

Progenesis QI for proteomics User Guide

Analysis workflow guidelines for DDA data

THE SCIENCE OF WHAT'S POSSIBLE.

Contents

Introduction	3
How to use this document	3
How can I analyse my own runs using Progenesis QI for proteomics?	3
LC-MS Data used in this user guide	3
Workflow approach to LC-MS run analysis	4
Restoring the Tutorial	5
Stage 1: Import Data and QC review of LC-MS data set	6
Stage 2A: Automatic Processing of your data	7
Stage 2B: After Automatic Processing	12
Stage 3: Licensing	13
Stage 4: Review Alignment	14
Reviewing Quality of Alignment	15
Stage 5A: Filtering	18
Stage 5B: Reviewing Normalisation	22
Stage 6: Experiment Design Setup for Analysed Runs	25
Stage 7: Review Peak Picking and editing of results	28
Stage 8: Peptide Ion Statistics on selected peptide ions	37
Stage 9: Identify peptides	41
Stage 10: QC Metrics	45
Stage 11: Refine Identifications	47
Resolving Conflicts	49
Protein Grouping	49
Protein Quantitation options	51
Stage 12: Review Proteins	53
Stage 13: Exporting Protein Data	60
Exporting Protein Data to Pathways Tool(s)	61
Stage 14: Exporting identifications for submission to PRIDE	64
Stage 15: Protein Statistics	66
Stage 16: Reporting	67
Creating an Inclusion list	69
Appendix 1: Stage 1 Data Import and QC review of LC-MS data set	71
Appendix 2: Stage 1 Processing failures	75
Appendix 3: Stage 1 Data QC review and addition of exclusion areas	77
Appendix 4: Licensing runs (Stage 3)	78
Appendix 5: Manual assistance of Alignment	79
Appendix 6: Within-subject Design	83
Appendix 7: Power Analysis (Peptide Ion Stats)	85
Appendix 8: Resolve Conflicts	86
Appendix 9: Using Clip Gallery to Save and Export Pictures and Data	90
Appendix 10: Exporting protein identities to IPA	92
Appendix 11A Search engine parameters for Mascot	95
Appendix 11B Use Additional Export Fields in Mascot for PRIDE	96
Referenœs	96

Introduction

This user guide takes you through a complete analysis of 6 LC-MS runs with 2 groups (3 replicate runs per group) using the unique Progenesis QI for proteomics workflow. It starts with LC-MS data file loading then Alignment, followed by Analysis 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 Reporting.

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 71) then start at page 7.

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 (dependant on PC spec) and is divided into two sections. 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. Alternatively if you would like to arrange a demonstration in your own laboratory contact <u>support@nonlinear.com</u> and we will help you.

LC-MS Data used in this user guide

NLD would like to thank Dr Robert Parker and Prof Haroun Shah at the Health Protection Agency, London, UK for providing the example data used in this user guide as well as invaluable discussion on the handling of the data.

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.

	Review		Experiment	Review Peak	Peptide Ion	Identify		Refine	Review	Protein	
Import Data	Alignment	Filtering	Design Setup	Picking	Statistics	Peptides	QC Metrics	Identifications	Proteins	Statistics	Report
	-	0		0	-	0			-	NAU	Callor-

Stage	Description	Page
Import Data	LC-MS Import Data: Selection and review of data files for analysis	6
	Automatic Processing: setting up steps for automatic processing	7
Ļ	After Automatic Processing: how to work with auto analysed data	12
Licensing	Licensing : allows licensing of individual data files when there is no dongle attached (Appendix 4 (page 78))	13
Review Alignment	Review Alignment: automatic and manual LC-MS run alignment	14
Filtering	Filtering : defining filters for peaks based on Retention Time, m/z, Charge State and Number of Isotopes.	18
	Review Normalisation: exploring LC-MS normalisation	22
Experiment Design Setup	Experiment Design Setup : defining one or more group set ups for analysed aligned runs	25
Review Peak Picking	Review Peak Picking : review and validate results, edit peak detection, tag groups of peaks and select peaks for further analysis	28
Peptide Ion Statistics	Peptide Ion Statistics : performing multivariate statistical analysis on tagged and selected groups of peptide ions	37
Identify Peptides	Identify Peptides : managing export of MS/MS spectra to, and import of peptide ids from Peptide Search engines.	41
QC Metrics	QC metrics: quality control charts for experimental/analysed data	45
Refine	Refine Identifications: manage filtering of peptide ids	47
Identifications	Resolving Conflicts : validation and resolution of peptide id conflicts for data entered from Database Search engines	49
Review Proteins	Review proteins : review protein and peptide identity and data export	53
Protein Statistics	Protein Statistics: multivariate statistical analysis on proteins	66
Report	Report : generate a report for proteins and/or peptides	67

Waters

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 you can restore the uncompressed Tutorial archive file. To do this, first locate the **Progenesis QI.P_Tutorial DDA.Progenesis QIP Archive** file using the **Open** button and press Open.

File		nonlinea
•		A Waters Compar
rform analysis Combine analysed fracti	ons Search	New to Progenesis QJ for proteomics? Here are some resources to help you et started with Progenesis QJ for
QIP Open Experiment		 proteomics: <u>The Progenesis OI for proteomics</u> workflow
 ShareFile BitTorrent Sync Desktop 	▲ Name Date m Progenesis QLp Tutorial for DDA.Progen 18/04/2	• Frequently-asked questions
ibraries Documents Just Music Eifertures Videos		running by zooming to a known ion?
Andy Borthwick Computer Ketwork ASSESVA	~ <	R 1
File name: Progenesis QLp	Tutorial for DDA.Progenesis(Experiments and Archives (*.Pr. Open Cancel	Jump to a specific m/z and RT using the <u>Go To Location tool</u> in the top-left corner of the ion maps.
ther experiments	3.0.	Latest blog posts Missing values: the Progenesis co- detection solution Identification scoring in Progenesis OI S978.29

This opens the 'Import Experiment from archive' dialog.

Select the **Create a new experiment** option and select the folder in which you placed the archive, using Browse.

IP Import Experiment from Archive							
Import experiment from archive After importing the experiment from this archive, any changes to the experiment will be saved to the location below, not back to the archive.							
Replace an existing experiment							
Experiment to repla	ace: Progenesis QI.p Tutorial for DDA 🔹						
Create a new expe							
Experiment name:	Progenesis QI.p Tutorial for DDA						
Save to folder: and Demo Suites\Progenesis QLp Tutorial for DDA Browse							
Import							

Then	press	Import.
mon	p1000	import.

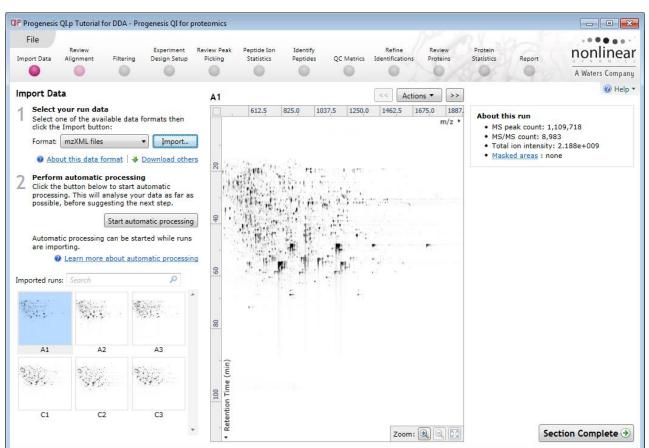
Loading: Progenesis QI.P_Tutorial DDA						

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.

Note: the **Experiment Properties** are available from the File menu. These were selected when the experiment was created (see Appendix 1, page 71). The tutorial data is profile data.

Tip: the **Mask areas for peak picking'** 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 (see Appendix 3, page 77). This is not required for this data set.

<<	Act	ions 🔹 🔁 >>		
1500		Mask areas for	peak picking	
	×	Remove run	Delete	00
	_		 Lotal ion intensity: 2.1 <u>Masked areas</u>: none 	886+009

OPP Experiment Properties

Runs in this experiment: 6 Machine resolution: High resolution Peak processing: Profile data

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

Now start the Automatic Processing.

×

Close

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:

- Automatic alignment of all runs to a reference run
- Automatic peak picking
- Creating an Experiment design
- Choosing a Quantitation method
- Identification of peptides (only available for MSe and HDMSe data formats)

In this tutorial example you have 6 Data Dependently Acquired (DDA) LC-MS runs, so the automatic identification of peptides is unavailable in the automatic processing.

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

QP Start automatic processing	- • • ×					
Select an alignment reference To compensate for drifts in retention time, all runs in the e to a single reference run.	experiment must be aligned					
How do you want to choose your alignment reference?						
Assess all runs in the experiment for suitability						
\odot Use the most suitable run from candidates that I select						
🔘 Use this run:						
A1 *						
For information on choosing the alignment reference, and why you might want to select your own candidates, please see the <u>online guidance</u> .						
< Back	Next > Cancel					

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 selects 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 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 first option, (See Appendix 1, page 71 for more details on using the other options).

You will now be asked if you want to align your runs automatically.

QP Start automatic processing								
Automatic alignment After selecting the experiment's alignment reference, the software can also automatically align all runs.								
After the alignment reference is chosen, do you want to start automatic alignment? Yes, automatically align my runs								
< Back Next > Cancel								

The default is for automatic alignment, click Next.

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

QIP Start automatic processing	- • ×					
Peak picking Peak picking is the process by which we locate the peptide ions and th peaks in your samples.	neir isotopic					
After the automatic alignment is finished, do you want to start automatic peak picking? Perform peak picking Set parameters						
< Back Next >	Cancel					

For the purposes of this User guide we will use the default settings for peak picking except in the case of Retention Time limits where we will limit the peak picking to between 10 and 75 min.

Note: for more details on setting Peak Picking parameters refer to the section on Filtering (page 18) 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 10 and 75 min and tick the boxes as shown below.

QP Peak Picking Parameters		a	Peak Picking Paramete	ers			X
Runs for peak picking Peak picking lim	its Maximum charge Retention time limits		Runs for peak picking	Peak picking limi	ts Maximum charge	Retention tim	e limits
Choose runs for peak picking			Retention time li	mits			
You can tick or un-tick each run to control which will be used by the peak picking algorithm. Although any run which is left un-ticked will not affect the peptide ion outlines, it will still have outlines added to it and will be available in the experiment design setup. Learn more about why you might not want to select all runs.	 ✓ Run ✓ A1 ✓ A2 ✓ A3 ✓ C1 ✓ C2 ✓ C3 		You can set the minim maximum retention ti picking. Ions that elut after these values will	num and me for peak te before or	☑ Ignore ions before ☑ Ignore ions after		minutes minutes

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.

QP Start automatic processing	×
Experiment design Experiment designs allow you to group and compare your samples according to the experimental conditions.	eir
By defining an experiment design in advance, statistical measures such as ANOVA can t calculated automatically.	e
Set up an experiment design	
Enter a name for the experiment design:	
AC	~
Load the criteria for grouping runs from this file:	
Brow	se
Group runs by: <pre></pre>	-
< Back Next > Cancel	

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 **Tutorial Groups.csv** file available with the Tutorial Archive you restored at the beginning of this tutorial exercise.

Sample Name	Conditions	Date of Collection	Location
A1	А	02/03/2011	Fridge A
A2	А	02/03/2011	Fridge A
A3	А	02/03/2011	Fridge B
C1	С	06/03/2011	Fridge B
C2	С	06/03/2011	Fridge B
C3	С	06/03/2011	Fridge A

Give the experiment design a name (i.e. AC) and then use the **Browse** function to locate the Tutorial Groups.csv file.

The 'Group runs by' drop down will update to reflect the possible fields in the csv file that you can use to 'group' your runs by: Conditions, Date of collection or Location.

OP Start automatic pro	ocessing
Experiment design Experiment designs experimental condit	allow you to group and compare your samples according to their
By defining an experi calculated automatic	ment design in advance, statistical measures such as ANOVA can be ally.
🔽 Set up an exper	iment design
Enter a name for	r the experiment design:
AC	(m)
Load the <u>criteria</u>	for grouping runs from this file:
C:\Users\andy.b	oorthwick\Documents\Customer Data\Progenesis QI.p v2_1 Browse
Group runs by:	Conditions 🔹
	Conditions
	Date of Collection
	Location
	< Back Next > Cancel

Select Conditions and then click Next.

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

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.

OP Start automatic processing	OP Start automatic processing
Protein Quantitation If you've included a known amount of a calibrant protein in each of your samples, you can calculate the absolute amounts of each identified protein.	Protein Quantitation If you've included a known amount of a calibrant protein in each of your samples, you can calculate the absolute amounts of each identified protein.
Quantitation method:	Quantitation method:
Relative Quantitation using Hi-N	Relative Quantitation using Hi-N
Number of peptides to measure per protein (N):	Absolute Quantitation using Hi-N • Requires a calibrant protein to calculate absolute amounts • Uses the most abundant N peptides • Allows comparison between proteins within a run
	Relative Quantitation using Hi-N Uses the most abundant N peptides Allows comparison between proteins within a run
	Relative Quantitation using non-conflicting peptides Uses only peptides which have no conflicting protein identifications Allows comparison of a single protein across runs
✓ Use protein grouping i.e. hide proteins whose peptides are a subset of another protein's.	Relative Quantitation using all peptides • Uses all peptides identified as part of a protein • Allows comparison of a single protein across runs protein's.
< Back Finish Cancel	< Back Finish Cancel

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

Select the Default option Relative Quantitation using Hi-N, with Use protein grouping ticked then click Finish.

The Alignment process starts with the automatic selection of C1 as the alignment reference

Progenesis QI for proteomics User Guide

QIP Automatic Processing Automatic processi Current step: Choosing a	ing		$\overline{\}$	7		
 Importing runs: Selecting reference: Aligning runs: Peak picking: 	6 of 6 processed Choosing an alignment reference Pending Pending			QP Automatic Processing Automatic processi Current step: Aligning 'A:	ing	
Creating design: Protein quantitation:	Pending Pending	Cancel		 Importing runs: Selecting reference: Aligning runs: Peak picking: Creating design: Protein quantitation: 	6 of 6 processed C1 2 of 5 processed Pending Pending Pending	

Once Alignment completes Peak Picking commences

QIP Automatic Processing (62%)	
Automatic processing Current step: Analysing	
✓ Importing runs: 6 of 6 processed ✓ Selecting reference: C1 ✓ Aligning runs: 5 of 5 processed Oreak picking: Picking Creating design: Pending	QIP Processing Complete Automatic processing complete. Time taken: 3 minutes 10 seconds
Protein quantitation: Pending Creating design: Pending Cancel	 ✓ Importing runs: 6 of 6 processed ✓ Selecting reference: C1 ✓ Aligning runs: 5 of 5 processed ✓ Peak picking: 14556 peptide ions found ✓ Creating design: Created ✓ Protein quantitation: Relative Quantitation using Hi-3

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

In this example, as the data is DDA it is going to open at **Identify Peptides** if all possible processing steps are selected.

You can either:

- Continue with the analysis, as the Processing dialog is not displaying any warnings, and perform **Identify Peptides**. In which case you can go to page 41
- Open the analysis at Identify Peptides and immediately return to the Review Alignment stage by clicking on it in the Workflow to review alignment quality (page 14)
- Alternatively you can **Close** the dialog. This will **not** move you to a later stage in the workflow but instead will allow you to review the 'Run Chromatography' then navigate systematically through all the stages yourself or jump from stage to stage.

File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Ion Statistics	Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report	nonlinear
Identify P	eptides		-	MS/MS S	5pectra	•						A Waters Company

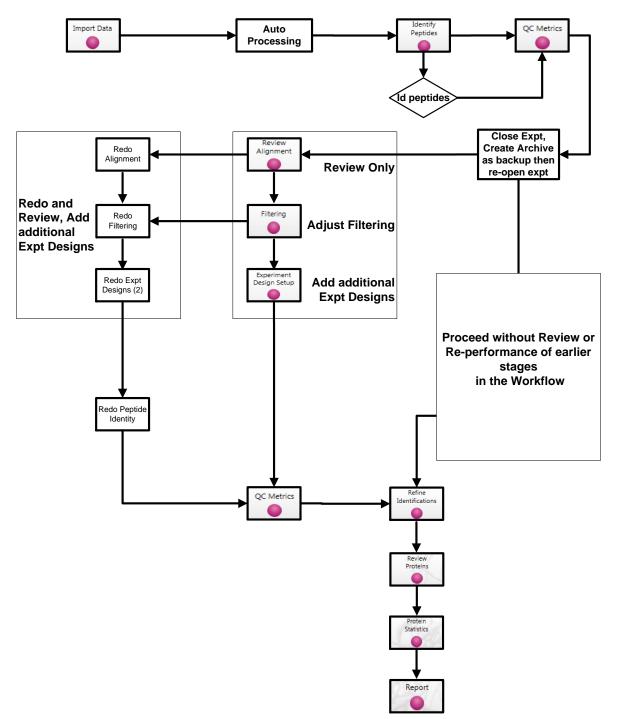
Note: if processing fails to complete successfully there are a number of suggested strategies you can use to proceed with your analysis. These are outlined in Appendix 2 (page 75)

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

Cancel

Stage 2B: After Automatic Processing

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



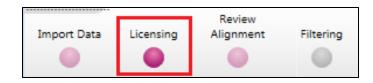
The flow chart gives you an overview of the various steps you can take to proceed with your automatically processed data.

In this example, as the data is DDA it is going to open at Identify Peptides so you can either:

- Continue with the analysis, as the Processing dialog is not displaying any warnings, and perform **Identify Peptides**. In which case you can go to page 41
- Open the analysis at Identify Peptides and return to the Review Alignment stage by clicking on it in the Workflow (page 14).

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, and you close Progenesis QI for proteomics you will be warned that the analysis will be lost.

OF Progenesis QI.p Tutorial for DDA - Progenesis QI for pro	teomics	
File Review Experi Import Data Licensing Alignment Filtering Design		nonlinear
		A Waters Company
Dongle License Runs		
This installation is currently restricted to analyse licensed runs only.	Run name	Licence License state this run
To license your runs, you need an evaluation or	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge	Unlicensed 🔽
lease licence code which can be obtained from	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge	Unlicensed 🔽
a sales representative.	$\label{eq:c:lustomer} C:\label{eq:c:lustomer} Users\ \ QI.p\ v3.0\ \ \ Utorials\ \ and\ \ Demo\ \ Suites\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	Unlicensed
Once licensed, your runs can be analysed on	$\label{eq:c:Users} C: Volume to Vo$	Unlicensed
any installation of the software. The licence is	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge	
automatically included when archiving an experiment.	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge	Unlicensed 🔽
If your runs have been licensed on another computer, <u>click here</u> to make the licences available on this computer. If you have one, you can <u>open a licence file</u> to install. If you have just installed a dongle, <u>click here</u> .	UP Unable to save experiment You cannot save the analysis without a valid license. If you close now your analysis will be lost.	
, , , , , , , , , , , , , , , , , , ,	OK Cancel	
	Run licence code:	Use Licence Code
		Section Complete 🦻

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 4 (page 78)

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 Alignment opens displaying the alignment of the runs to the Reference run (C1).

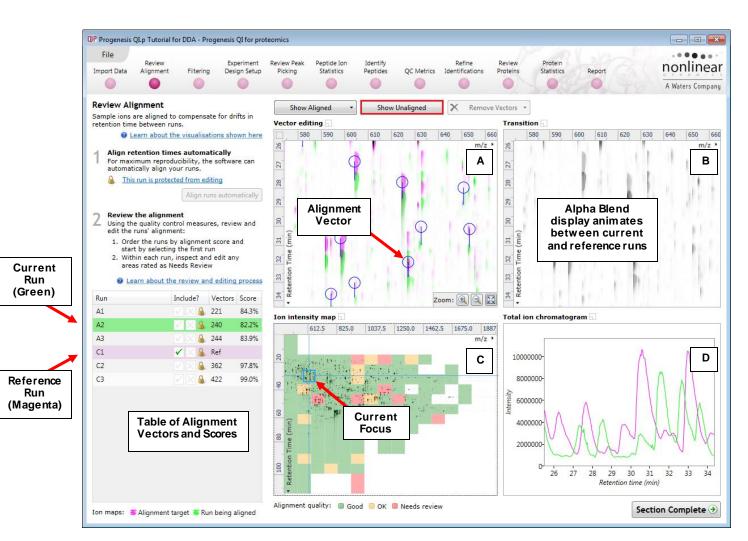
Layout of Alignment

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

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

Run	Include?	Vectors	Score
A1	🗹 🗙 🌲	221	84.3%
A2	🗸 🗙 🌡	240	82.2%
A3	🗹 🗙 🌲	244	83.9%
C1	🖌 🖂 🌡	Ref	
C2	$\mathbb{Z}\times\mathbb{A}$	362	97.8%
C3		422	99.0%

• In the Run table click on Run A2 to make it current. You will now be looking at the alignment of A2 to C1 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.

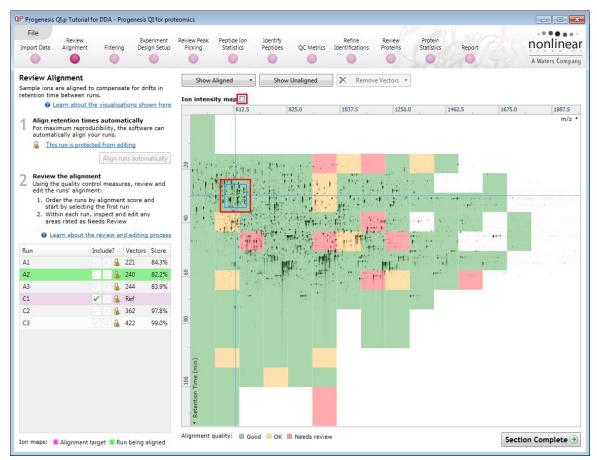
Transition (Window B): uses an **alpha blend** to animate between the current and reference runs. Before the runs are aligned, the peptide ions appear to move up and down. Once correctly aligned, they will appear to pulse. During the process of adding vectors, this view can be used to zoom in and also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



If the alignment has worked well then the lon Intensity Map will appear predominantly green and a score of greater than 80% will be reported in the table. Also when **Show Aligned** is selected, in Window A (vector editing) vector length should appear minimal and in Window B (Transition) will show peptide ions pulsing slightly but not moving up and down.

