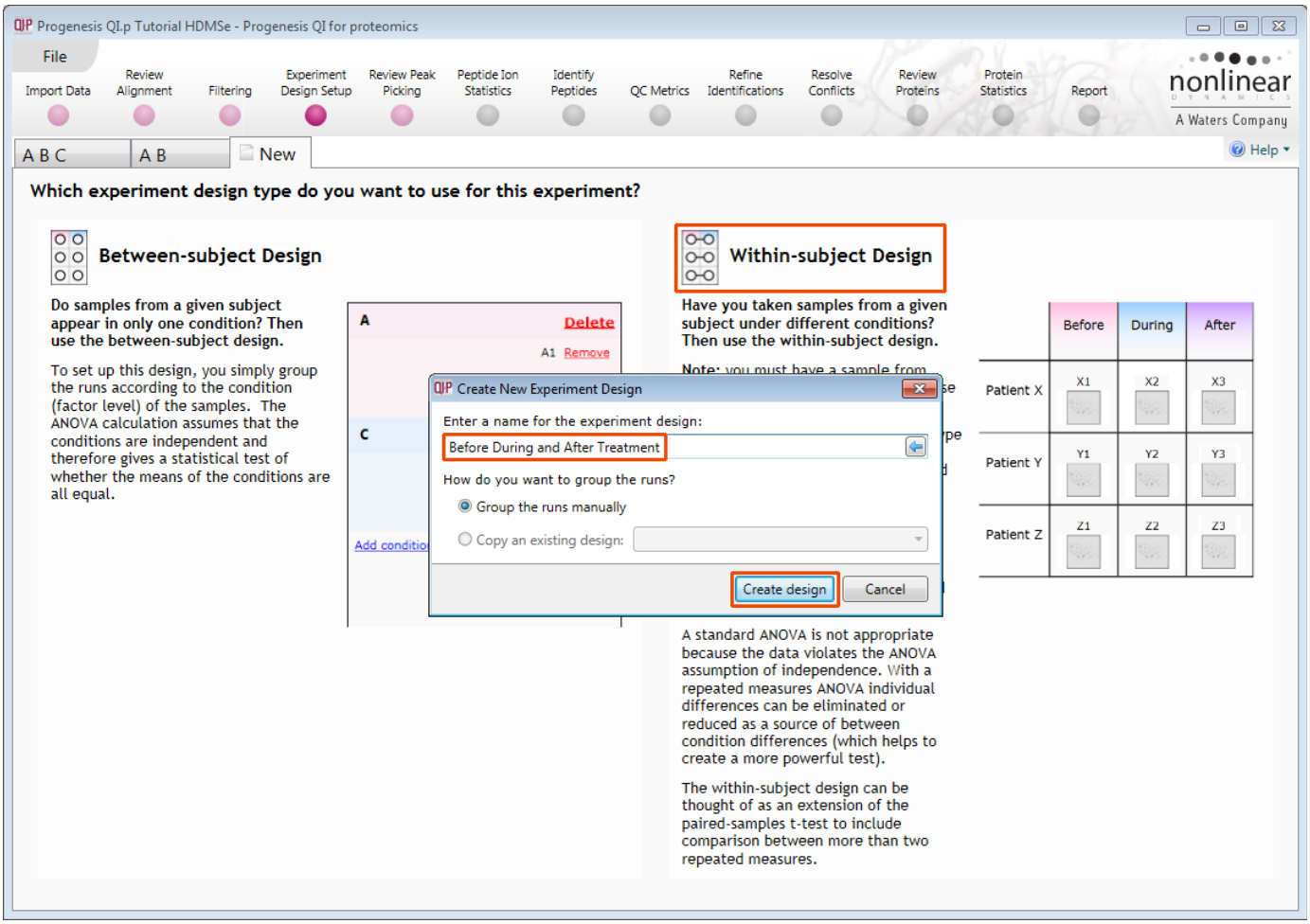
Also the 'ease' of addition of vectors is dependent on the actual differences between the LC-MS runs being aligned.