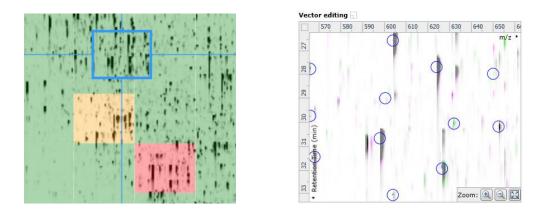
Note: you can use the icon to the right of the panel name to expand or contract each view

Reviewing Quality of Alignment

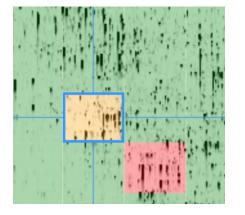
At this point the quality metric, overlaid on the lon Intensity Map as coloured squares, acts as a guide drawing your attention to areas of the alignment. These range from Good (Green) through OK (Yellow) to Needs review (Red). When reviewing individual squares double click on a coloured square to set the focus.

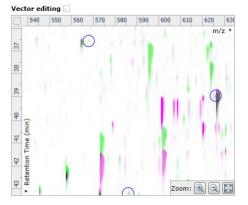
Three example squares are examined here.

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.

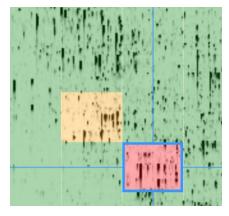


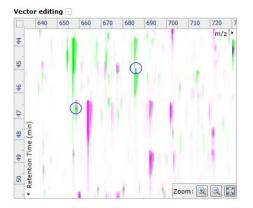
For a 'yellow' square, some of the data appears overlapped (black) indicating OK alignment. When viewed in the Transition view some of the data appears to pulse.





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

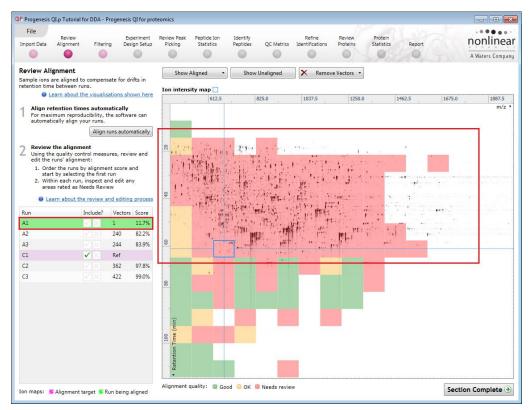




Note: the coloured metric **should be used as a guide**. In cases where there are a few 'isolated' red squares this this can also be indicative of 'real' differences between the two runs being aligned and should be considered when examining the overall score and surrounding squares in the current alignment.

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

Note: a marked area of red combined with a low score clearly indicates a 'misalignment' and may require some manual intervention (see Appendix 5, page 79).



Having performed the analysis automatically, in the course of reviewing the quality of alignment you decide that the alignment requires editing then you can unlock the analysis.

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

Review Alignment	
Sample ions are aligned to compensate for drifts in retention time between runs.	QIP Protected from editing
Learn about the visualisations shown here Align retention times automatically For maximum reproducibility, the software can automatically align your runs. This run is protected from editing Align runs automatically	Delete existing analysis? If you change the alignment, it will invalidate the current analysis including peptide ion pattern and IDs, editing, and tags. If you want to keep these, you should archive this experiment before changing the alignment and moving forward.
2 Review the alignment Using the quality control measures, review and edit the runs' alignment:	Delete analysis and allow editing Cancel
 Order the runs by alignment score and start by selecting the first run Within each run, inspect and edit any areas rated as Needs Review 	
Q Learn about the review and editing process	

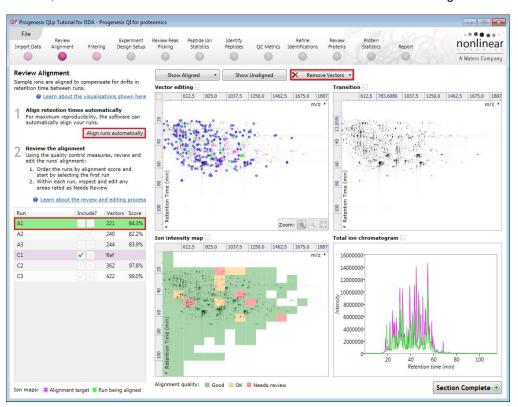
Note: as you click on the link, you will be warned that you are discarding the current analysis (all steps beyond alignment).

Details on editing alignment are described in Appendix 5 (page 79)

The alignment quality of this tutorial data set does not require any manual intervention.

If you have chosen to discard the current automatic analysis and have been exploring the alignment of one or more of the runs using manual vectors (for the purposes of this tutorial) make sure you first remove all manual vectors and then re-perform the Automatic alignment.

To do this for A1, first select Remove 'All vectors in the whole run' and then click Align runs automatically.



Having re-performed the Automatic alignment the Vectors and scores will appear as shown above.

Since you have unlocked your analysis you must now perform the Peak Picking that was originally performed as part of the **Automatic processing**.

To move to Peak Picking click on Filtering on the Workflow or Section Complete (bottom right).

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

OP Peak Picking Parameters	QIP Peak Picking Parameters
Runs for peak picking Peak picking limits Maximum charge Retention time limits Choose runs for peak picking Image: Choose runs for peak picking Image: Choose runs for peak picking You can tick or un-tick each run to control which will be used by the peak picking algorithm. Although any Image: Choose runs for peak picking	Runs for peak picking Peak picking limits Maximum charge Retention time limits Sensitivity You can adjust the sensitivity of the peak picking algorithm using these different methods. Each C Absolute ion intensity
run which is left un-ticked will not affect the peptide ion outlines, it will still have outlines added to it and will be available in the experiment design setup. Learn more about why you might not want to select all runs. A2 A2 A2 C1 C2 C2 C3	sensitivity method examines the intensities of groups of MS peaks to judge whether they are likely to form part of an ion or whether they represent noise and so should be ignored. Peaks that are rejected as noise will not be used to build ion outlines.
	fewer default more witch an ion has eluted. fewer default Minimum width: 0 minutes period will be rejected. minutes minutes
Start peak picking Cancel	Start peak picking Cancel

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 'replicate' 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 peak picking 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.

Chromatographic peak width					
The chromatographic peak width gives the length of time over which an ion has eluted. If you set a minimum peak width, any ion that has eluted over a shorter	Apply a minimum peak width				
	Minimum width:	0.15	minutes		
period will be rejected.					

For the runs in this user guide we will use the default settings for the Automatic method and NOT apply a minimum peak width.

The third tab allows you to set the **maximum charge** of the peptide ions, which will be detected. The default setting is a charge state of 20. For this example, leave this set as default.

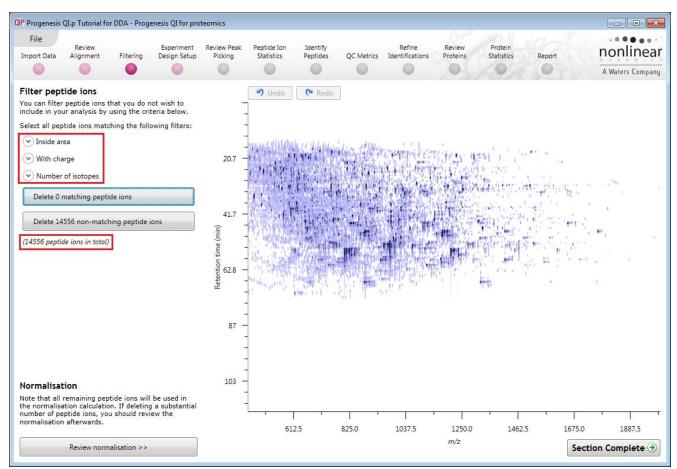
Note: you can either, leave this set as default and remove the high charge state peptide ions at the Filtering stage or you can choose not to detect them in the first place by reducing the charge state threshold here.

Finally, you can set **Retention time limits** for the detection. Default values are displayed in the unticked boxes and detection will use these values whenever the box(s) are unticked. Enter values of 10 and 75 min and tick the boxes as shown below.

QIP Peak Picking Parameters	QIP Peak Picking Parameters
Runs for peak picking Peak picking limits Maximum charge Retention time limits	Runs for peak picking Peak picking limits Maximum charge Retention time limits
Maximum allowable charge	Retention time limits
You can set the maximum charge of ions to be detected. Ions with a charge greater than this value will be rejected.	You can set the minimum and maximum retention time for peak picking. Ions that elute before or after these values will be ignored.
	Ignore ions after 75 minutes
Start peak picking Cancel	Start peak picking Cancel

Click Start Peak Picking to start the detection process.

On completion of detection, the Filtering stage will open displaying the number of peptide ions detected, in this example there are 14556 peptide ions.



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 state'.

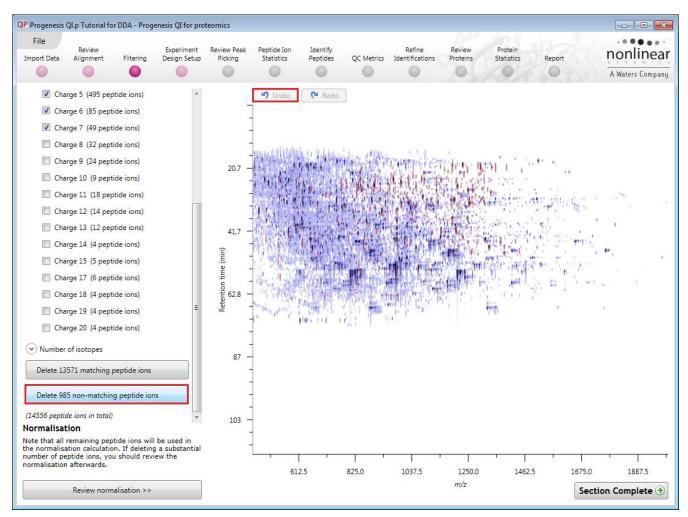
When **With charge** is selected the number of peptide ions present at each charge state is displayed, these can be selected accordingly. In this case we will retain peptide ions with a charge state of 2 to 7.

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

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

Filter peptide ions
You can filter peptide ions that you do not wish to include in your analysis by using the criteria below.
Select all peptide ions matching the following filters:
✓ Inside area
With charge
Charge 1 (849 peptide ions)
Charge 2 (5845 peptide ions)
Charge 3 (5351 peptide ions)
Charge 4 (1746 peptide ions)
Charge 5 (495 peptide ions)
Charge 6 (85 peptide ions)
Charge 7 (49 peptide ions)
Charge 8 (32 peptide ions)
Charge 9 (24 peptide ions)

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 985 Non Matching Peptide ions.

Note: if you have deleted peptide ions unintentionally you can use the **Undo** button to bring them back, however, when you move to the next section you will lose the capacity to undo the filter.

In this example do not undo the charge filter you performed as described above.

Before moving on from filtering you can review the normalisation of the data.

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

Review normalisation is accessed from the button at the bottom left corner of the filtering page.

Normalisation	103 -
Note that all remaining peptide ions will be used in the normalisation calculation. If deleting a substantial number of peptide ions, you should review the normalisation afterwards.	-
Review normalisation >>	

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

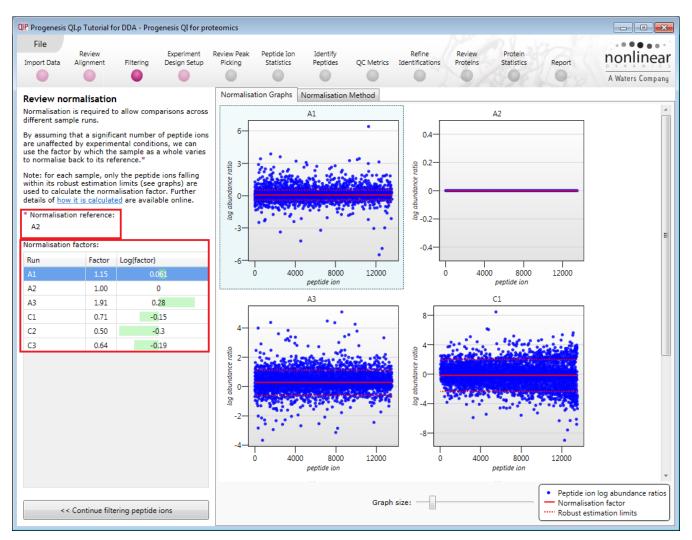
Recalculating normalisa	ion

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).

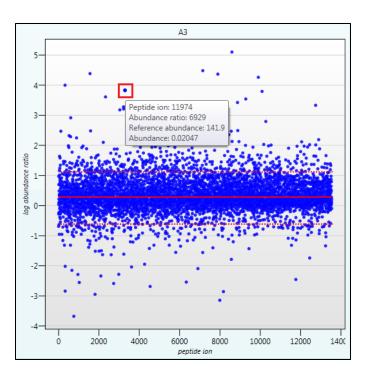
Normalisation factors are reported in the table to the left of the plots.



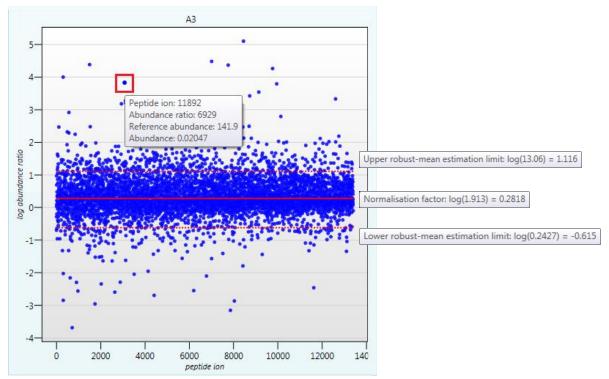
Calculation of Normalisation Factor:

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

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. Waters 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.



Note: if you do not wish to work with normalised data then Select 'Don't use any normalisation'.

IP Progenesis	QI.p Tutorial for	DDA - Prog	enesis QI for pro	teomics								
File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Ion Statistics	Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report	A Waters Company
Normalisation different sam By assuming are unaffecte use the factor	ple runs.	nt number o ntal conditio sample as a		Normalise to Normalise to	all proteins	keepina prot						

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

For this experiment, you should leave the Normalise to all peptide ions option selected.

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

	0
<< Continue filtering peptide ions	

For this example, we **DO NOT** do any additional Filtering beyond the Charge state filter you applied on page 21, 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.

QIP	Progenesis	QI.p Tutorial fo	r DDA - Proge	enesis QI for pro	oteomics										×
I	File nport Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Ion Statistics	Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report	-	A Waters Compar	s
E	New													🕜 Help	, •
1	Which ex	periment o	lesign typ	e do you v	ant to use	for this ex	periment	?							
	0 0 0 0 0 0	Between-s	ubject De	esign		Cr	eate	0-0 0-0 0-0	/ithin-subje	ct Design					
	appear use the	oles from a gi in only one c between-sub	ondition? T ject design	hen A		Di A1 <u>Re</u>	elete move	given sub	u taken sample bject under dif hs? Then use th design.	ferent		Before	During	After	
	group th conditio samples	et up this design, you simply the runs according to the ition (factor level) of the les. The ANOVA calculation				A2 <u>Re</u> A3 <u>Re</u>	emove	Patient X	X1	X2	X3				
	indepen statistic	dent and ther al test of whe	c c c c c c c c c c c c c c c c c c c				elete move move	For exam type of de experime been sam	Patient Y	Y1	Y2	Y3			
				Add co	ndition	C3 <u>Re</u>	<u>:move</u>	To set up software (factor le also whic The softw	o this design, yo not only which evel) each run b h subject it can vare will then p measures ANO	u tell the condition elongs to bu me from. perform a	Patient Z	Z1	Z2	Z3	
								because t assumptio repeated differenc reduced a condition	rd ANOVA is not the data violate on of independe measures ANO es can be elimi as a source of b differences (w more powerful	es the ANOVA ence. With a VA individual nated or between rhich helps to	l				
								thought o paired-sa comparise	in-subject desig of as an extensi mples t-test to on between mo measures.	on of the include					

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 6 page 83

This experiment contains 2 conditions: A and C 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 move the cursor over this option and click to open the dialog.

QIP Create New Experiment Design	—
Enter a name for the experiment design:	
AC	~
How do you want to group the runs?	
Group the runs manually	
○ Copy an existing design:	•
O Import criteria from a file:	Browse
Group runs by: <a>(<no groups="" valid="">)</no>	▼
What file formats are supported? Create design	Cancel

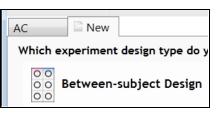
Give the new experimental design a name and then click Create design.

OP Progenesis	QI.p Tutorial fo	or DDA - Pro	genesis QI for p	roteomics								- • •
File	Review	-	Experiment	Review Peak		Identify		Refine	Review	Protein	921	nonlinear
Import Data	Alignment	Filtering	Design Setup	Picking	Statistics	Peptides	QC Metrics	Identifications	Proteins	Statistics	Report	DYNAMICS
	_	-	· · · ·		-		-				1.8	A Waters Company
ACIX	New											🕡 Help 🔻
Setup con	ditions			Runs	Add Selected F	Runs to Condi	tion V Se	earch		2		
(e.g., contro	onditions that y ol, drug A, etc) is to the correc	, and then a	compare below assign each of	v 🗖	Add to nev	v condition		C2			C3	
A			Delete		A	_						
			A1 <u>Remove</u>		1.24.24			- Selfer Selfer				
			A2 <u>Remove</u>									
			A3 Remove									
Add conditi	ion											
											Secti	on Complete Э

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 required 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 spreadsheet at the time of sample collection and/or preparation.



For this example there is a **Tutorial Groups.csv** 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 Tutorial Groups file and select what to **Group by**, for example: **Conditions**.

QP Create New Experiment Desig	n 💌
Enter a name for the experime	ent design:
AC-2	(+)
How do you want to group the	runs?
Group the runs manually	
Copy an existing design:	AC •
Import criteria from a file:	DDA_v3.0 Tutorial Groups.csv Browse
Group runs by:	Conditions
	Conditions Date of Collection
What file formats are suppor	Location

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

Note: currently this second method is the only method of creating an Experiment Design that can be applied when using the Automatic Analysis process. Additional Designs can be applied post Automatic processing.

QIP Progenesis	QI.p Tutorial f	or DDA - Pro	genesis QI for p	roteomics								
File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Ion Statistics	Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report	nonlinear
									0			A Waters Company
AC	AC-2 I	× 🖻 N	ew									🕜 Help 🔻
	nditions that I, drug A, etc)	, and then a	compare belov issign each of Delete		Add Selected F	Runs to Cond	ition ▼ Se	earch		٩		
			A1 <u>Remove</u> A2 <u>Remove</u> A3 <u>Remove</u>									
C C			C1 <u>Remove</u> C2 <u>Remove</u>									
Add condition	<u>on</u>											
											Sect	ion Complete 🏵

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

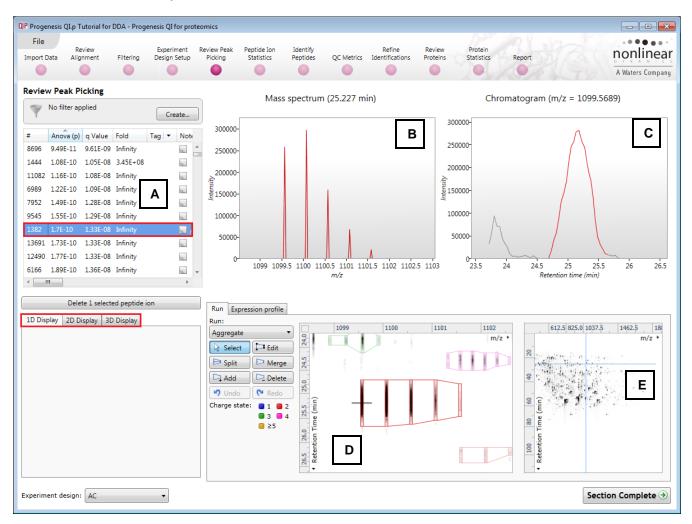
The purpose of this stage in the Workflow is to review the list of peptide ions using the visual tools provided and edit peptide ions if required.

The review stage has 4 display modes: 1D, 2D, 3D 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 grouping.

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 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.

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

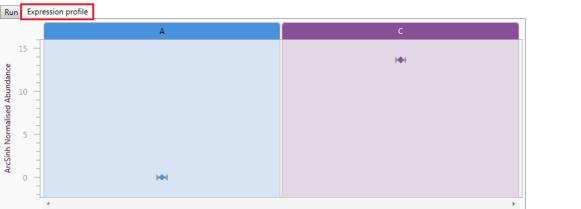
Window C: displays the Chromatogram for the current peptide ion on the selected Run (in window D). Hold the cursor over the peak to indicate Retention time and intensity

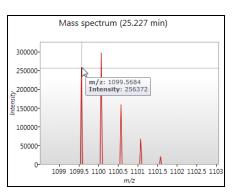
Window D: displays the details of the currently selected run. By default the selected run is an Aggregate of all the aligned runs.

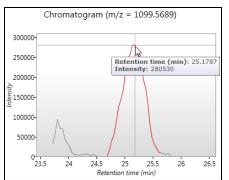
Details of individual runs can be viewed by using the 'Run' link and selecting the run you wish to view.

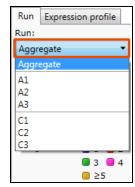
The peptide ion editing tools are located in this window (see page 32 for functional explanation).

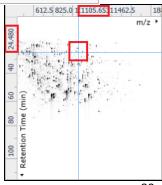
Clicking on the Expression Profile tab in Window D shows the comparative behaviour of the peptide ion across the various biological groups based on group average normalised volume. The error bars show +/- 3 standard errors.



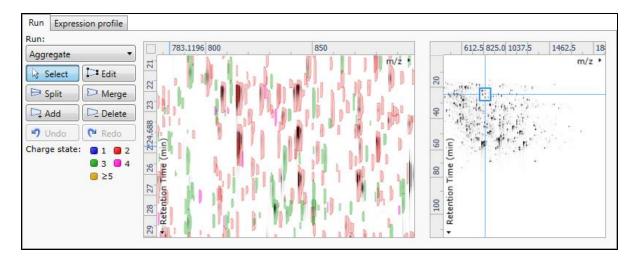








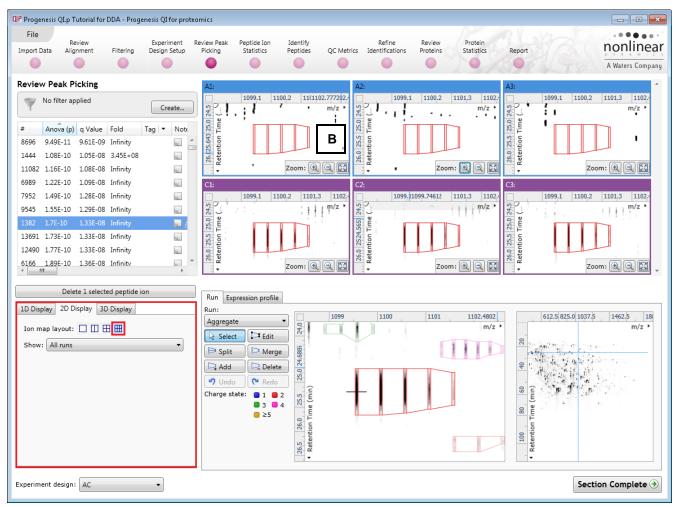
You can also drag out a larger area on this view that will refocus the other windows.



The 2D Display

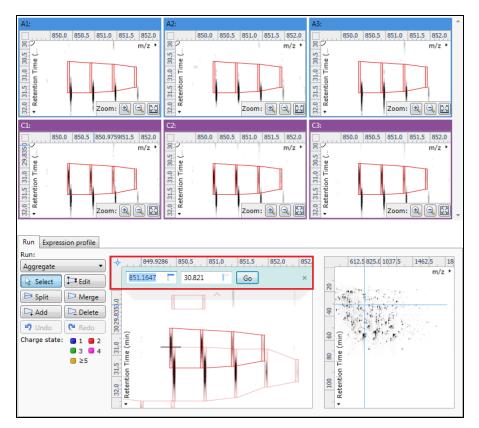
Windows A, D and E: perform the same functions across all 4 display modes.

In the 2D Display 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.

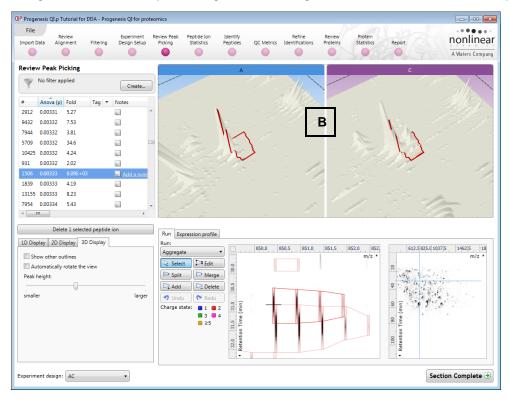
Using the the various views in the 2D display one can examine the peptide ion detection in detail to validate the correct detection of even fully overlapping peptide ions as shown below.



Note: Use **Go To Location** tool, in the Run window, to find peptide ion of interest then click on the monoisotopic envelope to refocus view.

The 3D Display

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

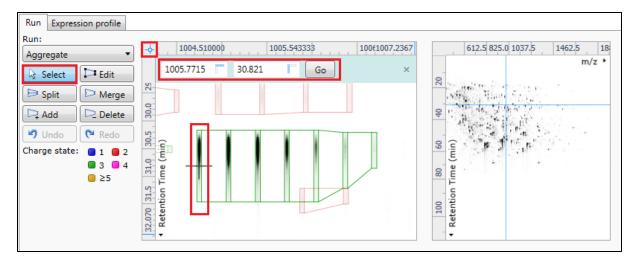


The views can be set to **Rotate** automatically or you can rotate them manually by clicking and dragging them with the mouse.

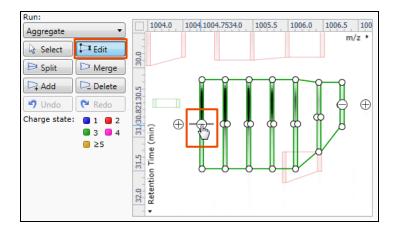
Editing of peptide ions in the View Results stage

As an example of using the editing tools which are located on the left of the **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.

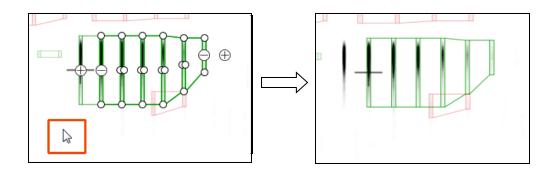
1. Locate the peptide ion at approx 1005.7 m/z and 30.8 min using the **Go To Location** tool (at the intersection of the m/z and RT axis).



2. Select the Edit tool and click on the peptide ion to reveal the 'edit handles'

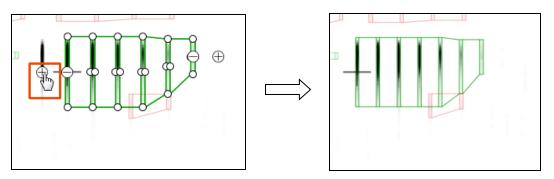


3. Click on the 'minus' handle over the monoisotopic peak to remove it.



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.

Reviev	v Peak Pi	cking					Review	v Peak Pi	cking				
Y	No filter ap	plied		Cre	ate		7	No filter ap	plied			Cre	ate
#	Anova (p)	Fold	Tag 💌	Notes			#	Anova (p)	Fold	Ta	•	Notes	
7527	0.591	1.11		10	*		7527	0.591	1.11				-
5152	0.591	4.3		10		P	5152	0.591	4.3			D.	
12717	0.591	1.37		0		Edit	12717	0.591	1.37			a.	
2191	0.591	1.24		a.		<u>_</u>	2191	0.591	1.24				_
181	0.591	1.07		Add a	a note		14390	0.591	1.07			🔜 Add a	a note
3790	0.591	1.16		10.			3790	0.591	1.16			4	
6856	0.591	1.2		12			6856	0.591	1.2				
3917	0.592	1.28		1		🔊 Undo	3917	0.592	1.28			4	
11519	0.592	1.73		1		, ondo	11519	0.592	1.73			4	
3222	0.592	1.2		12	-		3222	0.592	1.2				
•	"				- F		<	"					•

Note: use Undo to reverse this editing process, this restores the original # and removes the Edit tag.

Also to remove the Edit tag right click on the table, select Edit tags, and delete the Edit Tag from 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 25). 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.

Y	No filter ap	plied		Create	2
#	Anova (p)	Fold	Tag 💌	Notes	
7527	0.591	1.11		10	
5152	0.591	4.3		10	
12717	0.591	1.37		10	
2191	0.591	1.24		10	
181	0.591	1.07		🔜 Add a n	ote
3790	0.591	1.16		2	
6856	0.591	1.2		10	
3917	0.592	1.28		10	
11519	0.592	1.73		12	_
3222	0.592	1.2		12	-
•					F

Then order on **Abundance** and select all peptide ions with an Abundance of 1E+05 and greater, (the exact number is not important).

port Data Al	Review ignment	Filtering	Experim Design S		eview Peak Picking	Peptide Ion Statistics	Identify Peptides	QC Metrics		Refine tifications	Review Proteins	Protein Statistics	Report		K		A Waters Comp
view Peak	-																Create
Anova (p) q Value	Fold	Tag 🔻	Notes		Highest Mean	Lowest Mean	m/z	z	Mass	RT (mins)	RT window (mins)	Abundance	Intensity	Max CV (%)	MS/MS	Protein Peptic
0.105	0.0263	1.1		Add	l a note	A	С	805.4410	3	2413.301	54.875	7.65	1.25E+08	1.08E+08	7.6	152	
8 0.547	0.103	1.04		la.		С	A	1207.6552	2	2413.296	54.899	4.9	9E+07	5.04E+07	7.89	102	
3.12E-06	7.42E-06	5.66E+04		la.		С	A	1100.5867	3	3298.738	44.928	2.43	8.3E+07	7.87E+07	53.3	18	
3.8E-06	8.67E-06	1.1E+03		la.		A	С	1176.2270	3	3525.659	48.238	9.44	6.92E+07	2.22E+07	23.5	58	
4.39E-07	1.75E-06	844		la.		A	С	656.8612	2	1311.708	44.137	4.21	6.21E+07	1.17E+08	15	38	
2.02E-07	9.91E-07	777		ha.		с	A	988.9849	2	1975.955	50.605	4.65	5.09E+07	9.19E+07	14.2	46	
8.17E-08	5.34E-07	128		la.		с	A	663.8693	2	1325.724	46.597	4.1	4.95E+07	1.69E+08	6.67	63	
4.83E-06	1.04E-05	5.46E+03		la.		С	A	900.9713	2	1799.928	39.272	2.97	4.52E+07	8.16E+07	46.8	18	
6.3E-06	1.27E-05	3.75E+03		la.		A	С	1061.0070	2	2119.999	53.288	6.75	4.31E+07	2.52E+07	34.3	67	
5.27E-06	1.11E-05	1.35E+03		ha.		A	c	997.4477	2	1992.881	31.814	2.57	4.25E+07	3.81E+07	34	29	
	lete 1 selecti		on		Run Expr Run: Aggregate Select Split Add Ocharge sta	Edit Edit Merge C Delete Redo	min) 55	805		806	807		9 m/z •	100 80 60 40 20 Retention Time (min) 2 2 3 3	12.5 825.0 10	37.5 1250	0 1462.5 1675.0 18 m/z ♪

With the peptide ions still highlighted right click on them and select 'New Tag'.

Review Vesk Picking ✓ Nefter splet Softer splet <th< th=""><th>File</th><th>Re</th><th>eview inment</th><th>Filtering</th><th>Experim Design Se</th><th></th><th>Peptide Ion Statistics</th><th>Identify Peptides</th><th>QC Metrics</th><th></th><th>Refine ntifications</th><th>Review Proteins</th><th>Protein Statistics</th><th>Report</th><th></th><th>K</th><th></th><th>A Waters Compa</th></th<>	File	Re	eview inment	Filtering	Experim Design Se		Peptide Ion Statistics	Identify Peptides	QC Metrics		Refine ntifications	Review Proteins	Protein Statistics	Report		K		A Waters Compa	
4433 0.000655 0.000641 10.1 Add a note. A C 841.1300 4 3360.491 32.829 0.473 1E+05 223E+05 38.1 0 2711 0.0239 0.000651 4.55 Add a note. C A 766.0764 3 2001.213 46.185 1.33 1E+05 223E+05 38.1 0 2203 0.000517 4.55 Add a note. A C 6179252 1653.380 39.467 0.743 1E+05 257E+05 21.9 6 2203 0.000550 0.000461 12.9 Add a note. A C 669.0073 3 2004.000 28.825 0.644 1E+05 367E+05 37.7 0 4179 0.00263 0.00155 103 C A 1220.6313 3 3568.872 39.245 Net tags to assign 14 407E+05 31 12 9306 0.0247 0.00666 1.68 A C 1226.8008 4 5047.014 53.225 14 407E+05 31 12			•															Create	
2771 0.039 0.00844 3.19 Add a note C A 768.0784 3 230.213 46.85 1.33 1E+05 3.89E+05 47.2 6 2269 0.00056 0.00041 1.29 Add a note C A C 660.077 3 200.203 0.467 0.743 1E+05 3.87E+05 37.7 0 2724 0.33 0.0075 1.49 Add a note C A 862.4877 2 163.830 3.467 0.743 1E+05 3.87E+05 37.7 0 2724 0.33 0.00751 1.49 Add a note C A 862.4877 2 163.837 19.24 17.475 4.44E+05 94.9 9 1179 0.00263 0.00155 103 C A 122.06313 3 3658.872 39.245 No bags to assign plat 2.72E+05 84.9 0 122.35 2060 0.0247 0.00866 16.8 A C 1262.8608 4 139.735 2.400 2.5E+05 29.2 9	;	Anova (p)	q Value	Fold	Tag 💌	Notes	Highest Mean	Lowest Mean	m/z	z	Mass	RT (mins)	RT window (mins)	Abundance	Intensity	Max CV (%)	MS/MS	Protein Peptic	
10000021 0.0000517 4.55 Add a note A C 817,9225 2 1633.830 39.467 0.743 1E-05 2.57E-05 2.19 6 203 0.000566 0.001615 12.9 Add a note A C 669.0073 3 2004.000 28.825 0.644 1E-05 3.57E-05 37.7 0 179 0.00263 0.00155 103 C A 862.4887 2 1220.633 48.371 0.899 1E-05 4.44E-05 94.9 9 170 0.00263 0.00155 103 C A 1220.6313 3 3568.872 39.245 No tags to assign 4 4.44E+05 84.9 0 160 0.0154 0.00609 1.68 A C 1262.8604 2 1395.735 24.605 New tag	433	0.000865	0.000641	10.1		Add a note	A	с	841.1300	4	3360.491	32.829	0.473	1E+05	2.23E+05	38.1	0	4	
203 0.000461 12.9 Add anote A C 669.073 3 2004.000 28.825 0.644 1E+05 3.67E+05 37.7 0 374 0.333 0.0673 1.49 Add anote C A 8624887 2 1722.963 48.371 0.859 1E+05 3.67E+05 94.9 9 179 0.00263 0.00155 103 C A 1220.6313 3 3658.872 39.245 No tags to assign p4 2.72E+05 84.9 0 170 0.00263 0.0019 1.34 A C 698.8746 2 1395.735 24.000 New tags: p4 4.07E+05 31 12 126 0.0154 0.00609 2.11 A C 876.4850 2 1750.956 34.204 Edit tags: p4 2.5E+05 29.2 9 141 166 0.0154 0.00609 2.11 A C A 837.4590 2 1672.04 50.879 12 9.98E+04 6 5 <td cob<="" td=""><td></td><td></td><td></td><td></td><td></td><td>Add a note</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td>	<td></td> <td></td> <td></td> <td></td> <td></td> <td>Add a note</td> <td></td>						Add a note												
374 0.333 0.0673 1.49 Add a note:: C A 862.4887 2 1722.963 48.371 0.859 1F-415 4.44F-05 94.9 9 179 0.00263 0.00155 103 C A 1220.6313 3 3558.872 39.245 No tag: to assign 94 2.72E+05 84.9 0 716 0.301 0.6619 1.34 A C 698.8746 2 1395.735 24.605 New tag 94 4.07E+05 31 12 305 0.0247 0.00860 16.8 A C 1262.8608 4 5047.414 53.245 Quick Tag 94 2.5E+05 29.2 9 744 1.665-05 3.07E-05 2.0.8 C A 837.4590 2 1672.904 50.879 1.2 9.98E+04 3.95E+05 1.8 6 Intervention profile Run: Expression profile Note: 1 2 3 6 90 90 90 90 9						Add a note												C	
179 0.00263 0.00155 103 C A 1220.6313 3 3658.872 39.245 No tag to assign to						🔜 Add a note													
1/3 0.00235 103 C A 122.0513 3 305.8/2 3 23.25 14 2.124:05 84.9 0 126 0.001 0.0619 134 A C 698.874 2 1395.731 24.00 New tag 14 4.07E+05 31 12 05 0.0247 0.00660 1.68 A C 1262.8608 4 50.2450 New tag 14 4.07E+05 31 12 05 0.0247 0.00669 2.11 A C 876.4850 2 175.0956 34.20 Quick Tags 14 4.07E+05 29.2 9 166 0.0154 0.00609 2.11 A C 876.4850 2 175.0956 34.20 Edit tags 14 2.5E+05 29.2 9 144 1.86E+05 3.07E+05 2.0.8 C A 837.4590 2 1672.904 50.879 1.2 9.98E+04 3.5E+05 18 6 150 Display 3D Display 3D Display 30 4 9 806 807						Add a note						48.371							
All 0.00247 0.00866 1.6.8 A C 1262.8608 4 50.47.414 53.245 Quick Tags p4 7.69E+04 60.1 0 0.00660 2.11 A C 876.4850 2 1750.956 34.20 Edit tags p4 2.5E+05 29.2 9 244 1.86E+05 3.07E+05 20.8 C A 837.4590 2 1672.904 50.879 1.2 9.98E+04 3.95E+05 1.8 6 Toelete 2330 selected peptide ions Toelete 2330 selected peptide ions Num Expression profile Run: Add Delete Select 1 9 Select Pedet 9.86 809 612.5 825.0 1037.51250.01462.45.675.018 Normality Select Pedet Pedet<	179	0.00263	0.00155	103		10	С	A	1220.6313	3	3658.872	39.245	No tags to assig	n 04	2.72E+05	84.9	0		
000 0000	16	0.301	0.0619	1.34		19	A	С	698.8746	2	1395.735	24.606		04	4.07E+05	31	12		
244 1.86E-05 3.07E-05 2.0.8 C A 8.37.4590 2 1672.904 50.879 1.2 9.98E+04 3.95E+05 1.8 6 Delete 2330 selected peptide ions Run Expression profile Run: 6 1.02	306	0.0247	0.00866	16.8		10	A	С	1262.8608	4	5047.414	53.245	Quick Tags	• 04	7.69E+04	69.1	0		
Delete 2330 selected peptide ions Run Expression profile D Display 2D Display 3D Display	166	0.0154	0.00609	2.11		10	A	С	876.4850	2	1750.956	34.204 😭	Edit tags	04	2.5E+05	29.2	9		
Delete 2330 selected peptide ions Run Expression profile D Display 2D Display 3D Display B Display 2D Display 3D Display	744	1.86E-05	3.07E-05	20.8		9	С		837.4590	2	1672.904	50.879	1.2	9.98E+04	3.95E+05	18	6	•	
20 De 10 DE) Disp				: ions	Run: Aggreg Sel Ad	ate ect THE Edit it DHerge d DE Delete do CHE Redo state: 1 0 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2 4 8004 49 49 49 49 49 49 49 49 49 49 49 49 49				807	808 80		80 60 40 20 on Time (min)	12.5 825.0 10	137.5 1250		

Give the Tag a name. i.e. 'Most abundant'.

Create new tag	×
Most abundant	
	OK Cancel

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

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**.

-	No filter ap	plied							
Y	No Inter ap	piieu							
#	Anova (p)	q Value	Fold		Tag 🔻	Notes	5	Highest Mean	Lo
2369	0.000821	0.000617	4.55	;		ha.		Α	С
2203	0.000566	0.000461	12.9)		10		Α	С
2374	0.333	0.0673	1.49)		Ac	ld a note	С	А
4179	0.00263	0.00155		Most :	abundant			С	А
1716	0.301	0.0619		Newt	ag			Α	С
9306	0.0247	0.00866		Quick	Tags	•	Anova	p-value	
2166	0.0154	0.00609	P	Edit ta	gs		Max fo	ld change	
2744	1.86E-05	3.07E-05	20.8	3		1a	Modifi	cation	
1921	0.00415	0.00219	2.96	5E+03		la.	No MS	/MS data	
1164	0.0351	0.0113	1.56	;			No pro	tein ID	

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 have an **Anova p-value \leq 0.05** click on **Create** on the filter panel above the table.

No filter applied	Create
-------------------	--------

Y	No filter ap	plied		Cre	ate
#	Anova (p)	q Value	Fold	Tag 💌	Not
2711	0.0239	0.00844	3.19		
2369	0.000821	0.000617	4.55		
2203	0.000566	0.000461	12.9		
2374	0.333	0.0673	1.49		
4179	0.00263	0.00155	103		12
1716	0.301	0.0619	1.34		12
9306	0.0247	0.00866	16.8		12
2166	0.0154	0.00609	2.11		12
2744	1.86E-05	3.07E-05	20.8		12
1921	0.00415	0.00219	2.96E+03		12
۰					•

QP New Quick Tag	×
Where a feature has:	
Anova p-value: ≤ ▼ 0.05	
Apply the following tag:	
Anova p-value ≤ 0.05	
Create tag	ancel

QIP New Quick Tag
Where a feature has: Max fold change: ≥ ▼ 2
Apply the following tag:
Max fold change ≥ 2
Create tag Cancel

Review	w Peak Pi	icking			
	No filter ap	plied		Crea	ate
#	Anova (p)	q Value	Fold	Tag 💌	Not
2004	0.208	0.0453	1.11		•
1451	0.000124	0.000141	4.58		10.
959	0.619	0.114	1.08		
3568	0.892	0.156	1.01		14
1648	0.363	0.0724	1.31		14
1158	0.0165	0.00639	442		
1425	0.023	0.00821	2.24		ha.
2626	0.00161	0.00105	3.51		14
7817	0.00951	0.00421	173		14.
6982	0.0222	0.008	226		·
					- F

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

Now order the current peptide ions in the table by the **Highest mean** so that all the peptide ions showing the highest mean for **Condition C** are at the top of the list.

Then highlight all the peptide ions with the highest mean for condition C and create a new Tag for them.

#	Anova (p)	q Value	Fold	Tag 💌	Notes	Highest Mean
3778	0.0074	0.00348	697	-	Add a note	С
2095	0.000461	0.000395	2.13	-	🔜 Add a note	С
2128	3.28E-05	4.93E-05	115	-	🔜 Add a note	С
96	0.0176	0.00674	2.39	-	🔜 Add a note	С
2113	1.21E-06	3.6E-06	Infinity	-	🔜 Add a note	C
10979	0.043	0.0131	55.1		19.	А
9791	0.000557	0.000455	45.6		10	А
1200	0.00165	0.00107	31.3		10.	А
11231	0.044	0.0134	31		10.	А
11142	1.2E-05	2.15E-05	Infinity			Α

Create a tag for them called Significantly up in C, tagging 3937 peptide ions.

Finally view the tags you have just created by clicking on Edit in the Tag filter panel, above the table.

		OIP Filter the peptide ions	
	Tag filter applied peptide ions may be hidden Edit	Create a filter Show or hide peptide ions based on a selection of their tags. Move tags to the appropriate boxes to create the filter. For more guidance, please see the <u>online reference</u> .	
		Available tags:	Show peptide ions that have all of these tags:
Make sure that only the tag for the Most abundant peptide ions is shown and press OK.		Anova p-value ≤ 0.05 (7943 peptide ions) Max fold change ≥ 2 (9898 peptide ions)	Most abundant (2330 peptide ions)
		Significantly up in C (3937 peptide ions)	Show peptide ions that have at least one of these tags:
			Hide peptide ions that have any of these tags:
_			
To move to the next stage in the workflow, Peptide Statistics, click Section Complete .		<u>Clear the filter</u>	OK Cancel

Stage 8: Peptide Ion Statistics on selected peptide ions

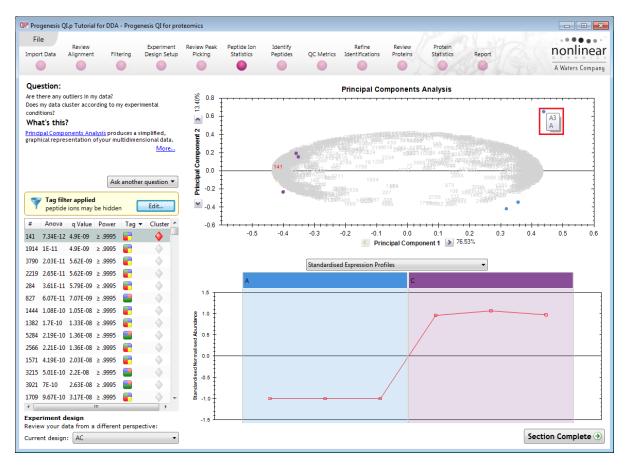
The user guide now describes the functionality of the Multivariate Statistics.



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.