To review the vectors, automatic and manual, return to page 18

Appendix 5: Within-subject Design

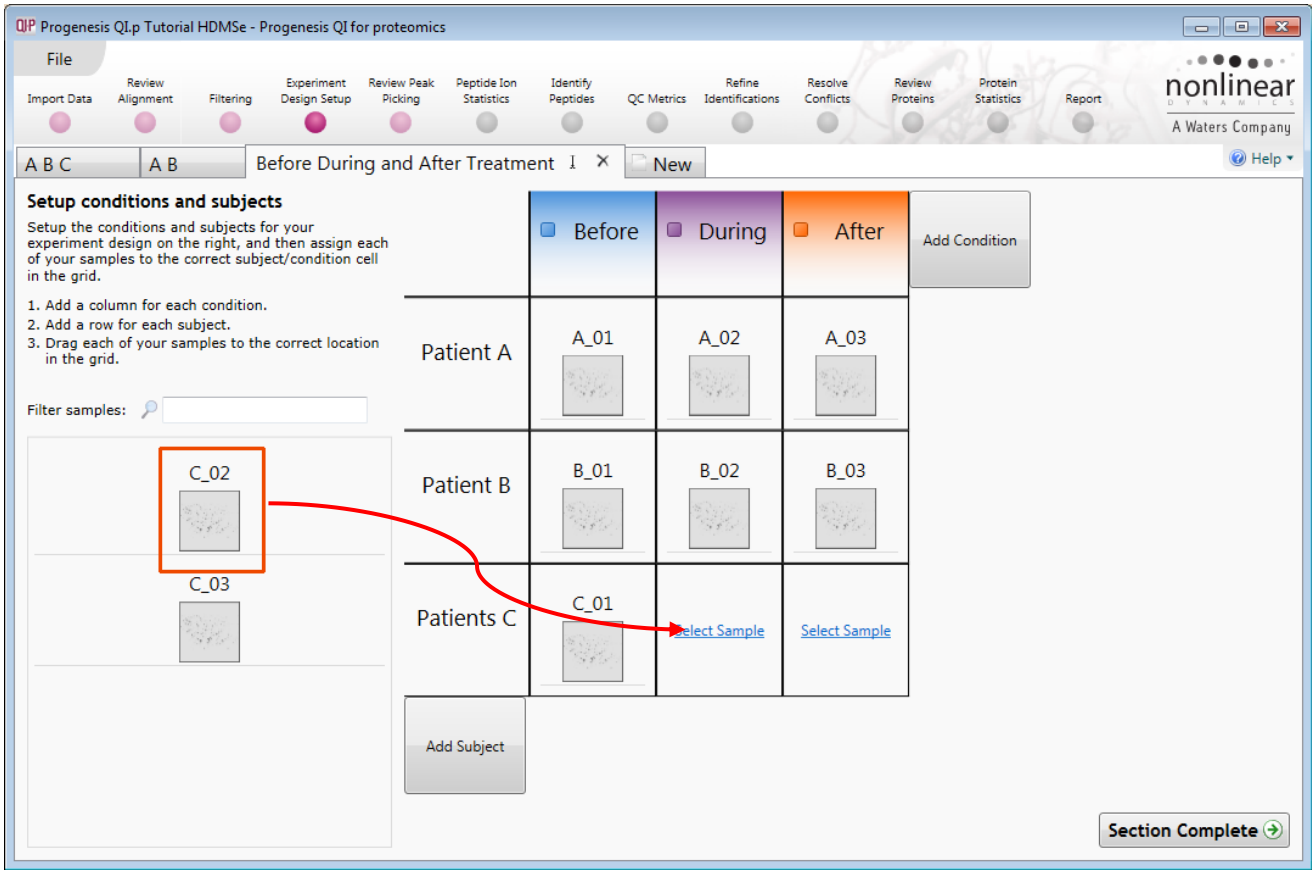
To create a **Within-subject Design** for your data set select this option on the **Experiment Design Setup** page and enter the name of the design.