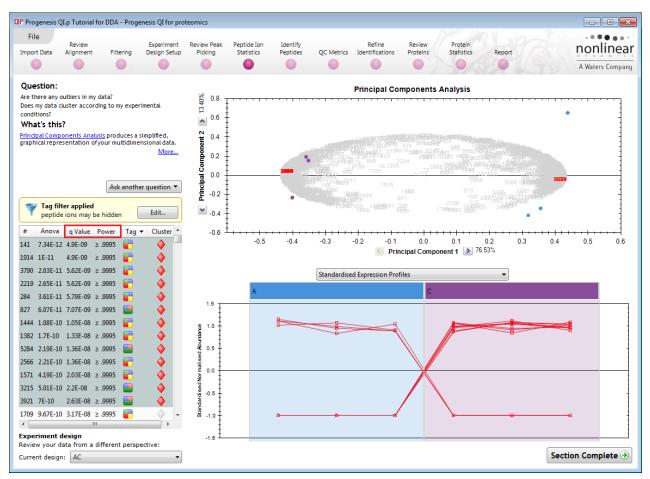
Principal Component Analysis (PCA)

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 observed on the 2D PCA plot can be compared to your experimental conditions and conclusions made 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 7 (page 85)

Correlation Analysis

Use the tags, created in Review Peak Picking, to filter the peptide ions in the table.

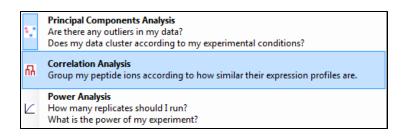
We are going to explore the Correlation Analysis for all the peptide ions that were tagged at the view results stage with an **Anova p-value≤0.05**.

On pressing OK the PCA will recalculate using these peptide ions, you can (to save time) stop this calculation by pressing **Cancel calculation** and then set up Correlation Analysis for the 7890 peptide ions.

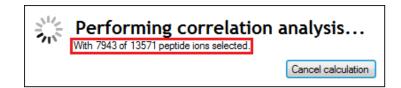
Create a filter Show or hide peptide ions based on a selection to create the filter. For more guidance, please s	
Available tags:	Show peptide ions that have all of these ta
Most abundant (2330 peptide ions) Max fold change ≥ 2 (9898 peptide ions) Significantly up in C (3937 peptide ions)	Anova p-value ≤ 0.05 (7943 peptide ion Show peptide ions that have at least one o these tags:
	Hide peptide ions that have any of these tags:

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'



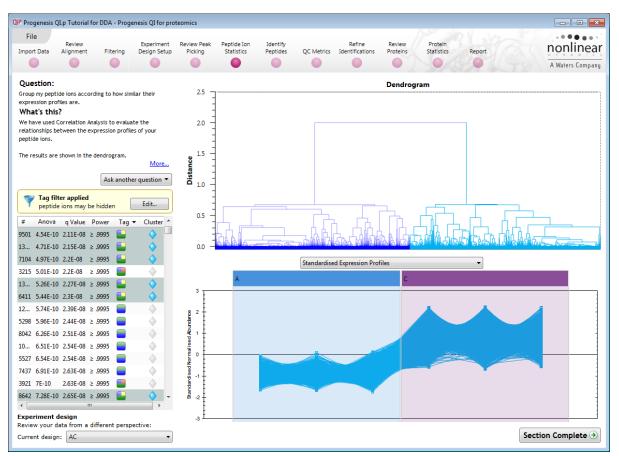
This time the statistically based question(s) being asked is:

'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.

To highlight all the peptide ions demonstrating **Increased expression in the C** group click on a 'node' for a branch of the Dendrogram (as shown above). As before, right click on the highlighted peptide ions in the table and create a Tag for these peptide ions (Up regulated in C).

Create new tag	—
Up regulated in C	
	OK Cancel

Also create a tag for those peptide ions showing **Increased expression in A** by first clicking on the other 'main' node then right click on the highlighted peptide ions in the table and creating the New tag (Up regulated in A).

Create new tag		—
Up regulated in A		
	ОК	Cancel

Comment: When you review the tags using **Edit** you can see that the Magenta and Yellow tags have been assigned to the same number of peptide ions. This shows how tabulated information about peptide ions can be used alongside interactive graphical plots of multivariate statistical analysis to explore your data.

Note: two groups is a special case, for more groups this will not be the case, and additional tagging will be required.

IP Filter the peptide ions Create a filter Show or hide peptide ions based on a selection to create the filter. For more guidance, please s	of their tags. Move tags to the appropriate boxes see the <u>online reference</u> .
Available tags:	Show peptide ions that have all of these tags:
Most abundant (2330 peptide ions) Max fold change ≥ 2 (9989 peptide ions) Significantly up in C (3937 peptide ions) Up regulated in C (3937 peptide ions) Up regulated in A (4006 peptide ions)	Anova p-value ≤ 0.05 (7943 peptide ions) Show peptide ions that have at least one of these tags: Hide peptide ions that have any of these tags:
Clear the filter	OK Cancel

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.

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

Stage 9: Identify peptides

Progenesis QI for proteomics does not perform peptide identifications itself for DDA data. Instead it supports identifications 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.

NOTE: If you have come straight to this stage having performed Automatic Processing then please ensure that you have first filtered your data as described on page 20, to leave only peptide ions with a charge state between 2 and 7. For completeness you may also wish to tag your data as described in the previous sections (pages 33 to 41) before performing the Peptide Identification.

File nport	e t Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Io Statistics		Identify Peptides		Refin Identifica		Protein Statistics Report	Te		A Waters Comp
un m		ptides n searches by entification pr		eak list files	MS/MS Spe Batch in		tions f	or creati	ng export list of	ms/ms spec	tra				
-	No filte	er applied	(Export	Rank #		Run	Scan number	Exported	Peptide ion intensity	Precursor intensity	(%)	Charge	Precursor m/z
T			l	Create	V	51		A1	2686	No	7.6e+007	5.2e+005	0.7	2	539.3007
	MS/MS	Proteins	Ta	g 🔻 Scor		12	1	A1	2726	No	7.6e+007	5.1e+007	66.5	2	539.3010
4	221	0	10		V	8	1	A1	2767	No	7.6e+007	6.3e+007	82.2	2	539.3007
	152	0				13	1	A1	2806	No	7.6e+007	3.9e+007	50.6	2	539.3008
2	109	0			v	19	1	A1	2855	No	7.6e+007	1.5e+007	20.2	2	539.3004
		-		-	V	26	1	A1	2906	No	7.6e+007	2.9e+006	3.8	2	539.3008
	102	0			V	33	1	A1	2948	No	7.6e+007	1.5e+006	1.9	2	539.3010
	95	0		-	v	43	1	A1	2990	No	7.6e+007	8.5e+005	1.1	2	539.3008
	88	0			V	47	1	A1	3040	No	7.6e+007	7.3e+005	1.0	2	539.3007
	85	0			V	32	1	A2	2846	No	7.7e+007	1.5e+006	2.0	2	539.3004
	80	0			V	6	1	A2	2890	No	7.7e+007	6.6e+007	85.5	2	539.3012
	79	0			V	10	1	A2	2931	No	7.7e+007	6.1e+007	79.6	2	539.3006
	75	0			V	17	1	A2	2979	No	7.7e+007	2.5e+007	33.1	2	539.3007
	72	0			4										•
31	71	0			Dentida	_	h		E20 2004		time 29.63 min, c		_	_	
	71	0		_				<u> </u>	2 339.3004, 1	etention	ume 29.65 min, c	narge +z			
				P.	Run:A1 S	can numl	ber:26	86	1	5000-					
asc elp E	ot	ch program yc 258 ms/ms spe ccessing	•		28.987- iiii 28.987- iiii 30.994-	ē			Intensit	0000 - 5000 -					
	-	ment ion coun		ution		539	540	541	542	0	200 41		800	L h.e 1	000
		t search result	s T				n	√z				m/z			ion Complete (

Note: ensure that NO tag filters are applied and that Mascot is the selected search engine.

Determining protein identification is dependent on the availability of MS/MS data for the LC-MS runs. This data may be available but limited if the LC-MS was performed in a data dependant MS/MS detection mode due to under sampling. Under these conditions MS/MS data acquisition is dependent on thresholds and parameters set prior to the acquisition of the LC-MS run.

For this example we are using LC-MS runs containing MS/MS data where the data was acquired in a data dependant mode (DDA).

The Peptide Search page shows the number MS/MS that have been matched to each peptide ion in the Peptide ion list (see above). MS/MS scans are matched to a peptide ion if their precursor m/z and aligned retention time fall within the area of one of the isotopes of the peptide ion. The MS/MS scans which are matched to the displayed peptide ions are shown in the MS/MS spectra list on the right.

The first step is to decide which MS/MS scans you wish to export to be identified. By default this is all the available spectra for the peptide ions displayed in the Peptide ions list. This number is visible on the Export button.

The set can be targeted using the tags and also refined with respect to quantity and quality of the spectra being sent to the search engine.

Note: by default the table is ordered on the number of MS/MS spectra available for each peptide ion.

The total number of spectra included in this set is **37258** as shown on the Export button.

Before exporting the spectra, the set can be further refined.

Note: many of the abundant peptide ions have a large number of spectra associated with them. To control the number of spectra for each peptide ion, expand the **Batch inclusion options**.

Select the search program you're using:
Mascot 👻
Help
Export 37258 ms/ms spectra
MSMS Preprocessing
Limit fragment ion count
Deisotoping and charge deconvolution
Import search results
Clear identifications 🔹

Progenesis Q	l.p Tutorial for D	DA - Proge	enesis QI for pro	teomics										, .
File mport Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Ion Statistics	Identify Peptide		Refine Identificat		Protein Statistics Report			no	nlinea
			•			•		•		0 0			A Wat	ters Compan
dentify Pep	tides			MS/MS Sp	ectra									
	searches by e intification proc		eak list files	 Batch in 	nclusion opti	ons for creati	ng export list of	ms/ms spect	tra					
🐨 No filte	r applied	0]		Rank	greater than	• 5		Peptide ion	intensity less than	•			
T			Create	Pe	ptide ion # (less than	•		Precursor	intensity less than	•			
# MS/MS	Proteins	Tag	g 👻 Scor		Charge	less than	•		Precursor	intensity less than	•			
4 221	0		-	50	an number	less than	•			(%)				
152	0	1	-			icas than				Run name contains	•			
2 109	0	1			Exported	equal to	•	•	Peptide	sequence contains	•			
B 102 3 95	0				Isotope	less than	•		Protein	accession contains	•			
395 488	0		-		ID score	loss than			Protein d	escription contains	•			
+ 00 5 85	0		_		ID SCOLE (less triair	•		_					
7 80	0								I	nclude in export	xclude fro	om export	Clear all filters	
79	0													
75	0													
72	0	1	•	Export	-	Run	Scan number	Exported	Peptide ion intensity	Precursor intensity	(%)	Charge		Isotope
81 71	0	1	-	V		2873 C2 0049 C2	763 8852	No No	2.1e+005 9.1e+004	6.7e+004 4.5e+004	31.3 49.6	3	589.2769 687.3604	1
71	0	1	-			1487 C3	5400	No	6.3e+004	4.5e+004 8.9e+004	14.1	4	760.3506	1
67	0	l.	•			1707 A3	5567	No	3.7e+005	1.8e+005	48.0	2	808.4498	1
67	0	1	-	٠										÷
66	0			Peptide	ion numb	er 1, m/z	z 539.3004, r	etention t	time 29.63 min, d	harge +2				
4 65	0		-	Run:C2	Scan numb	er:3061		2.5E+06-				1		
	''' h program you'	re usina:												
ascot		•		·툰 28.987		_		2E+06-						
<u>elp</u>				E 28.98/	6		1	1.5E+06-						
Export 180	68 ms/ms spect	tra		ilii 28.987 iyi 28.987 30.994			Intensity	1E+06-		1				
ISMS Prepro				:블 30.994										
Limit fragn	nent ion count	40	A.	"				5E+05-	d i					
Deisoto	ping and charg	e deconvol			539	540 541	542	0	 200	400 600		800	1000	_
Import	search results					m/z		0	200	400 600 m/z		000	1000	
	lentifications	-											Section Cor	nnlete
CiedF IC	renditications													

For example: We will make use of the 'Rank' value to reduce the number of Spectra being used for each peptide ion to a maximum of 5.

The 'Rank' of each MS/MS spectra is determined by comparing its % value against all other spectra matched to the same peptide ion.

Export	Rank	#	Run	Scan number	Exported	Peptide ion intensity	Precursor intensity	(%)	Charge	Precursor m/z	Isotope	Id s
	E Th	001		FTac	found by co	mparing its '%' values a			d to the ca	201 0775	1	*
	5	404		zouz		z.je+000	5.4e+003	13.9	o to the sa	01. 01. 01. 01.	1	
	6	1800	A3	4618	No	2.9e+005	1.1e+005	39.4	2	783.8677	1	
	6	1756	C2	8216	No	2.9e+005	1.6e+005	54.4	2	729.8898	1	-
*												+

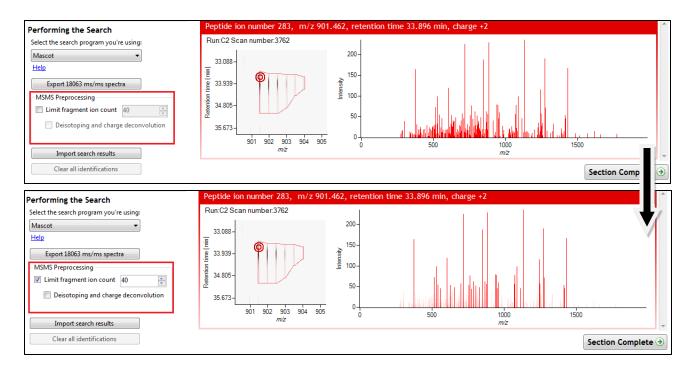
Note: the % value for each spectra is the Precursor intensity as a percentage of the Peptide ion intensity

Set the Rank filter to 'greater than' 5 and click **Exclude from export** this reduces the number to spectra to export to **18068**

Limiting the 'fragment ion count' (FIC) for the spectra being exported can improve the quality of the spectral data being used in the search by removing noisy peaks.

For example for the current spectra, reduce the FIC to 40.

Note: the effect this has on the number of peaks in the spectra. This 'limitation' is applied to all the spectra being exported; hence the export file size will be reduced.



For this example we will NOT limit the fragment count, so leave it un-ticked (the default setting).

Performing an MS/MS Ion Search

Having chosen 18068 spectra to export, as described above:

- 1. Select appropriate search engine i.e. Mascot
- 2. Click 'Export 18068 ms/ms spectra' to save search as file
- 3. Perform search on appropriate search engine and save results file
- 4. Click 'Import search results', locate results file and open

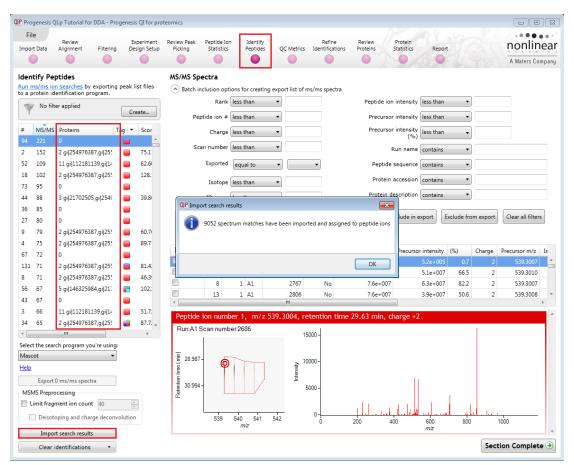
Please refer to Appendix 11 (page 95) for details of the 'Search Engine' parameters used in this example

Tip: For exporting mzIdentML of results to PRIDE where the Search Method is MascotSearch you **must select** additional **Mascot Export Fields** when exporting your search results from Mascot to Progenesis: (Protein sequence (Protein Hit Information) and Start and End (Peptide Match Information)), **Appendix 11 (page 95)**

Select the search program you're using:	
Mascot 👻	
Mascot	
PLGS (*.xml)	
SEQUEST (dta & out files) SEQUEST (dta & pepXml files)	
Phenyx	
Spectral Library Search	A
Ion Accounting	v
Deisotoping and charge deconverte	olution
Import search results	
Clear identifications 🔹	
Select the search program you're using:	
Mascot 👻	
Help	
Export 18068 ms/ms spectra	
MSMS Preprocessing	
Limit fragment ion count 40	A. V
Deisotoping and charge deconvolution	ution
Import search results	
Clear identifications 🔹	

Note: the blue link provides you with details on the appropriate formats for exporting search results and access to additional formats

Note: an example Search Results file, from a MS/MS lon search, is available in the folder you restored the Archive to (Search Results.xml). Select the 'Mascot' method and import this file to see results like those below.

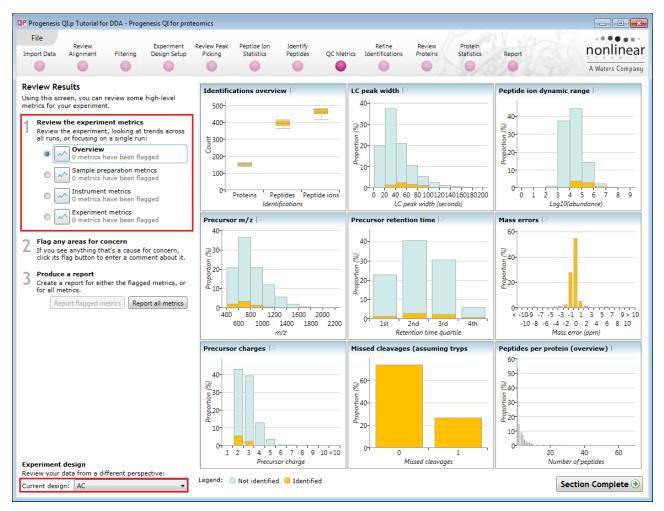


On importing the Search results the Peptide ions table updates to reflect the identified proteins and the relevant score for each searched peptide ion.

In order to review the quality of the **Peptide Search** results click on the next stage in the workflow, **QC Metrics**.

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.

<u>Metrics</u>

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.

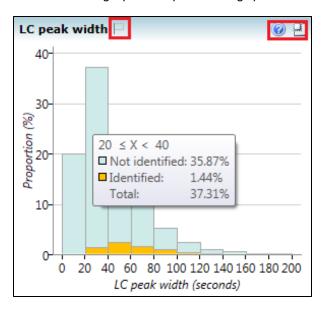
Interpretation and use

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):

Missed cleavages 🏳 PTMs 🏳 Abundance dynamic range 🏳	Identificatio	ons overview 🗎	(
---	---------------	----------------	---

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:

3 Produce a report Create a report for either the flagged metrics, or for all metrics. Report flagged metrics Report all metrics

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**.

2

Stage 11: Refine Identifications

In this example the organism under study is Clostridium difficile

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 40
- · Remove identifications where less than 2 hits were returned
- Remove all identifications where the Protein Description Contains 'hypothetical'
- Remove all identifications where the Protein Description **Contains** the following: 'Like', 'Putative', 'Probable', 'Potential', 'Predicted' and 'Partial'
- Remove all identifications where the Protein Description Doesn't contain 'Clostridium difficile'

File																	
	F	Review		Experiment	Review		Peptide Ior		Identify			Refine	Revie			nonli	ine
port	Data Ali	ignment	Filtering	Design Setup	Picki	ng	Statistics		Peptides	QC Metric	s Ider	tifications	s Protei	ins Statistics	Report	D Y N A	M I
		-							-			•			18018	A Waters	Comp
efin	e Identif	ications			Bate	h delet	ion criteria										
			ns include un emove them				[Score	less than	•	40			Sequence Ler	gth less than	•	
			tion criteria					Hits	less than	•				Cha	rge less than	•	
				ar the property				Mage	(Lana Albana	•				Econo		•	
	lues for a : lete.	set of iden	tifications ye	ou want to				Mass		•				Seque	nce contains		
					Abs	olute n	hass error	(ppm)	less than	•				Access	sion contains	•	
			identificat	tions on't want, click				m/z	less than	•				Descrip	tion contains	•	
	her:	e identifica		and wanty cilCK			Retention	Time	less than	•				Modificati	ons contains	•	
		Aatching Solighted ID:		s, to delete			Recention		less than		_			Hodinead	contains		
	Delete N	von-matchi	ng Search R									Delete m	atching se	arch results	Delete non-match	ing search results Reset the c	criteri
	delete ti	he IDs that	are <i>not</i> hig	hlighted		#	Score	Hit	s m/z	RT(mins)	Charge	Mass	Mass err	Sequence	Accession	Modifications	
			start again			97	73.51	5	668.38	20.25	2	1334.75		Sequence KVYLAGGV	Accession gi[1267004	precursor of the	e S-la
			ch of identifi Criteria and	cations to then return to		97	73.51	5	668.38	20.25	2	1334.75		KVYLAGGV	gi 226/66/iiii	S-layer protein	
	ep 1 above					97	73.51	5	668.38	20.25	2	1334.75		KVYLAGGV	gil2556567	cell surface pro	÷
-	No filter a	nolied				99	59.32	5	528.82	39.90	2	1055.62	-0.39	S VLELNELVK	gi[1503921	50S ribosomal	prote
Ψ.	no mer u	ipplied	(Create		99	59.32	5	528.82	39.90	2	1055.62	-0.39	VLELNELVK	gi 1266976	50S ribosomal	prote
	-					99	59.32	5	528.82	39.90	2	1055.62	-0.39	S VLELNELVK	🎯 gi 1583195	50S ribosomal	prote
	Total Hits			arge Tag		101	39.30	5	825.69	44.90	4	3298.7:	1.03	🔇 ELIITPASA	🎯 gi 87239956	s-layer protein,	, parti
	60		29.63 2			101	39.30	5	825.69	44.90	4	3298.73	1.03	🔮 ELIITPASA	🕥 gi 2067250	S-layer protein	A, pa
	10		54.88 3			101	39.30	5	825.69	44.90	4	3298.73	1.03	🔮 ELIITPASA	gi 92380869	s-layer protein,	parti
	55	598.321				101	39.30	5	825.69	44.90	4	3298.73	1.03	🔮 ELIITPASA	🕥 gi 2549763	cell surface pro	otein
	10	624.291				104	32.24	1	456.56	16.62	3	1366.67	-0.02	S LAGDDRYE	🕥 gi 2551019	cell surface pro	otein (
	35	604.337				104	32.24	1	456.56	16.62	3	1366.67	-0.02	S LAGDDRYE	🕥 gi 2549763	cell surface pro	otein
	15	626.314	38.94 2	•		108	114.64	5	1116.1(60.53	2	2230.18	0.33	AFVVGGTG	🎯 gi 70632806	S-layer protein	prec
	18		46.60 2	•	V	108	114.64	5	1116.1(60.53	2	2230.18		AFVVGGTG	🎯 gi 2549763	cell surface pro	
	10	462.27(•	V	108	114.64	5	1116.1(60.53	2	2230.18		AFVVGGTG	🎯 gi 2551019	cell surface pro	
	12	753.828		•		109	64.89	5	882.42	48.24	4	3525.6(DLTGASAD	il87239954	s-layer protein,	
0	20	595.319	36.87 2			109	64.89	5	882.42	48.24	4	3525.6(DLTGASAD	gi 87239952	s-layer protein,	
	18	573.803	24.19 2	•		109	64.89	5	882.42	48.24	4	3525.6(DLTGASAD	🥥 gi 2551019	cell surface pro	
-	60	573.324	41.74 2	•		112	62.69	5	702.86	29.22	2	1403.7(EVENELKD	🥥 gi 2549763	cell surface pro	
1 2	30	656.861	44.14 2			112	62.69	5	702.86	29.22	2	1403.7(-0.54	SVENELKD	🔮 gi 11496150	SlpA [[Clostridit	um] (
-	30																

On the Batch detection options panel, set the Score to less than 40, 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 (554 matching out of 2968)

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

Click Yes.

OP Dele	ete 554 search results?
?	Are you sure you want to permanently delete 554 peptide search results?
	Yes No

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

Then in the 'Description' field enter '**hypothetical**' and delete matching search results. Then also in Description field enter the 'regular expression': **regex: like|puta|prob|pote|pred|part** and delete matching search results.

File	2	Review				Review		Peptide Ior		Identify			Refine	Revie		rotein			
nport	t Data 🖌	lignment	Filterir		Experiment Design Setup	Picki		Statistics		Peptides	QC Metr	ics I	dentifications			atistic	Report		nonline
•		•	•		•	•				•	•		•	•		•	101		A Waters Comp
efin	e Identi	fications				Bate	h dele	tion criteria											
		dentificatio s, you can							Score	less than	•			Se	quence Le	ngth	less than 🔻		
		set of dele							Hits	less than	•				Ch	arge	less than 🔹	1	
Ir	n the batc	h deletion o	riteria, e	enter th	e property				Mass	less than	•				Segu	ence	contains 🔻	i	
	alues for a elete.	set of ider	ntification	ns you	want to														
				_		Abs	olute r	nass error	(ppm)	less than	•				Acces	sion	contains 🔹	J	
		unwante ne identifica			ns want, click				m/z	less than	•				Descrip	otion	contains 🔻	regex: like putz	a prob pote pred pa
	ither:							Retention	. Time	less than	•				Modificat	ions	contains 🔻	1	
		Matching S hlighted ID		esults, I	to delete							0							
	Delete	Non-match the IDs tha	ing Sear										Delete ma	atching se	arch results		Delete non-match	ng search results	Reset the criter
	uelete	the tos the	it are not	c myning	gnieu			elete 737 sea	irch re	sults?					23		Accession	Modifications	
		criteria to														DNK	gi 2549763	(ell surface protein
d	elete, clici				ions to en return to		?	Are you	sure)	ou want to p	permanen	tly del	ete 737 pept	tide search	results?	DNK	gi[71732944		-layer protein [Clo
st	tep 1 abov	e.						~								DNK	gi 1121811	5	layer protein, part
_	No filter	applied		_												рик	🕥 gi 1463259	5	-layer protein, part
Ŧ		applied		_	Create											DNK	🕥 gi 87239956	5	layer protein, part
				~	-								Yes		No	DNK	🎯 gi 2067250	S	S-layer protein A, pa
	Total Hi		RT (min		je lag											DNK	🎯 gi 21702505	S	S-layer protein [Clos
	60	539.30(29.63	2	_		1	59.57	5	539.30	29.63	2	1076.59	-0.37	S LLFTQ	/DNK	🎯 gi 2067250	5	S-layer protein A, pa
2	10	805.441		3			-	59.57	5	539.30	29.63	2	1076.59	-0.37	S LLFTQ	/DNK	🎯 gi 87239954	s	-layer protein, part
	55	598.321		2				59.57	5	539.30	29.63	2	1076.59	-0.37	S LLFTQ		🎯 gi 2551019	c	cell surface protein
	10	624.291		2			-	59.57	5	539.30	29.63	2	1076.59	-0.37	S LLFTQ	/DNK	gi 92380869	5	-layer protein, part
	35	604.337		2			-	59.57	5	539.30	29.63	2	1076.59	-0.37	S LLFTQ	/DNK	🕥 gi 87239952	5	-layer protein, part
	15	626.314		2	•			75.17	5	805.44	54.88	3	2413.3(2.48	LVSPA		🎯 gi 2551019	c	cell surface protein
	18	663.869	46.60	2			2	75.17	5	805.44	54.88	3	2413.3(2.48	LVSPA	PIVL	🎯 gi 2549763	c	cell surface protein
	10	462.27(34.33	2			3	51.72	5	598.32	30.51	2	1194.63	-0.21	YQVVI		🎯 gi 1463259	ş	S-layer protein, part
	12	753.828	30.46	2			3	51.72	5	598.32	30.51	2	1194.63		VQVVI		🔮 gi 87239956		-layer protein, part
0.	20	595.319	36.87	2				51.72	5	598.32	30.51	2	1194.6		S YQVVI		😒 gi 92380869		layer protein, part
1	18	573.803	24.19	2			-	51.72	5	598.32	30.51	2	1194.63		YQVVI		🎯 gi 2551019	c	cell surface protein
2	60	573.324	41.74	2			3	51.72	5	598.32	30.51	2	1194.63		YQVVI		🔮 gi 2067250		S-layer protein A, pa
4	30	656.861	44.14	2	📲 👻 .		3	51.72	5	598.32	30.51	2	1194.63	-0.21	S YQVVI	YPE	🔮 gi 87239952	5	-layer protein, part
					- F	4 =													• •

Finally alter the Description to 'doesn't contain Clostridium difficile' and delete the matching search results.

values for a set delete. Delete the unv To delete the id- either: Polete Mar Delete Mar Delete Nor Delete N	Ations Filt ations include the second	e them here. riteria , enter the proper- ions you want to tifications you don't want, cl Results, to delete tarch Results, to not highlighted again dentifications to ia and then return Create	etup Pir erty A dick te			Score Hits Mass (ppm) m/z n Time	less than less than less than less than less than less than s m/z belete 177 se	RT(mins arch result	Delete r Charge s?	Refine ientifications matching sear ye Mass Mi o permanently	rch results	Sequence Accession Description Modifications	Less than contains contains doesn't contain contains on-matching searc Accession 2505 2505	clostridium difficile
efine Identifica your peptide ident results, yo Specify action To delete the id either: Delete Mat To delete the id either: Delete Mat To delete the id either: Reset the crite To specify anoth delete, click Res step 1 above. No filter appl Total Hits m Total Hits m Total Hits m S 55 5 1 10 6 5 35 6 7 18 6	ations tifications incloue ou can remov of deletion critei eletion critei eletion critei eletion criteia wanted iden dentifications training Search hited 10s n-matching Search hited 10s n-matching Search there batch of i esert the Criter eletia to start there batch of i esert the Criter eletia m/z RT (m	ude unwanted or e them here. riteria (enter the proper ions you want to tifications you don't want, cl Results, to delete earch Results, to ob highlighted again dentifications to ia and then return Create	erty dick te	## # # # # # # 1 1 1 1 1 1 2	e mass error Retention \$9.57 \$9.57 \$9.57 \$9.57 \$9.57 \$9.57 \$9.57	Score Hits Mass (ppm) m/z n Time	less than less than less than less than less than less than less than less than		Delete r Charge s?	matching sear ye Mass Ma	See	uuence Length Charge Sequence Accession Description Modifications Cequence	less than less than contains contains doesn't contain contains con	clostridium difficile clostridium difficile definitions cell surface S-layer prot S-layer prot cell surface
your peptide ident results, set of In the batch del values for a set delete. Delete the um To delete the id either: Polete Neum To delete the id polete Neum Reset the crite To specify anoth delete, click Res step 1 above. No filter appl Total Hits m Total Hits m Total Hits m 10 6 5 5 5 5 15 6 7 18 6	tifications incl ou can remov of deletion of eletion criteria t of identificat wanted iden dentifications tching Search hited 10s n-matching Se IDs that are / tors to start ther batch of i esset the Criter olied m/z RT (m	e them here. riteria , enter the proper- ions you want to tifications you don't want, cl Results, to delete tarch Results, to not highlighted again dentifications to ia and then return Create	r A dick te	Absolute # ✓<	Retention Retention Score 59.57 59.57 59.57 59.57 75.17	Score Hits Mass (ppm) m/z n Time	less than less than less than less than less than less than s m/z belete 177 se		Charge s?	ge Mass Ma	rch results	Charge Sequence Accession Description Modifications	less than contains contains doesn't contain contains cont	clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile
relevant results, yo Specify a set of In the batch del values for a set delete. Delete the unv To delete the id either: Polete Mat To belete Mat To percify another To the crite To appediy another To tal Hits m Total Hits m Total Hits m 10 6 5 5 5 5 15 6 7 18 6	ou can remov of deletion c lettor, criteria t of identificat wanted iden dentifications tching Search hted IDs IDs that are <i>i</i> teria to start there batch of i search there batch of olied	e them here. riteria , enter the proper- ions you want to tifications you don't want, cl Results, to delete tarch Results, to not highlighted again dentifications to ia and then return Create	r A dick te	Absolute # ✓<	Retention Retention Score 59.57 59.57 59.57 59.57 75.17	Score Hits Mass (ppm) m/z n Time	less than less than less than less than less than less than s m/z belete 177 se		Charge s?	ge Mass Ma	rch results	Charge Sequence Accession Description Modifications	less than contains contains doesn't contain contains cont	clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile clostridium difficile
Specify a set of In the batch del values for a set delete. Delete the unit To delete the idi- either: Delete Mark the highling' Delete Mark To specify anoth delete, click Res step 1 above. Who filter appl No filter appl Total Hits Im 60 5 2 10 8 3 55 5 4 10 6 5 35 6 7 18 6	of deletion c leletion criteria t of identificat wanted iden dentifications tching Search hited IDs n-matching Se teria to start there has the criter bis set the Criter olied m/z RT (m	riteria , enter the proper ions you want to tifications you don't want, cl Results, to delete earch Results, to delete earch Results, to chighlighted again dentifications to ia and then return	A click te	# 1 1 1 1 1 1 1 2 1 2	Retention 59.57 59.57 59.57 59.57 59.57 75.17	Mass (ppm) m/z n Time	less than less than less than less than less than s m/z Delete 177 se	RT(mins arch result	Charge s?	ge Mass Ma	ass ern	Sequence Accession Description Modifications	contains contains doesn't contain contains on-matching searc Accession 23 763 2505 ch results? 2944	clostridium difficile clostridium difficile clostridium difficile Modifications cell surface S-layer prot S-layer prot cell surface
values for a set delete. Delete the unv To delete the id- either: Polete Mar Delete Mar Delete Nor Delete N	t of identificat wanted iden dentifications tching Search hited IDs n-matching Se IDs that are / teria to start ther batch of sest the Criter plied m/z RT (m	tifications you don't want, ci Results, to delete earch Results, to obt highlighted again dentifications to ia and then return Create	A click te	# 1 1 1 1 1 1 1 2 1 2	Retention 59.57 59.57 59.57 59.57 59.57 75.17	(ppm) m/z n Time Hits	less than less than less than s m/z Delete 177 se		Charge s?	ge Mass Ma	ass ern	Accession Description Modifications	doesn't contains contains contains an-matching searc Accession 2505 ch results? 2944	clostridium difficile clostridium difficile Modifications cell surface S-layer prot S-layer prot cell surface
Delete the un Dolete the un Dolete the idelet the idelity • Delete Nate • Delet	dentifications tching Search hted IDs n-matching Se IDs that are / theria to start ther batch of i sset the Criter	you don't want, ci Results, to delete earch Results, to oto highlighted again dentifications to ia and then return <u>Create</u>	click te	# 1 1 1 1 1 1 1 2 1 2	Retention 59.57 59.57 59.57 59.57 59.57 75.17	m/z n Time Hite	less than less than rs m/z Delete 177 se	RT(mins	Charge s?	ge Mass Ma	ass ern	Description Modifications	doesn't contain contains on-matching searc Accession 23 763 2505 ch results? 2944	clostridium difficile control of the criteri Modifications cell surface S-layer prot S-layer prot cell surface
To delete the identification either: • Delete Mats • Delete Mats • Delete Mats • Delete Mats • To specify anoth • delete, cick Res • To filter appl • No filter appl • Total Hits • Total Hits • Total Hits • Total S 55 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 35 • 10 • 10 • 10 • 10 • 10 • 10 • 10 • 10 • 10 • 10 • 10 • 10	dentifications tching Search hted IDs n-matching Se IDs that are / theria to start ther batch of i sset the Criter	you don't want, ci Results, to delete earch Results, to oto highlighted again dentifications to ia and then return <u>Create</u>	te rn to	✓ 1 ✓ 1 ✓ 1 ✓ 1 ✓ 1	 Score 59.57 59.57 59.57 59.57 75.17 	n Time Hite	less than s m/z Delete 177 se	RT(mins	Charge s?	ge Mass Ma	ass ern	Modifications	contains on-matching searc Accession 83 763 2505 ch results? 2944	n results) Reset the criteria Modifications cell surface S-layer prot S-layer prot cell surface
Delete Max the highligh Delete Nor	htted IDs n-matching Se IDs that are / teria to start ther batch of i sset the Criter olied m/z RT (m	arch Results, to hot highlighted again dentifications to ia and then return Create	rn to	✓ 1 ✓ 1 ✓ 1 ✓ 1 ✓ 1	 Score 59.57 59.57 59.57 59.57 75.17 	Hite	s m/z Delete 177 se	RT(mins	Charge s?	ge Mass Ma	ass ern	Delete no	Accession Accession 763 2505 ch results? 2944	Modifications cell surface S-layer prot S-layer prot cell surface
the highling • Delete Non- delete the I • Delete Non- delete, click Res step 1 above. • No filter appl • Total Hits m • 60 5 2 10 8 3 55 5 4 10 6 5 35 6 5 15 6 7 18 6	htted IDs n-matching Se IDs that are / teria to start ther batch of i sset the Criter olied m/z RT (m	arch Results, to hot highlighted again dentifications to ia and then return Create	rn to	✓ 1 ✓ 1 ✓ 1 ✓ 1 ✓ 1	59.57 59.57 59.57 59.57 59.57 75.17		elete 177 se	RT(mins arch result	Charge s?	ge Mass Ma	ass ern	Sequence	Accession 23 763 2505 2944	Modifications cell surface S-layer prot S-layer prot cell surface
Reset the crit To specify another delete, click Restep 1 above. No filter appl F Total Hits m 2 10 8 3 55 5 4 10 6 5 35 6 5 15 6 7 18 6	teria to start ther batch of i eset the Criter olied m/z RT (m	again dentifications to a and then return Create	rn to	✓ 1 ✓ 1 ✓ 1 ✓ 1 ✓ 1	59.57 59.57 59.57 59.57 59.57 75.17		elete 177 se	arch result	s?				2505 2944	cell surface S-layer prot S-layer prot cell surface
To specify another delete, circh et delete, circh e	ther batch of i eset the Criter olied m/z RT (m	dentifications to a and then return Create	rn to	✓ 1 ✓ 1 ✓ 1 ✓ 2	59.57 59.57 59.57 75.17	QP D				o permanently	/ delete 1	77 peptide searc	2505 th results? 2944	S-layer prot S-layer prot cell surface
delete, ciick Res step 1 above.	eset the Criter olied m/z RT (m	Create	rn to	✓ 1 ✓ 1 ✓ 2	59.57 59.57 75.17		Are you	u sure you	want to	o permanently	/ delete 1	77 peptide searc	th results? 2944	S-layer prot cell surface
Total Hits m 60 5 2 10 8 3 55 5 4 10 6 5 35 6 5 15 6 5 15 6 7 18 6	m/z RT (m			✓ 1 ✓ 2	59.57 75.17	()	Are you	ı sure you	want to	o permanently	/ delete 1	77 peptide searc		cell surface
Total Hits m 60 5 2 10 8 3 55 5 4 10 6 5 35 6 5 15 6 7 18 6	m/z RT (m			2	75.17								019	
60 5 2 10 8 3 55 5 4 10 6 5 35 6 5 15 6 7 18 6														
60 5 2 10 8 3 55 5 4 10 6 5 35 6 5 15 6 7 18 6		in Charge Tag		/ 2	75.17								019	
60 5 2 10 8 3 55 5' 4 10 6 5 35 6' 5 15 6 7 18 6'													763	cell surface
2 10 8 3 55 5 4 10 6 5 35 6 5 15 6 7 18 6		2		/ 3	51.72							Yes	No 763	cell surface
8 55 5 4 10 6 5 35 6 5 15 6 7 18 6	805.441 54.88		_	/ 3	51.72	-							2505	S-layer prot
4 10 6 5 35 6 5 15 6 7 18 6	598.321 30.51	-			51.72	5	598.32		2			YQVVLYPE	🥥 gi 2551019	cell surface
5 35 6 5 15 6 7 18 6	624.291 27.63			✓ 4	89.77	5	624.29		2	1246.5; 0.		DVENAIDD	🥥 gi 2549763	cell surface
5 15 6. 7 18 6				✓ 4	89.77	5	624.29		2	1246.57 0.		DVENAIDD	🥥 gi 2551019	cell surface
7 18 6	604.337 26.33	-		/ 5	76.10	5	604.34		2	1206.6€ 0.		VYLAGGV	🎯 gi 11496150	SIpA [[Clost
	626.314 38.94			✓ 5	76.10	5	604.34	26.33	2	1206.6€ 0.		VYLAGGV	🎯 gi 1267004	precursor o
3 10 4	663.869 46.60	2	_	/ 5	76.10	5	604.34	26.33	2	1206.6€ 0.	.02 🧃	VYLAGGV	🎯 gi 2549763	cell surface
	462.27(34.33	2	•		76.10	5	604.34	26.33	2	1206.6€ 0.		VYLAGGV	🎯 gi 2551019	cell surface
9 12 7	753.828 30.46	2 🧧	_	/ 5	76.10	5	604.34	26.33	2	1206.6€ 0.	.02 🧃	VYLAGGV	🎯 gi[70632806	S-layer prot
10 20 5	595.319 36.87	2 🧧	e 🛛	/ 5	76.10	5	604.34	26.33	2	1206.6€ 0.	.02 🧃	VYLAGGV	🎯 gi 2556567	cell surface
11 18 5	573.80: 24.19	2 🧯	Image: Second	/ 5	76.10	5	604.34	26.33	2	1206.6€ 0.	.02 🧃	VYLAGGV	🎯 gi 2964523	S-layer prot
L2 60 5	573.324 41.74	2		✓ 6	47.91	5	626.31	38.94	2	1250.61 0.	.21 🧃	DLKDYVD	🎯 gi 2549763	cell surface
L4 30 6	656.861 44.14	2	e 5	✓ 6	47.91	5	626.31	38.94	2	1250.61 0.	.21 🧃	DLKDYVD	🎯 gi 2551019	cell surface
10 5	- 42 00: 01 0C	-		/ 6	47.91	5	626.31	38.94	2	1250.61 0	.21 🧯		ail21702505	S-laver prot
			-					m						

Having applied all the filters there will be 1180 search results remaining

To validate the Peptide search results at the protein level click on Resolve Conflicts (bottom left).

Resolving 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 Resolve Conflicts stage is now accessed from the bottom left of the Refine Identifications stage.

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.

Details on performing conflict resolution are available in Appendix 8 (page 86)

For this guide: we will NOT resolve the conflicts.

The following pages describe the handling of Protein grouping and Protein Quantitation. Options for these are available at both Resolve Conflicts and Review Proteins

Note: the default settings are to **Use Protein Grouping** and **Relative Quantitation using HiN** (peptides =3)

QP Protein quantitation options
Quantitation method:
Relative Quantitation using Hi-N
Number of peptides to measure per protein (N): 3
✓ <u>Use protein grouping</u> , e. hide proteins whose peptides are a subset of another protein s.
OK Cancel

Protein Grouping

Where proteins are identified containing the same peptides then they are effectively indistinguishable aside from the 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 a group where the protein with the greater number is deemed the 'patriarch' and appears in the proteins table.

With protein grouping switched on (default setting) protein groups, the additional members are indicated by a bracketed number located after the Accession number. Taking **flagellin subunit** as an example, there are 9 additional group members. There are no conflicts and all the peptides are uniquely assigned.

Experiment design:				•													
Prote 🚽 No filter a	applied				Рер	tide i	ons	of g	gi 25497	3900			🚽 No filter	applied	_		
Т			Creat	:e									T			Create	
Accession	Peptides	Unique	Conflict:	Sci 🔺		#	Σ	몓	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	-
🎯 gi 260682215 (+1)	24	11	19	2		74	٠	1	63.9	5	0.968	1669.888	-0.0967	42.4	3		
🎯 gi 126700407	17	17	0	1	V	138	٠	0	103	5	0.966	1669.887	-1.11	42.4	2		
🎯 gi 254973900 (+9) 🔉	15	15	0	1	V	148	٠	2	101	5	0.925	1230.609	-0.407	22.9	2		
🎯 gi 126698450 (+1) gi	254973900	- flagellin	subunit [Clostrid	ium diff	icile Q	CD-6	6c26	125	5	0.958	2317.115	0.167	38.9	2		
	10	10	0	Ψ 		175	•	3	60.9	3	0.954	2317.115	0.157	38.9	3	.	
1 Protein: gi 254973	900 fla	aollin o	ubunit I(Nostri	dium	liffici	0	CD									
No protein selecte		genn s		10301		Jinici	eu		00020]								
	n resolutio	n															
					Denti	de ier		6	offiction	proto	i						
Conflicting protein									nflicting	·							_
Accession	Peptide	s Unique	Conflic	t: P	#	Σ	Ę	5 S	core	Hits C	Correlation	Mass N	Mass error (p	RT (mins)	Charge T	ag	-
•				•	•		-										P.
€ Refine Identifica	ations	Protein o	options											Section	Complet	te	€

Now open the protein options and untick Use protein grouping to examine the effect of the process.