In this example there are 3 Subjects (i.e. patients A, B and C) who have been individually sampled: Before(1), During (2) and After (3) treatment



When the design page opens use the **Add Subject** and **Add Condition** buttons to create the matrix that fits your experimental design, over typing the names as required.

Then Drag and drop the Samples on to the correct 'cell' of the matrix.



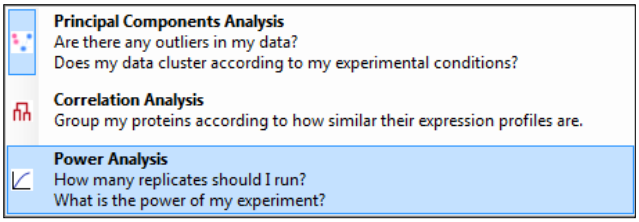
You can create additional Experimental Designs using the New tab

All of these Experimental Designs are available at all the following stages in the Progenesis QI for proteomics workflow.

Appendix 6: Power Analysis (Peptide Stats)

Power analysis is a statistical technique that is used to gauge how many replicates are needed to reliably see expression differences in your data. It is available through the Peptide Stats section of the workflow.

To perform a power analysis of the data click on **Ask another question** at the top of the table in the Peptide Stats screen. A selection of 3 tools will appear in the form of questions.



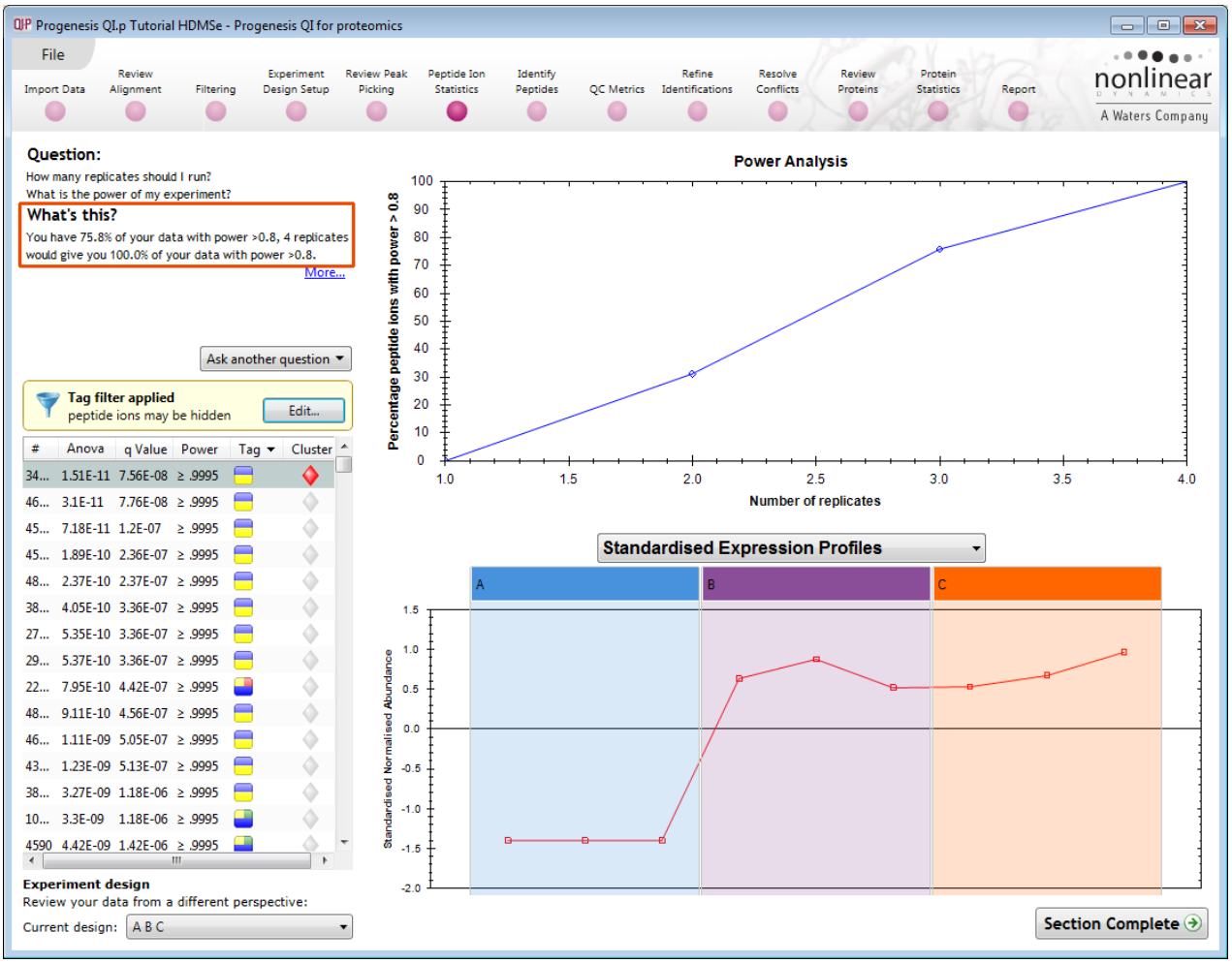
Select the option

'How many replicates should I run and what is the power of my experiment?'

It answers this question by informing you:

'How many replicates you need so that at least 80% of your peptide ions with a power >0.8'

Using the **Significant p<0.05 peptide ions (21038),** as an example, view the power analysis.



This is displayed graphically showing that 75.8% of the 21083 peptide ions have a power of 80% or that 4 replicates would give you 100% of your data with power > 0.8.

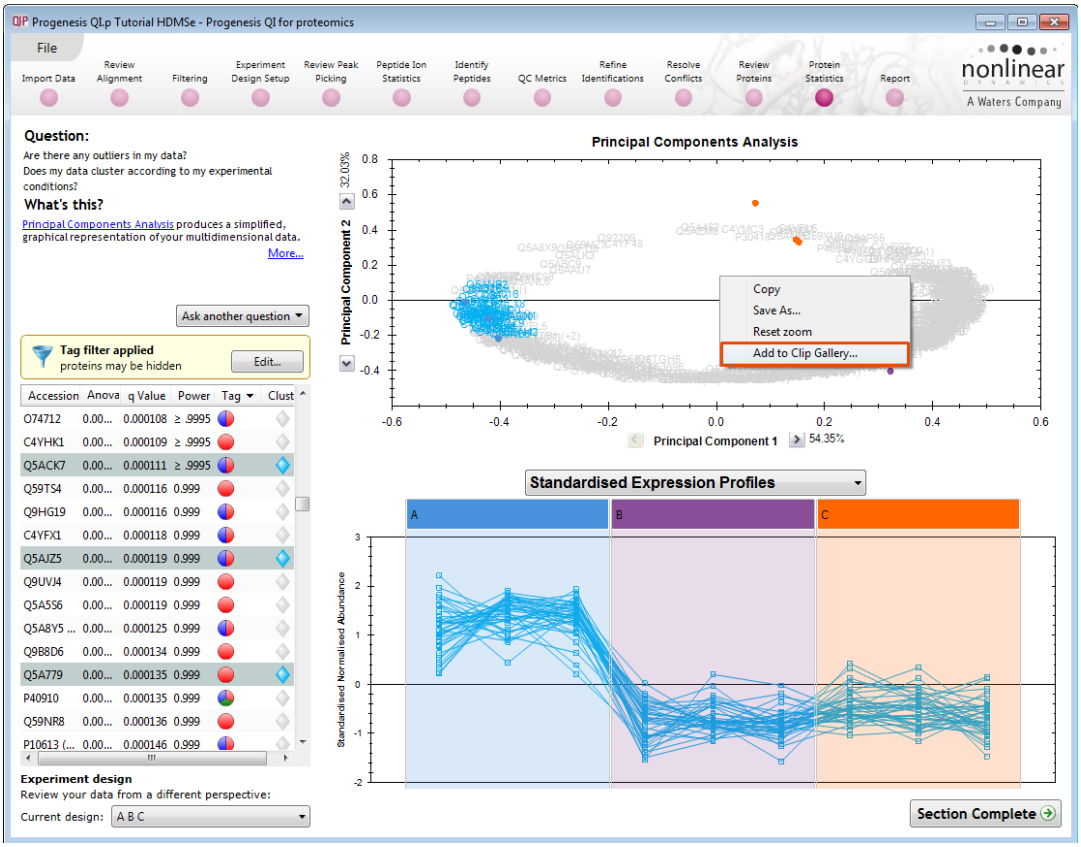
- **The power of a statistical test** reflects our confidence in the experimental data's ability to find the differences that do actually exist
- The power is expressed as a percentage, where 80% power is an accepted level, therefore allowing you to assess the number of sample replicates that would be required to achieve a power of 80%.