As grouping is switched off the grouped proteins appear with conflicts to the other group members and there will be no unique peptides for each group member.

oteins 🕎	No filter ap	plied		Create		Рер	tide i	ons	of gi 26	0682017			T	No filter appli	ed		Create
ccession	Peptides	Unique	Conflict:	Score	^		#	Σ	中 Sco	e Hits	Correlation	Mass	Mass error (p	RT (mins)	Charg	je Tag	÷ ,
gi 260682017 🜼	15	0	127	1.95E+03			1839	0	0 9	4	0.958	1407.656	0.593	38.7	2		1
gi 254973900	15	0	127	1.95E+03		V	887	0	0 1	95	0.968	1692.834	-1	20.7	2		4
gi 5668937	14	0	123	1.9E+03		V	3330	0	1 84	7 4	0.874	1423.649	-0.431	22.7	2		7
gi 261863741	14	0	121	1.76E+03		V	477	0	1 47	4 5	0.978	1692.835	-0.238	20.7	3		5
gi 10281485	10	0	94	1.27E+03		V	300	0	0 43	8 4	0.963	1676.838	-1.19	34.8	3		7
gi 126697810	10	0	91	1.19E+03		1	449	0	0 93	3 5	0.965	1676.838	-1.21	34.8	2		7
qi 10281487	8	0	75	965		v	148	0	2 1	1 5	0.925	1230.609	-0.407	22.9	2		3
		0	67	2.5E+03				~	0 1	5 5	0.958	2317.115	0.167	38.9	2		5
gi 21702505	26	0	07	2.36+05		\checkmark	157	0	0 1	, ,							
gi 73745726 Protein: gi 260 Protein: gi 254	5 0682017 1973900 Protein reso	0 flagelli flagelli lution	54 n subun n subun	662 it [Clostrid it [Clostrid	lium	diffi	175 cile C cile C	0 019 000	3 60 96]	9 3	0.954	2317.115	0.157	38.9	3		3
gij73745726 Protein: gij260 Protein: gij254 btide ion views P Conflicting pro	5 0682017 4973900 Protein reso	0 flagellin flagellin lution	54 n subun n subun e ion 183	662 it [Clostrid it [Clostrid	lium	diffi	175 cile C cile C de ior	o D19 CD	3 60 96] •66c26]	9 <u>3</u> "" 73900	0.954					-	4
gi 73745726 Protein: gi 260 Protein: gi 254 btide ion views	5 0682017 4973900 Protein reso	0 flagelli flagelli lution	54 n subun n subun e ion 183	662 it [Clostrid it [Clostrid	lium P	v diffi diffi Pepti	175 cile C cile C de ior	o D19 2CD	3 60 96] •66c26]	9 <u>3</u> "" 73900	0.954			38.9 RT (mins) 38.7			4
gi 73745726 Protein: gi 260 Protein: gi 254 stide ion views P Conflicting pro Accession gi 254973900	5 7 0682017 1973900 Protein reso oteins for Peptide	0 flagellin Iution peptide	54 n subun n subun e ion 183 Conflict	662 it [Clostrid it [Clostrid 39 : Protein Sc	lium P	veptive veptiveptive veptive veptive veptive veptive veptive veptiv	175 cile C cile C de ior	CD19 CCD is of CCD	3 60 96] -66c26] 9 gi 254 1 Score	9 3 III 73900 Hits (0.954	Mass I	Mass error (p	RT (mins)	Charge	Tag	 ▲ ▲
gi[73745726 Protein: gi]260 Protein: gi]254 stide ion views P Conflicting pro Accession gi]254973900	5 0682017 1973900 Protein reso oteins for Peptide 15	0 flagellin lution peptide s Unique 0	54 n subun n subun e ion 183 Conflict 127	662 it [Clostrid it [Clostrid 39 Protein Sc 1.95E+03	P	veptio # v 18 v 18 v 18 v 8	175 cile C cile C de ior Σ	CD19 CD2 NS of 0 0 0	3 60 66] 66c26] gi 254 1 Score 99	73900 Hits 0	0.954	Mass 1 1407.656	Mass error (p 0.593	RT (mins) 38.7	Charge 2	Tag	 Al 1.2
gi[73745726 "Protein: gi[260 Protein: gi[254 bide ion views P Conflicting pro Accession gi[254973900 gi[26682017	5 0682017 4973900 Protein reso Deteins for Peptide 15 • 15	0 flagellin flagellin lution peptide s Unique 0 0	54 n subun n subun ion 183 Conflict 127 127	662 it [Clostrid it [Clostrid :9 Protein Sc 1.95E+03 1.95E+03	P	vepti vepti	175 cile C cile C de ior 339 α 87 α		3 60 66c26] 66c26] 99 109	73900 Hits 0 4 5	0.954	Mass 1407.656 1692.834	Mass error (p 0.593 -1	RT (mins) 38.7 20.7	Charge 2 2	Tag	1.2 4.6
gi[73745726 " Protein: gi[260 Protein: gi[254 otide ion views P Conflicting pro Accession gi[254973900 gi[26082017 gi[5668937	5 26682017 4973900 Protein reso oteins for 15 15 14	0 flagellin flagellin lution peptide 0 0 0 0 0	54 n subun n subun : ion 183 Conflict 127 127 123	662 it [Clostrid it [Clostrid 39 Protein Sc 1.95E+03 1.95E+03 1.9E+03	P	 ✓ ✓	175 cile C cile C de ior 239 0 87 0 330 0	0 CD19 QCD as of P 0 0 0 1	3 6(96] -66c26] 99 109 84.7	73900 Hits 0 4 5 4	0.954	Mass I 1407.656 1692.834 1423.649	Mass error (p 0.593 -1 -0.431	RT (mins) 38.7 20.7 22.7	Charge 2 2 2	Tag	► AI ~ 1.2 4.6 7.3
gij73745726 Protein: gij260 Protein: gij254 titide ion views P Conflicting pro Accession gij254973900 gij260682017 gij5668937 gij261863741	5 1973900 Protein reso oteins for Peptide 15 14 14	0 flagellin flagellin lution peptide 0 0 0 0 0 0 0	54 n subun n subun ion 183 Conflict 127 127 123 121	662 it [Clostrid 19 Protein Sc 1.95E+03 1.95E+03 1.9E+03 1.76E+03		♥ ♦ diffi • diffi • diffi ♥ ♥ 18 ♥ ●	175 cile C cile C de ior Σ 339 0 87 0 330 0 77 0	CD19 CD19 CCD as of 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0	3 6(66] 66c26] 91 254 Score 99 109 84.7 47.4	73900 Hits 0 4 5 4 5	0.954 Correlation 1 0.958 : 0.968 : 0.874 : 0.978 : 0.963 :	Mass 1 1407.656 1692.834 1423.649 1692.835	Mass error (p 0.593 -1 -0.431 -0.238	RT (mins) 38.7 20.7 22.7 20.7	Charge 2 2 2 3	Tag	 ▲ AI 1.2 4.6 7.3 5.6
gij73745726 Protein: gij260 Protein: gij254 tide ion views P Conflicting pro Accession gij254973900 gij264973900 gij2668937 gij261863741 gij10281485	5 1973900 Protein reso Neterins for Peptide 15 14 14 10	0 flagellin flagellin flagellin flagellin flution peptide Unique 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	54 n subun n subun e ion 183 Conflict 127 127 123 121 94	662 F it [Clostrid it [Clostrid 19 Protein Sc 1.95E+03 1.95E+03 1.95E+03 1.95E+03 1.76E+03 1.27E+03	P	♥ ↓ diffi ↓ di	175 cile C cile C de ior Σ 339 0 87 0 330 0 77 0 00 0	CD19 2CD as of 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3 60 66] 66c26] 99 109 84.7 47.4 43.8	73900 Hits 0 4 5 4 5 4 5 4	0.954 Correlation 0 0.958 : 0.958 : 0.968 : 0.978 : 0.963 : 0.963 :	Mass 1407.656 1692.834 1423.649 1692.835 1676.838	Mass error (p 0.593 -1 -0.431 -0.238 -1.19	RT (mins) 38.7 20.7 22.7 20.7 34.8	Charge 2 2 2 3 3 3	Tag	 ✓ AI 1.2 4.6 7.3 5.6 7.6
gij73745726 Protein: gij260 Protein: gij254 tide ion views P Conflicting pro Accession gij254973900 gij254973900 gij2668937 gij2668937 gij261863741 gij10281485 gij126697810	5 90682017 1973900 Protein reso oteins for Peptide 15 14 14 10 10	0 flagellin flagellin lution peptide Unique 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	54 n subun n subun i ion 183 Conflict 127 123 121 94 91	662 it [Clostrid it [Clostrid 39 Protein Sc 1.95E+03 1.95E+03 1.9E+03 1.76E+03 1.27E+03 1.19E+03		♥ ♥ ● diffi ● diffi ● diffi ♥ ■ diffi ♥ ● ● ● ● ●	175 cile C cile C de ior Σ 339 87 330 77 00 49	CD19 CD19 CD19 CD19 CD 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3 60 66 66c26 9 1 5 5 66c26 9 1 1 5 1 1 5 1 1 1 1 1 1 1 1	73900 Hits 0 4 5 4 5 4 5 4 5	0.954 Correlation 0 0.958 : 0.968 : 0.968 : 0.978 : 0.963 : 0.965 : 0.925 :	Mass 1407.656 1692.834 1423.649 1692.835 1676.838 1676.838	Mass error (p 0.593 -1 -0.431 -0.238 -1.19 -1.21	RT (mins) 38.7 20.7 22.7 20.7 34.8 34.8	Charge 2 2 3 3 2 2	Tag	 ✓ AI ✓ AI ✓ 1.2 4.6 7.3 5.6 7.6 7.5
gij73745726 Protein: gij260 Protein: gij254 stide ion views Conflicting pro Accession gij254973900 gij264973900 gij2668937 gij26863741 gij10281485 gij126697810 gij10281487	5 2682017 1973900 Protein reso teins for Peptide 15 14 14 10 10 8	0 flagellin flagellin flution peptide 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	54 n subun n subun ion 183 Conflict 127 127 123 121 94 91 75	662 it [Clostrid it [Clostrid 39 : Protein Sc 1.95E+03 1.95E+03 1.76E+03 1.27E+03 1.19E+03 965		♥ ♥ ● diffi ● diffi ♥ ● ● <	175 cile C cile C de ior Σ 339 C 339 C 330 C 330 C 49 C 48 C	CD19 CD19 CD19 CD19 CD19 CD19 CD19 CD19	3 60 66[66c26] 99 109 84.7 47.4 43.8 93.3 101	73900 Hits 0 4 5 4 5 4 5 5 5	0.954 Correlation 0 0.958 : 0.968 : 0.978 : 0.963 : 0.963 : 0.965 : 0.925 : 0.958 :	Mass 1407.656 1692.834 1423.649 1692.835 1676.838 1676.838 1676.838 1230.609	Mass error (p 0.593 -1 -0.431 -0.238 -1.19 -1.21 -0.407	RT (mins) 38.7 20.7 22.7 20.7 34.8 34.8 22.9	Charge 2 2 3 3 2 2 2 2	Tag	 ▲ AI <l< td=""></l<>
gij73745726 Protein: gij260 Protein: gij254 stide ion views Conflicting pro Accession gij254973900 gij264973900 gij2668937 gij26863741 gij10281485 gij126697810 gij10281487	5 2682017 1973900 Protein reso teins for Peptide 15 14 14 10 10 8	0 flagellin flagellin flution peptide 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	54 n subun n subun ion 183 Conflict 127 127 123 121 94 91 75	662 it [Clostrid it [Clostrid 39 : Protein Sc 1.95E+03 1.95E+03 1.76E+03 1.27E+03 1.19E+03 965		▼ ▼ • diffi • diffi <td>175 cile C cile C de ior Σ 339 α 330 α 330 α <t< td=""><td>CD19 2CD as of 0 0 0 0 0 0 0 0 0 0 0 0 0</td><td>3 60 66 66 66 66 66 66 66 66 66</td><td>73900 Hits 0 4 5 4 5 4 5 5 5 5</td><td>0.954 Correlation 0 0.958 : 0.958 : 0.968 : 0.978 : 0.963 : 0.965 : 0.925 : 0.958 : 0.954 :</td><td>Mass 1 1407.656 1692.834 1423.649 1692.835 1676.838 1676.838 1230.609 2317.115</td><td>Mass error (p 0.593 -1 -0.431 -0.238 -1.19 -1.21 -0.407 0.167</td><td>RT (mins) 38.7 20.7 22.7 20.7 34.8 34.8 22.9 38.9</td><td>Charge 2 2 2 3 3 2 2 2 2 2</td><td>Tag</td><td> AI 1.2 4.6 7.3 5.6 7.5 3.1 5.6 </td></t<></td>	175 cile C cile C de ior Σ 339 α 330 α 330 α <t< td=""><td>CD19 2CD as of 0 0 0 0 0 0 0 0 0 0 0 0 0</td><td>3 60 66 66 66 66 66 66 66 66 66</td><td>73900 Hits 0 4 5 4 5 4 5 5 5 5</td><td>0.954 Correlation 0 0.958 : 0.958 : 0.968 : 0.978 : 0.963 : 0.965 : 0.925 : 0.958 : 0.954 :</td><td>Mass 1 1407.656 1692.834 1423.649 1692.835 1676.838 1676.838 1230.609 2317.115</td><td>Mass error (p 0.593 -1 -0.431 -0.238 -1.19 -1.21 -0.407 0.167</td><td>RT (mins) 38.7 20.7 22.7 20.7 34.8 34.8 22.9 38.9</td><td>Charge 2 2 2 3 3 2 2 2 2 2</td><td>Tag</td><td> AI 1.2 4.6 7.3 5.6 7.5 3.1 5.6 </td></t<>	CD19 2CD as of 0 0 0 0 0 0 0 0 0 0 0 0 0	3 60 66 66 66 66 66 66 66 66 66	73900 Hits 0 4 5 4 5 4 5 5 5 5	0.954 Correlation 0 0.958 : 0.958 : 0.968 : 0.978 : 0.963 : 0.965 : 0.925 : 0.958 : 0.954 :	Mass 1 1407.656 1692.834 1423.649 1692.835 1676.838 1676.838 1230.609 2317.115	Mass error (p 0.593 -1 -0.431 -0.238 -1.19 -1.21 -0.407 0.167	RT (mins) 38.7 20.7 22.7 20.7 34.8 34.8 22.9 38.9	Charge 2 2 2 3 3 2 2 2 2 2	Tag	 AI 1.2 4.6 7.3 5.6 7.5 3.1 5.6

All the group members will appear in both tables with large numbers of conflicts with the other group members.

Note: flagellin has **no unique** peptides as they are all present in flagellin subunit protein hence the reason for grouping. As a result all the conflicts are internal to the group.

Now set the Protein Options back to Use	
protein grouping.	

OP Protein quantitation options
Quantitation method:
Relative Quantitation using Hi-N
Number of peptides to measure per protein (N):
Use protein grouping i.e. hide proteins whose peptides are a subset of another protein's.
OK Cancel

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.

٩	P Protein quantitation options
	Quantitation method:
	Relative Quantitation using Hi-N
	Absolute Quantitation using Hi-N Requires a calibrant protein to calculate absolute amounts Uses the most abundant N peptides Allows comparison between proteins within a run
	Relative Quantitation using Hi-N • Uses the most abundant N peptides • Allows comparison between proteins within a run
	Relative Quantitation using non-conflicting peptides • Uses only peptides which have no conflicting protein identifications • Allows comparison of a single protein across runs
	Relative Quantitation using all peptides • Uses all peptides identified as part of a protein • Allows comparison of a single protein across runs

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.

OP Protein quantitation options
Quantitation method:
Relative Quantitation using Hi-N
Number of peptides to measure per protein (N): 3
Use protein grouping i.e. hide proteins whose peptides are a subset of another protein's.
OK Cancel

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.

		QP Protein quantitation options	
QP Protein quantitation options		Quantitation method:	
Quantitation method:		Absolute Quantitation using Hi-N	•
Absolute Quantitation using Hi-N		Number of peptides to measure per protein (N):	3 💌
Number of peptides to measure per protein (N):	N	Calibrant accession:	gi[1267007
Calibrant accession:		▲ The calibrant protein can't be found. How will measurements be calculated?	gil126700790 enolase (2-phosphoglycerate dehydratase) (2-phospho-D-q
Amount (fmol):		Amount (fmol):	gi[126700794 glyceraldehyde-3-phosphate dehydrogenase [Clostridium d
Use protein grouping i.e. hide proteins whose peptides are a subset of another protein's.		Employ protein grouping, i.e. hide proteins whose subset of another protein's.	
OK Cancel			K Cancel

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.

Review Proteins Using this screen, you can find the proteins of interest in your experiment.	No filter a	pplied Creat	e	ch 🔎		@ H	elp
Set the quantitation options	Max fold change	Highest Mean	Lowest Mean	Description	Amount (fmol) - A	Amount (fmol) - C	2
If you've not already done so, choose between relative and absolute guantitation, use of Hi-N,	3.64	Α	С	ferredoxin [Clostridium difficile 630]	19.8	5.43	-
protein grouping and more.	3.1			fructose-1,6-bisphosphate aldolase [Clostridium difficile 630]	1.37	4.24	1
Protein options	4.84	С	A	RNA-binding protein Hfq [Clostridium difficile 630]	4.72	22.8	
	9.58	Α	С	phosphatase, 2C family [Clostridium difficile 630]	1.23	0.128	1

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

QP Protein quantitation	options 💌
Quantitation method:	
Relative Quantitation	using non-conflicting peptides
Use protein groupin of another protein's	g i.e. hide proteins whose peptides are a subset
	OK Cancel

The relative Quantitation can also be performed using all peptides.

OP Protein quantitation options
Quantitation method:
Relative Quantitation using all peptides
Use protein grouping i.e. hide proteins whose peptides are a subset of another protein's.
OK Cancel

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.

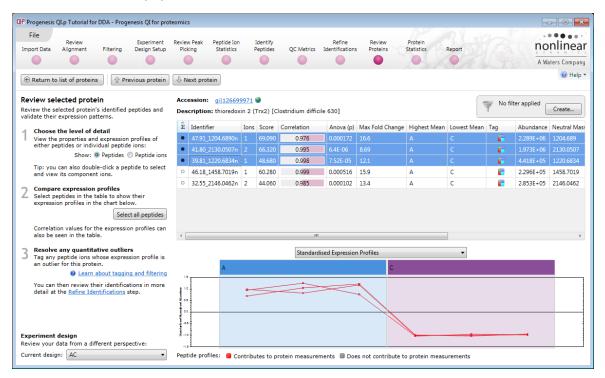
Stage 12: Review Proteins

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

P Progenesis QI.p Tutorial for DDA - Progenesis QI for prot	eomics										
File Review Experiment Import Data Alignment Filtering Design Setup	Review Peak Peptide Ion Picking Statistics	Ident Pepti		Refine s Identifications	Review Proteins	Prote Statist		Report		ņ	onlinea
							2	- 180	A L	A	Waters Compa
Review Proteins Jsing this screen, you can find the proteins of nterest in your experiment.	No filter applied	Creat	e Search		9						Help
Set the quantitation options	Accession	Peptides	Unique peptides	Confidence score	Anova (p)	q Value	Tag 🔹	Max fold change	Highest Mean	Lowest Mean	Description
If you've not already done so, choose between relative and absolute guantitation, use of Hi-N,	🎯 gi 126699939	3	3	236	2.56E-06	4.95E-06		8.66	A	с	transketolas
protein grouping and more.	gi 126699756	3	3	190	4.67E-06	7.73E-06		9.5	A	С	elongation (
Protein options	🔮 gi 126701103	1	1	51.7	7.08E-06	1.03E-05		6.07	A	С	ribose-5-ph
	🎯 gi 54781345 (+1)	5	5	398	7.97E-06	1.03E-05		7.65	A	С	(R)-2-hydro
Create a shortlist to review In the table, sort and filter the proteins based	🔇 gi 126699971	5	5	334	9E-06	1.04E-05	4	11.7	А	С	thioredoxin
on their measurements, to generate a shortlist	🔇 gi 126701179	3	Anova p-valu		1.42E-05	1.5E-05		5.17	A	с	transcription
for further review.	gi 126697690	6	Max fold char	ige ≥ 2	1.58E-05	1.5E-05		5.94	A	с	ferredoxin/f
How are the measurements calculated?	gi 126700634	3	New tag		1.69E-05	1.5E-05		6.49	A	с	PTS system
To sort the table by a given value, simply click the relevant column header.	gi 384359782 (+1)	24	Quick Tags	•	Anova p-	value		37	С	Α	hemagglutii
Benden the endeduce	gi 126699940	1	Edit tags		Max fold	change		7.53	A	с	transketolas
For each protein of interest, review its peptide	gi 126700297	2	Add to Clip G	allery	Modificat	ion		2.42	A	с	propanedio
measurements and correlations:	gi 254973900 (+9)	15	15	1.95E+03	Sequence			3.78	A	С	flagellin sub
View peptide measurements	gi 126697684	2	2	109	Peptide ta	ags contain		3.47	A	С	phosphate I
You can also double-click to review a protein.	gil254976383 (+2)	9	1	797	4.69E-05	2.74E-05		464	С	A	cell surface
Export data for further processing	4										Þ
By exporting your data to external tools, there's no limit to your analysis.	Selected protei		redoxin 2 (1	rx2) [Clostr	idium d	ifficile	630]				
Export to pathways tool			А						с		
Export protein measurements	- 16 -										
Export peptide measurements	V 15 -		} \$4								
	"16								康		
xperiment design eview your data from a different perspective:	4										
Current design: AC 🔹	Quantifiable proteins									Section (Complete 🤆

As an example let us explore Thioredoxin 2. The table indicates that this protein is most highly expressed in Condition A by 11.7 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** below the table.



The solid icon in Σ (click to order on) column indicates that the peptide contributes to protein measurements. **Note**: a correlation score for each peptide is available to facilitate the validation of peptide expression.

By selecting all the peptides you can compare the pattern of expression across all the samples allowing you to identify any 'atypical' behaviour assigned to the current protein.

File	Review	Experiment	Review Peak	Peptide Ion	Identify		Refine	Review	Protein			
mport Data	Alignment Filter		Picking	Statistics	Peptides	QC Metrics	Identificatio			Report		nonlinea
	•		•	•	•			•	0./	•		A Waters Compa
le Return to	list of proteins	⑦ Previous protein	🕀 Next pro	tein								🕢 Help
eview sel	ected protein		Accession	gi 12669997	11 🧐						S No filter app	lied
	elected protein's ident expression patterns.	ified peptides and	Descriptio	n: thioredoxin :	2 (Trx2) [Clost	ridium diffi	cile 630]				Ψ	Create
Choose	the level of detail		old Change	Highest Mean	Lowest Mean	Tag	Abundance	Neutral Mass	Retention Time (mins) Peptid	e Sequence	Modifications
View the	properties and expre							1204.689		S VL		
either pe	eptides or individual p	eptide ions: des © Peptide ions					1.973E+06		41.804	S VD	EVTKDDATVPNIENMIK	
_						-	4.418E+05	1220.6834		S VL	GLPTMAIYK	[7] Oxidation (M)
	can also double-click its component ions.	a peptide to select				-	2.296E+05	1458.7019		S DD	ATVPNIENMIK	
	e expression profile					- E		2146.0462		S VD	EVTKDDATVPNIEN <mark>M</mark> IK	[17] Oxidation (M
Correlati	on profiles in the char ion values for the exp seen in the table.	Select all peptides	4						m			
	any quantitative o					Standar	dised Expressi	on Profiler		•		
. Tan	peptide ions whose ex r for this protein.	pression profile is				Standar	uiseu expressi					
	1 C C C C C C C C C C C C C C C C C C C	agging and filtering	20	A					с			
								-				
an outlie You can	then review their ider the <u>Refine Identificat</u>		10 10 03	8					<u></u>			
an outlie You can detail at xperiment	the <u>Refine Identificat</u>	<u>ons</u> step.		8								

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.