Appendix 7: Using Clip Gallery to Save and Export Pictures and Data

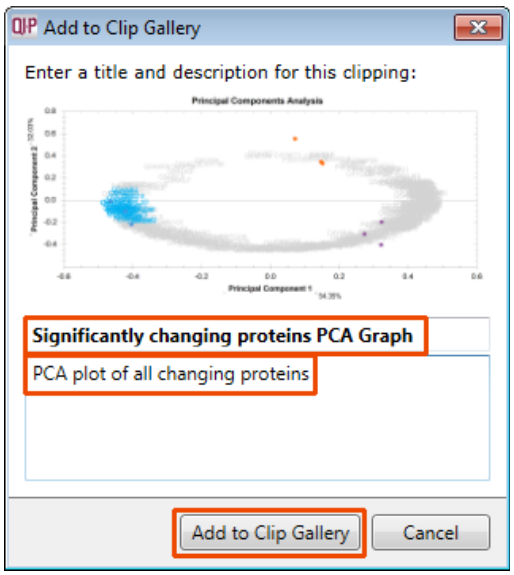
At every stage of the Progenesis QI for proteomics workflow the views and data tables can be added to the Clip Gallery.

The saved images of the Views and the tables are retained as part of the experiment and are stored accordingly. This facility allows you to capture (high resolution) images that can be used in the development of specific reports and/or used as part of the process of publishing your experimental findings.

As an example of using the Clip Gallery, at the **Protein Statistics** view while displaying the PCA plot right click on the **Biplot** View and select **Add to clip gallery…**

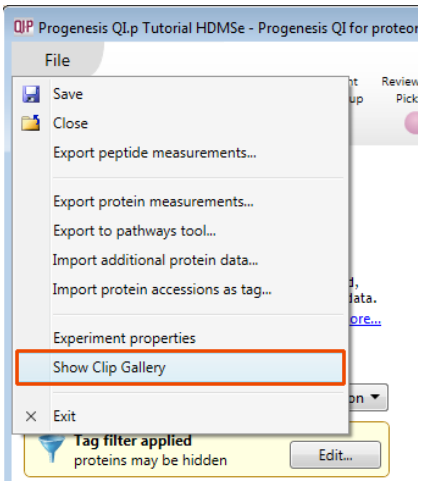


This will open a dialog displaying what is to be saved and allows you to alter the title and provides a description of the image for later reference.