Note: as the quantitation is based on the top 3 peptides, the remainder are shown grey on the graph

The behaviour of individual peptide ions can also be reviewed by selecting **peptide ions** as the level of detail to review.

File	Review		Experiment			Peptide Ior	n	Identify				Review	Protein	2			nonlinea
mport Data	Alignment	Filtering	Design Setup	PI	cking	Statistics		Peptides	QC Metrics	Identifi	cations P	Proteins	Statistics R	eport			DYNAMIC
												•		.		18	A Waters Compar
Return to	list of proteins	🕆 Pro	evious protein	Û	Next protei	in											🕜 Help
leview sele	ected protei	n		Ac	cession:	gi 12669	997	1 🕥							ſ	S No filter app	lied
	lected protein's expression patt		peptides and	De	scription:	thioredo:	xin 2	(Trx2) [Clos	tridium diffi	cile 630]						Y no much opp	Create
Choose	the level of de	tail		n	Lowest Me	an Tag	•	Abundance	m/z	Charge	Retention	Time (mins)	Mass error (ppm)	Pep	tide Se	equence	Modifications
View the	properties and	expression			с			2.296E+05	730.3582	2	46.185		-0.33	۲	DDAT\	/PNIENMIK	
either pe	ptides or individ		e ions: Peptide ions		С			1.402E+06	711.0242	3	41.823		-0.26	9	VDEVT	KDDATVPNIENMIK	
_		- 1 - -					1		1066.0325	2	41.785		-0.43	9	VDEVT	KDDATVPNIENMIK	
	can also double its component i		otide to select		с			2.17E+05	716.3561	3	32.622		0.03	9	VDEVT	KDDATVPNIEN <mark>M</mark> IK	[17] Oxidation (M
Company	e expression p	rofilos			с			6.836E+04	1074.0302	2	32.487		-0.22	9	VDEVT	KDDATVPNIENMIK	[17] Oxidation (M
Select pe	ptide ions in the	e table to s			с			2.289E+06	603.3518	2	47.914		0.04	9	VLGLP	TMAIYK	
expressio	on profiles in the				С			4.418E+05	611.349	2	39.811		-0.37	9	VLGLP	TMAIYK	[7] Oxidation (M)
		Select al	l peptide ions														
	on values for the		on profiles can							_							
also be s	een in the table			×									"				
	any quantitati								Standar	dised Exp	ression Prof	iles		•	•		
	r for this protein		sion prome is			Α						с			_		
	② Learn at	oout taggin	g and filtering			~											
	then review thei the Refine Ident										-						
uctail at	the <u>Kenne Idem</u>	ancations a	kep.	and and													
				Tellino.													
				adhed he													
xperiment (desian			2000	1.0										_		
	lata from a diffe	rent persp	ective:														

Note: peptide ions with different charge states are now displayed, these are deconvoluted at the peptide level.

Returning to the protein level, with protein grouping switched on (default setting) 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.

File Review Experiment R moort Data Alignment Filtering Design Setup	eview Peak Peptide Ion Picking Statistics	Ident Peptio		Refine Identifications	Review Proteins	Prote Statist		Report		n	onlinea
										A	Waters Compan
Review Proteins lasing this screen, you can find the proteins of iterest in your experiment.	No filter applied	Creat	e		Q						🕜 Help
Set the quantitation options	Accession	Peptides	Unique peptides	Confidence score	Anova (p)	q Value	Tag 💌	Max fold change	Highest Mean	Lowest Mean	Description
If you've not already done so, choose between relative and absolute quantitation,	🎯 gi 126697972	2	2	130	0.000386	0.000109	۷	2.79	с	Α	fructose-1,6
use of Hi-N, protein grouping and more.	🎯 gi 126699754	3	3	250	0.00038	0.000109	۹	2.11	A	с	ribosome re
Protein options	gi 254974605	4	4	311	0.000364	0.000109		2.66	С	Α	electron trar
	🎯 gi 126701091	3	3	236	0.000314	9.84E-05		3.25	С	Α	FOF1 ATP sy
Create a shortlist to review In the table, sort and filter the proteins	🕲 gi 254973854 (+4)	5	5	428	0.00027	8.69E-05	۵	2.92	С	А	60 kDa chap
based on their measurements, to generate a	🎯 gi 126698631 (+ _{gil}	254973854	- 60 kDa chaperor	nin [Clostridium dif	ficile OCD-6	56c261		_	С	А	cell wall-bin
shortlist for further review.	🚳 gi 126699299 gi	126697767	- 60 kDa chaperor	nin (Protein Cpn60)	(GroEL pro	tein) [Clost	ridium diffi	cile 630]	A	с	dinitrogenas
How are the measurements calculated?				in [Clostridium dif GroEL [[Clostridiur		43255]			A	с	4Fe-4S ferre
To sort the table by a given value, simply click the relevant column header.				n GroEL [[Clostridiu					A	с	cell surface
Review the proteins	🎯 gi 126697654	3	3	215	0.00023	8.4E-05		2.56	A	с	30S ribosom
For each protein of interest, review its	gi[126700198	1	1	46.6	0.000225	8.4E-05	۷	9.64	A	с	phosphatase
peptide measurements and correlations:	٠			M.							+
You can also double-click to review a protein.	Selected protei		Da chaperoi	nin [Clostrid	ium dif	ficile Q	CD-660	:26]			
Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	ing 14.0		A						с		
Export to pathways tool Export protein measurements	ArcSinh Normali		R.						j æ t		

To view members peptides click on View peptide measurements.

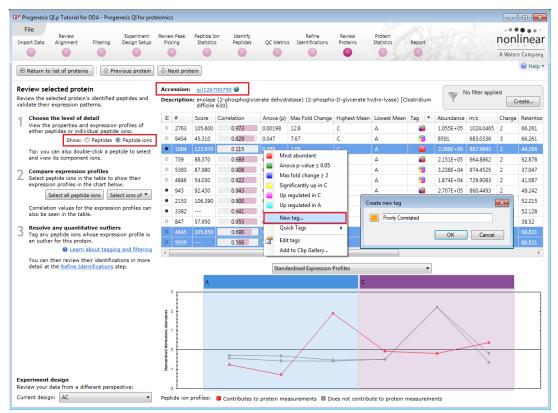
Click on **Select peptides of** to show the list of additional group members. Then as you click on one of the member proteins its peptides are highlighted in the table.

mport Data	Review Alignment	Filtering	Experiment Design Setup		w Peak Peptide Ion king Statistics		Identify Peptides	QC Metrics	Refine Identifications		Protein Statistics R	eport			non	linea
	•	•	•					•	•	•	•				A Wate	rs Compa
le Return to	list of proteins	Û	Previous protein	٩ ţJ	lext protein											🕜 Hel
Review sel	ected protei	n		Acc	ession: gi 25497	3854	۵						ĺ	🥯 No filte	r applied	
	elected protein's expression patt		ed peptides and	Des	scription: 60 kDa d	haper	ronin [Clos	tridium difficile	QCD-66c26]					Y		Create
4 Choose	the level of de	itail		Σ	Identifier	Ions	Score	Correlation	Anova (p)	Max Fold Change	Highest Mean	Lowest Mean	Tag	Abundance	Neutral Mass	Retent
View the	properties and	express		•	40.28_1531.8239n	1	58.830	0.976	4.49E-05	3.36	С	А	-	2.834E+05	1531.8239	40.285
either pe	ptides or individ		es O Peptide ions	٠	31.61_1237.7139n	1	66.390	0.962	0.00447	1.97	С	А	- 🔁	1.712E+05	1237.7139	31.611
Tin: you			peptide to select	•	53.74_2087.0860n	2	106.050	0.985	0.000267	3.22	С	Α		4.912E+05	2087.086	53.744
	its component		peptide to select	0	35.11_1069.6371n 71.79 2148.2152n	_	68.460 50.250	0.995	0.000651	25.8	с с	А		1.37E+05 3915	1069.6371	35.105
Select per expression	e expression p eptides in the ta on profiles in the elect all peptides	ble to s e chart	how their													
	on values for th seen in the table	. .	gi 254973854 60 kDa chaperoni QCD-66c26]	n [Clo	stridium difficile				III							
Resolve	any quantitat	ive	gil126697767					Sta	indardised Exp	ression Profiles		•				
an outlie	peptide ions wh r for this protein Q Learn a		60 kDa chaperoni protein) [Clostridi		tein Cpn60) (GroEL fficile 630]					C	c					
	then review the the <u>Refine Iden</u>		gi 255305190 60 kDa chaperoni ATCC 43255]	n [Clo	stridium difficile							_		_		
			gi 8886080 heat shock protei	n Groß	EL [[Clostridium]							/		$\overline{}$	-9	
			difficile]													

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. Example, at the Protein level of **Review proteins** use the search facility to locate the protein with accession gi|126700790.

Progenesis	QI.p Tutorial for	DDA - Proge	nesis QI for prot	teomics												×
File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking		Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Prot Stati		R	eport		A Waters Cor	ieal
	oteins een, you can fi ur experiment.		ins of	Vo No	filter applied	Create	gi 126700	790	×						0	Help 🔻
If you've relative	quantitation onot already do and absolute q grouping and m	ne so, choo: Jantitation, u		Accession	Peptic 00790 (+1) 9	des Unique 9		onfidence score 67		q Value 0.0118	Tag	_	Max fold change 1.84	Highest Mean C	Lowest Mean A	Desc enola

Then clear the search box (click on 'x') leaving the protein highlighted in the table. Double click on the protein and then click on Show: **Peptide ions** then select 3 of the poorly correlated



To delete the tagged peptide ions (Poorly Correlated). Click on **Refine Identifications**, then set up a tag filter to restrict the right hand table to only those peptide ions tagged with Poorly Correlated.

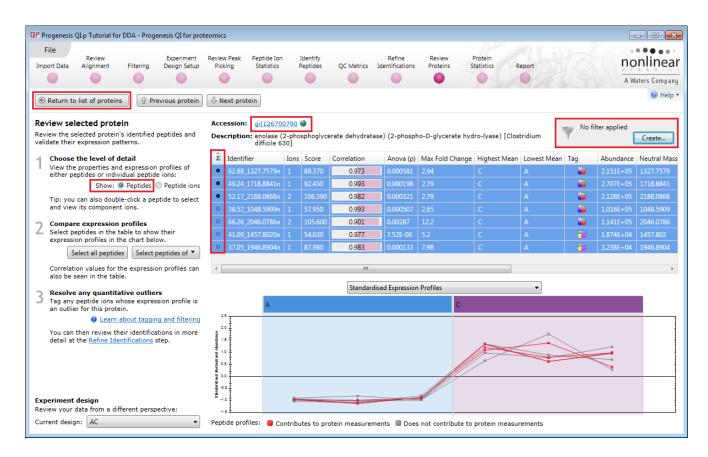
OP Progenesis QI.p Tutorial for DDA - Progenesis QI for prot	eomics		x
File Review Experiment Import Data Alignment Filtering Design Setup	Review Peak Picking	Peptide Ion Identify Refine Review Protein Protein Statistics Report	5
		A Waters Compar	ny
Refine Identifications	Batch dele	etion criteria	
If your peptide identifications include unwanted or irrelevant results, you can remove them here.	ſ	Conce long there a Converse Long the line there a	
Specify a set of deletion criteria		QP Filter the peptide ions	
In the batch deletion criteria, enter the property values for a set of identifications you want to delete.	Absolute	Create a filter Show or hide peptide ions based on a selection of their tags. Move tags to the appropriate boxes to create the filter. For more guidance, please see the <u>online reference</u> .	
2 Delete the unwanted identifications To delete the identifications you don't want, click		Available tags: Show peptide ions that have all of these tags:	
either:		Most abundant (2330 peptide ions)	
 Delete Matching Search Results, to delete the highlighted IDs 		Anova p-value ≤ 0.05 (7943 peptide ions)	
 Delete Non-matching Search Results, to delete the IDs that are not highlighted 		Max fold change ≥ 2 (9898 peptide ions) Show peptide ions that have at least one of earch results Reset the criteria	1
delete the 155 that are not highlighted	#	Significantly up in C (3937 peptide ions) these tags:	
3 Reset the criteria to start again To specify another batch of identifications to	934	Up regulated in A (4006 peptide ions) 50S ribosomal protein L	
delete, click Reset the Criteria and then return to	943	enolase [Clostridium dif	
step 1 above.	943	Hide peptide ions that have any of these tags: enolase (2-phosphoglyc	
Tag filter applied	952	cell surface protein [Clo:	
peptide ions may be hidden Edit	952	cell surface protein [Clo:	
# Total Hits m/z RT (min Charge Tag *	961	cell surface protein [Clo:	
1004 8 887.969 44.27 2	 ✓ 961 ✓ 962 	Clear the filter OK Cancel Clear the filter OK Cancel OK	
4645 10 974.001 60.83 2	962	46.54 3 527.96 38.76 3 1580.8' -0.41 W TALODAVT W gil2095707 cell wall protein V I/Clos	3
9939 649.674 60.83 3 🥃	963	46.54 3 527.96 38.76 3 1580.8: -0.41 • TALQDAVT • gij265767 tein wan protein v (jotos	
	973	42.67 2 514.33 24.10 2 1026.64 -0.38 🕥 NALVEIIKK 🕥 gi[2964514 5-nitroimidazole antibio	

Then highlight all three entries in the left-hand table, the corresponding rows will be selected on the righthand table. Right click on the highlighted rows and Delete selected peptides(s)

efine Identi your peptide in elevant results Specify a s In the batch	dentification s, you can n		Experiment Design Setup	Review F Pickin		Peptide Ion Statistics		lentify ptides	OC Metric	Ref		Revie	w Protein			nonline
efine Identi your peptide in elevant results Specify a s In the batch	ifications dentification s, you can r	is include un			g		Pe	epildes				Protei	ns Statistics	Denest		nonlinea
your peptide is elevant results Specify a s In the batch	dentification s, you can n			Detal						s Identifi	cations	Protei	ns Statistics	Report		DYNAMIC
your peptide is elevant results Specify a s In the batch	dentification s, you can n			Detel										0101		A Waters Compa
elevant results Specify a s In the batch	s, you can r			Batch	n deleti	ion criteria										
Specify a s In the batch		emove them				1	Score	less than	•				Sequence Le	ength less than	•	
In the batch							Hits	less than	•				Cł	harge less than	•	
	h deletion cr		the property				Mass	less than	•				Sequ	ence contains	•	
delete.	a set of Iden	uncations yo	ou want to	Abso	lute m	ass error (less than					Acce	ssion contains	•	
Delete the	unwanted	identificat	ione	A030	face in	luss ciror (
To delete th			n't want, click				m/z	less than	•				Descri	ption doesn't co	ntain 🔻	
either:	Matching St	earch Result	s to delete			Retention	Time	less than	•				Modifica	tions contains	•	
the hig	hlighted IDs	5									Delete	matching	search results	Delete non-m	atching search re	sults Reset the criteria
		ng Search R are <i>not</i> high									Delete	e matering	sedicirresuits		atching searchine	suits Neset the chiena
					#	Score	Hits	m/z	RT(mins)	Charge M	Mass	Mass err	Sequence	Accession	Modifications	
Reset the To specify a				V	979	62.95	5	571.28	34.55	2 11	L40.5	-1.59	FNLAGDTT	🕥 gi 2964519		D-proline reductase [Clc
delete, click	k Reset the (then return to		994	68.46	3	535.83	35.11	2 10	069.64	-1.09	AVIVAVEELK	🎯 gi 2549738		60 kDa chaperonin [Clos
step 1 abov	/e.				1002	90.35	4	704.36	40.15	2 14	406.71	-0.66	FSTTYSSAI	🎯 gi 1266990		FMN-binding protein [C
👕 Tag filter	r applied	_			1002	90.35	4	704.36	40.15	2 14	106.71	-0.66	FSTTYSSAI	🕥 gi 2964514		5-nitroimidazole antibio
	ons may be l	hidden	Edit		1004	122.97	4	887.97	44.27	2 17	772 0'	1.10	ATVPSGAST	🕥 gi 2964520		enolase [Clostridium dif
					1004	122.97	4	887.97	De	lete selecte	ed pep	otide(s)	SGAST	🕥 gi 1267007		enolase (2-phosphoglyc
Total Hit	ts m/z	RT (min Cha	arge Tag 💌		1013	44.61	3	692.38	41.47	2 1	382.74	-0.47	S LGIELQPTD	gi 1267004		D-alaninepoly(phosph
		44.27 2			1014	48.31	5	613.33	38.68	3 18	336.96	-1.24	S KYVAPAIG	gi 2095712		cell wall protein V [[Clos
					1015	91.11	5	1428.69	53.20	2 28	355.37	-0.36	LDNLGDG	gi 2549763		cell surface protein (S-la
					1017	69.73	3	541 32	27.50	2 10	080.62	-0.25	LIDIANPTPK	gi 1266976		30S ribosomal protein S

Return to **Review proteins** locate the protein using the Search, as above, and clear the filter at the level of the peptides

There are now 7 remaining peptides for this protein



Now click **Return to the list of proteins** and create a tag for those proteins that have oxidised Methione residues.

Modified proteins can be located by specifically searching for proteins containing modified peptides. Right click on the list of proteins and select **Modification** from the list of **Quick Tags**.

Accession	Peptides	Unique peptides	Confidence score	Anova (p)	q Value	Tag 💌	Max fold change	Highest Mean	Lowest Mean
🎯 gi 54781345 (+1)	5	5	398	7.97E-06	1.03E-05		7.65	Α	С
🚳 gi 54781347	6	6	544	0.00402	0.000686	•	1.94	Α	С
🕥 gi 126697583	1	1	79.3	4.72E-05	2.74E-05		4.04	Α	С
S gil126697628	Anova p-va		153	0.981	0.0784		1.06	Α	С
🕥 gi 126697629 🔍	Max fold cl	nange≥2	319	0.128	0.0125		1.41	С	Α
🕥 gi 126697630	New tag		308	0.253	0.0236		1.15	С	Α
🎯 gi 12669763:	Quick Tags	•	Anova p-valu	ie	135	•	1.49	А	С
🎯 gi 126697636 🚰	Edit tags		Max fold cha	nge	231	-	1.32	А	С
🎯 gi 126697637	Add to Clip	Gallery	Modification		10291		7.28	Α	С
🕥 gi 126697639	3	3	Sequence		-28		1.15	С	А
🎯 gi 126697640	3	3	Peptide tags	contain	43	•	1.73	Α	С
🎯 gi 126697643	3	3	196	0.117	0.0116		1.4	С	Α
🎯 gi 126697650	1	1	80.9	0.528	0.0457		1.19	С	Α

To find those proteins containing peptides with Oxidation on Methionine residues enter **Oxidation (M)**. This will automatically provide a named tag when you click **Create tag**.

OP New Quick Tag			×
Where any peptide of	f a protein has		
Modification with:	Oxidation M		
Can I use wildcards?			
Apply the following ta	ag:	_	
Modification	with Oxidation M		
		Create tag	Cancel
L			

To reduce the table to displaying only these proteins with modified peptides (Oxidation on methionine) use the tag filter to focus on these proteins.

QIP Filter the proteins	X
Create a filter Show or hide proteins based on a selection of t the filter. For more guidance, please see the <u>or</u>	their tags. Move tags to the appropriate boxes to create nline reference.
Available tags:	Show proteins that have all of these tags:
 Anova p-value ≤ 0.05 (110 proteins) Max fold change ≥ 2 (83 proteins) 	Modification with Oxidation (M) (15 proteins) Show proteins that have at least one of these tags: Hide proteins that have any of these tags:
Clear the filter	OK Cancel

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

🥂 Progenesis QI.p Tutorial for DDA - Progenesis QI for prot	eomics										
File Review Experiment Import Data Alignment Filtering Design Setup	Review Peak Peptide Ion Picking Statistics	Iden Pepti		Refine s Identifications	Review Proteins	Prote Statist		leport		ņ	onlinea
		•				•				A	Waters Company
Review Proteins Using this screen, you can find the proteins of interest in your experiment.	Tag filter applie proteins may be		Edit	arch	Q						🕜 Help
Set the quantitation options	Accession	Peptides	Unique peptides	Confidence score	Anova (p)	q Value	Tag 🛛 💌	Max fold change	Highest Mean	Lowest Mean	Description
If you've not already done so, choose between relative and absolute guantitation, use of Hi-N,	🕥 gi 209571234	25	12	2.5E+03	7.76E-07	3E-06	۲	56.6	Α	с	cell wall pro
protein grouping and more.	🔮 gi 126699971	5	5	334	9E-06	1.04E-05	٨	11.7	А	С	thioredoxin
Protein options	🎯 gi 126697690	6	6	578	1.58E-05	1.5E-05	۲	5.94	A	с	ferredoxin/fl
	🕥 gi 384359782 (+1)	24	11	2.13E+03	2.03E-05	1.57E-05	۲	37	с	А	hemagglutir
2 Create a shortlist to review In the table, sort and filter the proteins based	🕥 gi 254973900 (+9)	15	15	1.95E+03	3.49E-05	2.38E-05	۲	3.78	Α	С	flagellin sub
on their measurements, to generate a shortlist	🕥 gi 254976387 (+5)	63	18	7.07E+03	6.5E-05	3.32E-05		2.44	с	А	cell surface (
for further review.	gi 255101963 (+2)	61	16	6.71E+03	9.9E-05	4.78E-05		2.62	A	С	cell surface (
We have the measurements calculated?	gi 126700129	3	3	267	0.000148	6.61E-05	۲	2.43	Α	с	translation in
To sort the table by a given value, simply click the relevant column header.	gi 126697654	3	3	215	0.00023	8.4E-05	۲	2.56	Α	с	30S ribosom
Review the proteins	gi 126700372	2	2	126	0.0037	0.000664	۲	5.37	Α	с	PTS system I
For each protein of interest, review its peptide	gi 54781347	6	6	544	0.00402	0.000686		1.94	Α	с	2-hydroxyist
measurements and correlations:	gi 126697631	7	7	626	0.00934	0.00135		1.49	Α	с	50S ribosom
View peptide measurements	gi 126700078	6	6	582	0.0107	0.00142	•	1.32	с	A	molecular cł
You can also double-click to review a protein.	<			m							Þ
A Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	Selected protei		redoxin 2 (1 A	「rx2) [Clostr	idium d	lifficile	630]		С		
Export to pathways tool Export protein measurements	16 - 15 - 14 -		Her								
Export peptide measurements Export peptide ion measurements	P 15								141		
Experiment design	4										1
Review your data from a different perspective:	Quantifiable proteins	displaye	d: 15							Section	Complete 🖯

Note: the **Sequence** Quick tag can be used to locate Proteins containing peptides with specific sequence motifs, (i.e potential phosphorylation sites).