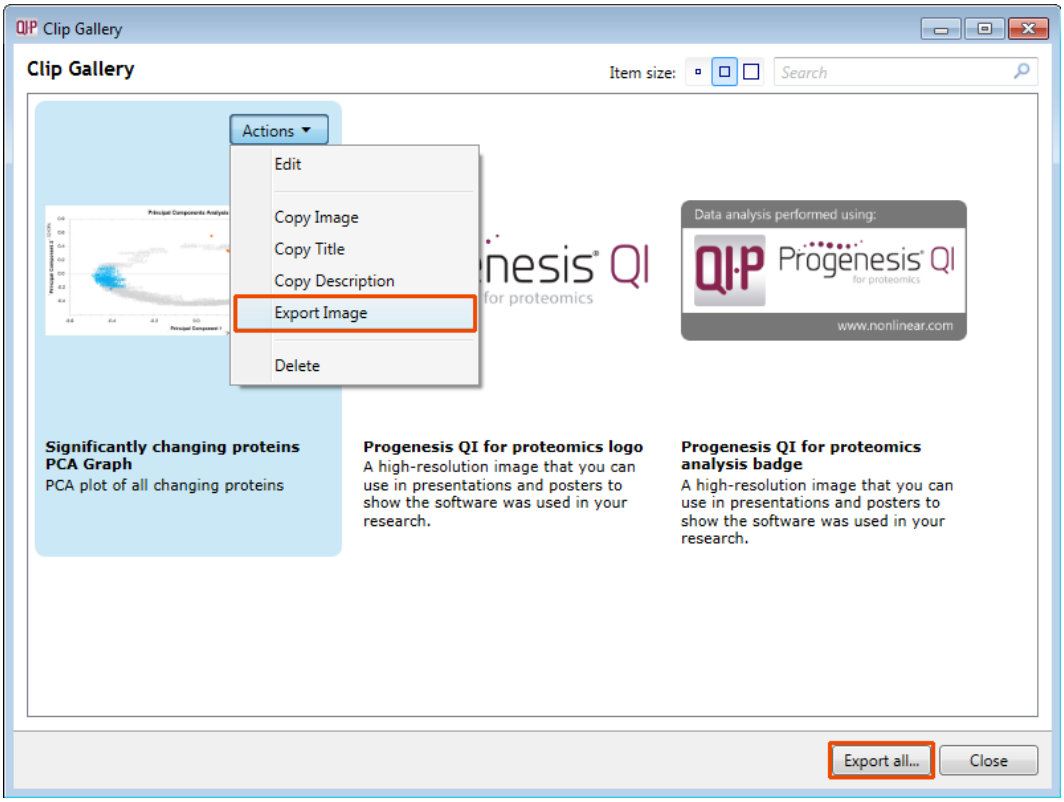


Enter details as required and click **Add to clip gallery**

To view, edit and/or export from the clip galley the gallery can be accessed from the **File** menu.



Selecting an image in the gallery makes available an **Actions** menu that allows you to manipulate the output of the image.



Note: there is also the capacity to **Export all…** the images in the experiments clip gallery which creates a list of files in a folder of your choice where the file name is based on the image title.

Note: right clicking on a table and adding it to the Clip Gallery allows you to export the current content to Excel.

Appendix 8: Waters Machine Specification

This appendix provides information on the approximate time(s) taken at each stage and the total time taken to analyse a set of 9 (Phase 1) HDMSe runs on a Waters Demo Spec PC.

Machine Spec: Lenovo

Processor: Intel® Xeon® CPU 2.66GHz 12core X5650 @ 2.67GHz

RAM: 24.0 GB

System Type: 64-bit Operating System

File Folder Size: Each file folder (.RAW): 40.9 Gig

Analysis Stages:		**Per file**	**Total**	
Import Data:	Loading of Raw data per file	12min	2hr 6min	Total for 9 files
	Apex Background processing	1hr 05min(max)	9hr 45min	Total for 9 files
	(re-opening at Import Data)		20s	
Alignment:	Automatic alignment of data		8min 30s	
	(re-opening at Alignment)		10s	
Peak Detection:	Automatic Detection of data		14min 15s	
	(re-opening at Peak Detection)		10s	
Identify Peptides:	Performing MS^E Search		14min 30s	
	(re-opening at Identify Peptides)		10s	

Total Analysis Time	Excluding Background Apex Processing	2hr 42min
	Including Apex processing assuming pause for Apex	9hr 58min

Restoring	Progenesis QI.p Tutorial HDMSe.ProgenesisQIPArchive	4min

References

Silva, JC, Gorenstein, MV, Li, G-Z, Vissers, JPC and Geromanos. Molecular and Cellular Proteomics (2006); 5 : 144-156 Absolute Quantification of Proteins by $LCMS^E$