Stage 13: Exporting Protein Data

Protein data can be exported in a csv file format. You can either export the **Protein, peptide or peptide ion 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 for only proteins that have Oxidised Methionine residues.

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

QIP Filter the proteins	
Create a filter Show or hide proteins based on a selection of the filter. For more guidance, please see the <u>o</u>	their tags. Move tags to the appropriate boxes to create nline reference.
Available tags:	Show proteins that have all of these tags:
Anova p-value ≤ 0.05 (110 proteins) Max fold change ≥ 2 (83 proteins)	Modification with Oxidation (M) (15 proteins)
	Show proteins that have at least one of these tags:
	Hide proteins that have any of these tags:
	OK Cancel

· ·	for further processing your data to external tools, there's ur analysis.
	Export to pathways tool
	Export protein measurements
	Export peptide measurements
	Export peptide ion measurements

QIP P	Progenesis QI.p Tutorial for DDA - Progenesis QI for prot	teomics
Г	File	
	Save	Review Peak Picking
	Close	
	Export peptide ion measurements	🜍 Ta
	Export peptide measurements	T pro
	Export protein measurements	Accession
	Export to pathways tool	🔇 gi 209
	Import additional protein data	🔇 gi 126
	Import protein accessions as tag	🎯 gi 126
	Export mzIdentML for PRIDE submission	🎯 gi 384
		🎯 gi 254
	Experiment properties	🎯 gi 254
	Show Clip Gallery	🎯 gi 255
×	Fxit	🎯 gi 126
Ļ	Exit	🛛 🎯 gi 126

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.

Export protein measurements	
Choose properties to be included in exported file	
Accession Peptide count Unique peptides	
Confidence score Anova (p)	
 ✓ q Value ✓ Max fold change 	
Power Highest mean condition	
Lowest mean condition Description Normalized abundance	
Raw abundance Amount (fnol) Iaos	
OK Cancel	J

OP Export	protein measurements
	Export complete
	Open File Open Folder Close

Note: if you have performed **Absolute Quantification** then the 'Amount' field will be available.Excel will open displaying the exported protein measurements

A	• : × <	f_{x}											٧
	А	В	С	D	E	F	G	н	I	J	к	L	
1							Normalize	ed abundar	nce				
2							Α			С			
з	Accession	Peptide count	Unique peptides	Anova (p)	Max fold change	Description	A1	A2	A3	C1	C2	C3	
4	gi 209571234	27	13	6.83E-07	62.79369729	cell wall protein V [[Clostridium] difficile	2789425	2970726	3305345	53699.16	44389.53	46280.83	3
5	gi 126698450;gi 255654924	13	13	2.26E-06	22.05186495	ABC transporter sugar-family extracellula	47609.09	57641.84	59778.92	1190086	1163209	1285921	1
6	gi 126699971	5	5	9.36E-06	11.66850585	thioredoxin 2 (Trx2) [Clostridium difficile	1531732	1617510	1554755	112771.2	140965.7	149399.3	3
7	gi 126697690	6	6	1.61E-05	5.926205751	ferredoxin/flavodoxin oxidoreductase su	541761.6	622072.8	489616.9	89471.36	95303.22	94232.15	5
8	gi 254973900;gi 260682017;	14	14	3.58E-05	3.769225308	flagellin subunit [Clostridium difficile QC	7134978	6250012	7703039	1841126	1800413	1953252	2
9	gi 384359782;gi 209570719	25	11	4.87E-05	34.10724963	hemagglutinin/adhesin [Clostridium diff	15313.93	21869.31	28048.02	671330.6	699702.7	853825.5	5
10	gi 254976387;gi 296452394;	63	17	8.42E-05	2.337336712	cell surface protein (S-layer precursor pro	36387455	38943949	36547316	87665176	79830732	94002331	L
11	gi 255101963;gi 21702505;g	62	16	9.08E-05	2.399770929	cell surface protein (S-layer precursor pro	76442562	81974214	73615536	33461634	29397327	33830398	3
12	gi 126700129	3	3	0.000152412	2.427619895	translation inhibitor endoribonuclease [0	1904730	1859457	2109501	842499.2	727363.1	849662.5	;
13	gi 126697654	3	3	0.000221494	2.552432634	30S ribosomal protein S8 [Clostridium dif	339592.7	383138.6	401331	152480.2	156522.2	131386.2	2
14	gi 54781347	7	7	0.004201931	1.929454479	2-hydroxyisocaproate-CoA transferase [C	536600.6	429139	422380.2	236515.8	207800.5	275120.1	L
15	gi 126697631	7	7	0.00888281	1.496908451	50S ribosomal protein L7/L12 [Clostridiun	835727.3	706354	938772.6	563937.3	550111.2	543269.9)
	gi 126700078	7	7	0.010201297	1.325272359	molecular chaperone DnaK [Clostridium (173402.4	159057.1	143730.8	220741	211288	199053	3
17	gi 126697969;gi 255654423	10	10	0.022073139	1.854961791	Beta-subunit of electron transfer flavopr	2301233	2178468	3367460	1471693	1160940	1597730)
18	gi 126697659	3	3	0.059638789	2.099815096	50S ribosomal protein L15 [Clostridium di	61674.14	71262.92	138850.9	46640.17	46024.65	36769.39)
19	gi 126700790;gi 296452046	9	9	0.109419122	1.863572589	enolase (2-phosphoglycerate dehydratas	94166.12	63813.4	225567.4	241840.1	232040.9	240886.5	;

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 QI 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.

Ingenuity Pathways Analysis (IPA): Explores experimental data by identifying relationships, mechanisms, functions, and pathways of relevance.

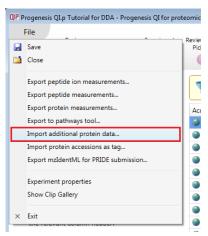
Note: Plugins for these tools are provided as standard.

Note: Access to IPA is dependent on having a local licence for IPA. Details of exporting data to IPA are described in Appendix 10 (page 92).

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 **IMPaLA** as the tool followed by the test to be performed.

File Review Experiment Import Data Alignment Filtering Design Setup	Review Peak Peptide Ion Identify Refine Review Protein Picking Statistics Peptides QC Metrics Identifications Proteins Statistics Report		nonlinea
			A Waters Compan
Review Proteins Using this screen, you can find the proteins of interest in your experiment.	Tag filter applied proteins may be hidden Edit.		🕡 Help
Set the quantitation options	Accession Peptides Unique peptides Confidence score Anova (p) Tag 🔻 Max fold change	Highest Mean Lowest Mea	n Description
If you've not already done so, choose between relative and absolute quantitation, use of Hi-N,		A C	decarboxylase
protein grouping and more.	QIP Export Pathways Information	C A	FMN-binding r
Protein options	Select a pathways tool	A C	cell wall proteir
	Choose a pathways tool from the list below. You can find out more or download new plugins using the links below.	C A	ABC transporte
Create a shortlist to review		A C	transketolase, «
 In the table, sort and <u>filter the proteins</u> based or their measurements, to generate a shortlist for 	Which pathways tool do you want to use?	A C	elongation fact
further review.	IPA: Ingenuity Pathway Analysis 🔹	C A	glyceraldehyde
How are the measurements calculated	IMPaLA: Integrated Molecular Pathway Level Analysis IPA: Ingenuity Pathway Analysis	A C	transcription el
To sort the table by a given value, simply click the relevant column header.	Panther Classification System	A C	propanediol ut
		A C	thioredoxin 2 (
Review the proteins For each protein of interest, review its peptide		A C	(R)-2-hydroxyis
measurements and correlations:		A C	PTS system ma
View peptide measurements		A C	ferredoxin/flav
You can also double-click to review a protein.		A C	ferredoxin-NAI
			*
4 Export data for further processing By exporting your data to external tools, there's no limit to your analysis. Export to pathways tool	s		
Export protein measurements	< Back Next > Cancel	с	
Export peptide measurements			
Export peptide ion measurements	R2 15 - NM 14	рфя	
Experiment design			
teview your data from a different perspective:			
Current design: AC 🔹	Quantifiable proteins displayed: 82	Section	on Complete 🖯

Select either Pathway over-representation analysis or Wilcoxon pathway enrichment analysis.



4	Export data for further processing ¹ By exporting your data to external tools, there's no limit to your analysis.							
	Export to pathways tool							
	Export protein measurements							
	Export peptide measurements							
	Export peptide ion measurements							

Make sure the **Open IMPaLA in my browser** is ticked and then click **Copy proteins to clipboard** When IMPaLA opens paste in the exported values and perform the test.

IMPaLA: Integrated Molecula	ar Pathway Level Analysis
pathway over-representation and enrichment ana	lysis with expression and / or metabolite data
genes/proteins - example input for over-representation analysis - example input for enrichment analysis	metabolites - example input for over-representation analysis - example input for enrichment analysis
paste genes or proteins below	paste metabolites below
g11254976387 0.12 g1384359782 0.28 g112604077 0.19 g11269971234 -1.8 g1126497970 -0.58 g1126497970 -0.31 g1126497970 0.27 g1126497970 0.27 g1126497970 0.27 g11264979788 0.08 g1126499718 -1 g1126499718 -1 g1126499718 -0.12 or upload a file with genes or proteins Browse	or upload a file with metabolites
optionally, provide genes/proteins background for over-representation analysis	optionally, provide metabolites background for over-representation analysis
Browse	Browse
Unigene	specify metabolite identifier 🗸
choose analysis type: pathway over-representati Wilcoxon pathway enrichn START ANA or clear the	nent analysis

Note: currently the Clostridium difficile protein set being used in this user guide does not yield any pathway 'hits' with **IMPaLA** or **Panther**.

Although the previous page outlines the process of exporting

(Impala_Wilcoxon_demo.txt) has been included in the tutorial download to allow the demonstration of exporting data

data to pathway tools, a text file

to Impala for pathways analysis.

QIP Export Pathways Information
Configure your export Choose which identifications to export and the type of analysis you want to perform.
Select the type of analysis to perform:
Pathway over-representation analysis
Pathway over-representation analysis Wilcoxon pathway enrichment analysis
To perform the pathway analysis, copy the protein data to the clipboard and paste it into the genes/proteins section of the IMPaLA search page.
< Back Copy proteins to clipboard Close

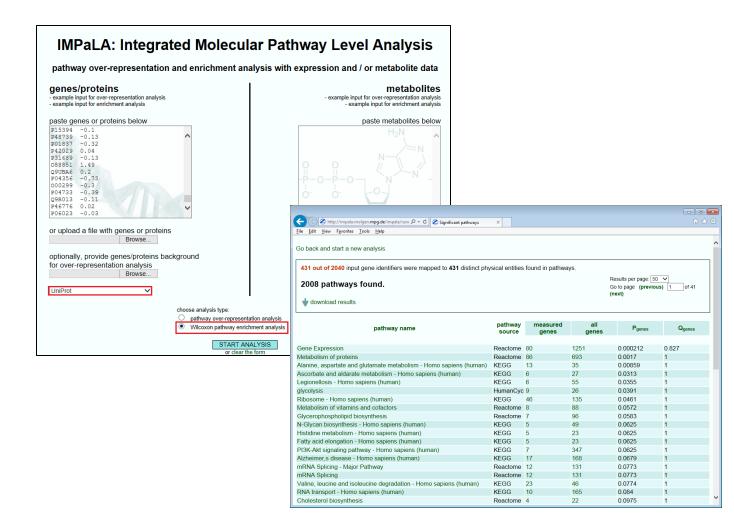
In Progenesis QI for proteomics set up the **Export to pathways tool** dialog to export the data to **IMPaLA** to perform a Wilcoxon pathway enrichment analysis. Click copy to clipboard to open **IMPaLA**.

To **simulate** the **Copy proteins to clipboard** open the text file in 'Notepad', select all and copy. In the **IMPaLA** window, open in your browser, right click on the genes/protein panel and paste the contents of the file Impala_Wilcoxon_demo.txt.

Select UniProt as the identifier, tick Wilcoxon pathway enrichment analysis then Start Analysis.

Note: if you are using your own data then pressing **Copy proteins to clipboard** will open **Impala** and allow you to paste directly into Impala, without saving to a separate file.

QIP Export Pathways Information
Configure your export Choose which identifications to export and the type of analysis you want to perform.
Select the type of analysis to perform:
Wilcoxon pathway enrichment analysis
For Wilcoxon enrichment analysis, choose two experimental conditions that you would like to compare.
Baseline: Comparison:
🗖 A 🔹 式 🗖 C 🔹
If greater, gives a negative fold change. If greater, gives a positive fold change.
To perform the pathway analysis, copy the protein data to the clipboard and paste it into the genes/proteins section of the IMPaLA search page.
Open IMPaLA in my browser
< Back Copy proteins to clipboard Close



Note: in the case of exporting to **Panther** the normal process requires a file to be created and stored then opened in Panther to perform the pathway analysis.

Clear any applied tag filters then move to the next section/stage.

Stage 14: Exporting identifications for submission to PRIDE

Following the analysis and review of your data you can export identification results as mzldentML.

Note: this only supports identifications obtained from searches performed in Progenesis QI for proteomics version 4.0 or later, using the search methods **Mascot** or **Ion Accounting**.

Export Identifications is carried out at the Review Proteins stage of the workflow. Click on the **File** menu and select **Export mzldentML for PRIDE submission...**

A dialog will appear allowing you to include the protein sequence or not in your export. Choosing to export the sequence will allow the display of the coverage map in PRIDE.

Enter a file name. Progenesis will save two files the mzldentML as a .mzid file and also a corresponding .mgf file containing the spectra.

Open PRIDE Inspector and locate and select the mzid file to import.

	IDE Inspector 2.5.2		
	Quick Start	Try Examples	
Y	Open Ident	Select mzML/mzXML/mzid/PRIDE xml Files Look In: PRIDE submission DDA Tutorial.MSMS.mgf DDA Tutorial.mzid ML	
	Review Pro	File Name: DDA Tutorial.mzid	amples
	Feedback	Open Cancel	
	🖂 Give Us Your Feed	adback	
	When use PRIDE Inspector, p	please cite:	

Click open.

You will then be asked if you want to load the corresponding spectrum files. Click yes and locate the .mgf file.

<mark>qip</mark> p	QIP Progenesis QI.p Tutorial for DDA - Progenesis QI for prote									
	File	F								
	Save	ľ								
1	Close									
	Export peptide ion measurements									
	Export peptide measurements									
	Export protein measurements									
	Export to pathways tool									
	Import additional protein data									
	Import protein accessions as tag									
	Export mzIdentML for PRIDE submission									
	Experiment properties									
	Show Clip Gallery									
×	Exit									

Export mzIdentML for PRIDE submission
Choose properties to be included in exported file
V Protein sequence
OK Cancel

A dialog opens and allows you to **Add spectra files.** Add the corresponding .mgf file and click **Set**

🗤 mzldentMl Files				🛁 He 🚱
🧪 DDA Tutorial.mzid	File Name	Size (M)	Туре	Remove
	DDA Tutorial.MSMS.mgf	242.9612	MGF	×
	Add spectra files			
			Cancel	Set

PRIDE Inspector will open allowing you to check the data.

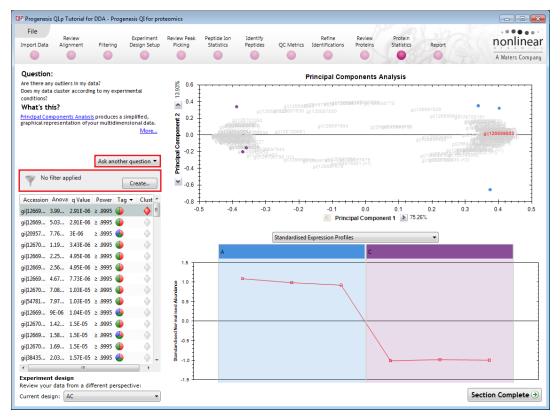
en Exp	port <u>H</u> elp									
		Overview Protein Peptide	Spectrum Summ	nary Charts						
I.j As		Overview Flotein Feptide	Spectrum Sum							
🧪 C	DDA Tutorial.mzid 🛛 🛄 🗙	Protein List					🛓 Up	date Protein Details	Y Decoy Filter	1
		Protein Group ID	Protein	#	#PSMs		#Distinct Peptides	#PTMs		
		+ ProteinGroup_1_gi_x007C_:	2 <u>qi 254976387</u>	4	476		58	7		
		+ ProteinGroup_2_gi_x007C_	2 <u>gi 255101963</u>	4	453		56	6		
		ProteinGroup_3_gi_x007C_	2 <u>gi 209571234</u>	1	133		23	2		
		+ ProteinGroup_4_gi_x007C_;	2 <u>qi 260682215</u>	1	112		22	2		
		ProteinGroup_5_gi_x007C_	1 <u>qi 126700407</u>	8	88		17	0		
		+ ProteinGroup_6_gi_x007C_3	2 <u>qi 254973900</u>	1	100		12	5		
		+ ProteinGroup_7_gi_x007C_	1 <u>qi 126698450</u>	7	70		11	1		
		ProteinGroup_9_gi_x007C_			38		9	1		
		+ ProteinGroup_10_gi_x007C			33		9	0		
		+ ProteinGroup_8_gi_x007C_	1 <u>qi 126697969</u>		50		9	2		
		ProteinGroup_11_gi_x007C	_qi 255101959		31		8	0		
		ProteinGroup_12_gi_x007C	_qi 126699128		49		7	0		
		ProteinGroup_13_gi_x007C	_qi 126698643		23		7	0		
		ProteinGroup_14_gi_x007C	_qi 126700790	3	31		7	2		
						_				
F 40	say Summary	PSM List [gi 126697970] Mod	lified residues: [C - 5	7.0215]					炎 Show Proteir	n Gro
M AS	ssay Summary	Peptide	Ranking	Delta m/z	Charge	Precursor m/	z Modifications	Length Start	t Stop	
ົ ຈ	Spectra found	IAPVVIELLGEGR	1	-0.0002	3	455.9413		13 24	36	
	Proteins found	IAPVVIELLGEGR	1	-0.0002	3	455.9413		13 24	36	
	roteins iounu	IAPVVIELLGEGR	1	-0.0002	3	455.9413		13 24	36	
🝚 F	Protein Group found	IAPVVIELLGEGR	1	0.0003	2	683.409		13 24	36	
	Peptides found	IAPVVIELLGEGR	1	0.0003	2	683.409		13 24	36	
	r epildes lound		1	0 0003	2	602 400		12 24	26	
• N	Mod: UNIMOD:4	Spectrum Fragmentation Ta	able Sequence							
	Mod: UNIMOD:35									
		🛕 Protein sequence is retriev	ed from the original	file				4	🛓 Update Protein (Detai
• A	Auto MS/MS Annotations									
							Selected P	PTM 📕 Fit 📕 F	uzzy Fit 📕 Ove	erlap
		Accession: 126697				· · · · · · · · · · · · · · · · · ·				
		38 peptides (38 ma	tched, 9 distinct), 98	/345 amino ac	ids (28.4%	6 coverage)				
		MNDIKDLSSY	KNVWIFAEOR	EGK <mark>IAP</mark>	VVTE	llgegr klak	EVDAELCAIL	LGK DVDGLAK	K 60	
		ELITFGADKV	YVADDALLEK			KDAIDE IKPE	IMLFGATHIG			
		VGTGLTADCT	KLEIDPEDKK			NIMATIICPN	HRPOMSTVRP			
		TRTGEVIALD	YKITODDIRT	TVLETV		DLVSLTDANV	IVSGGLGLGG			
		ADKLGGVVGS	SRAAVDAGWI			TTVKPNLYIA	CGISGAIQHL			
		AINKNPAAPI	LEIADYGVVG	_	-	EKLDSVDDLL	EAIKA	THORE DO DE LA	345	
1										

Once checked you can submit the data to PRIDE using the PX Submission Tool.

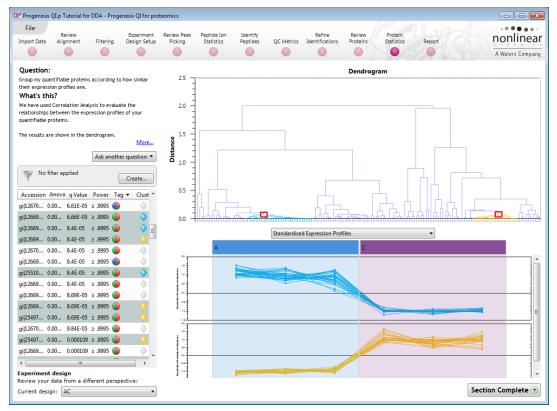
Note: For exporting mzIdentML of results to PRIDE where the Search Method is Mascot, you **must select** additional **Mascot Export Fields** when exporting your search results from Mascot to Progenesis: (Protein sequence (Protein Hit Information) and Start and End (Peptide Match Information)), **Appendix 11 (page 95)**

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 (in this example none). Allowing you to identify similar paterns of expression using the Correlation Analysis.



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.**

OP Filter the proteins	×
Create a filter Show or hide proteins based on a selection of t the filter. For more guidance, please see the or	heir tags. Move tags to the appropriate boxes to create line reference.
Available tags:	Show proteins that have all of these tags:
 Anova p-value ≤ 0.05 (110 proteins) Max fold change ≥ 2 (83 proteins) 	Modification with Oxidation (M) (15 proteins) Show proteins that have at least one of these tags: Hide proteins that have any of these tags:
Clear the filter	OK Cancel

As an example we will create a report for **only** the proteins containing peptides with Oxidation of Methionine residues.

- 1. First reduce the proteins to report on by selecting the 'Modification with Oxidation (M)' tag. In this example it reduces the number of proteins in the table to 15.
- 2. Expand the various Report Design options (by default they are all selected)
- 3. Un-tick as shown below
- 4. Click Create Report

QIP Progenesis Q	I.p Tutorial for D	DA - Proge	nesis QI for pro	oteomics								_	
File	Review		Experiment	Review Peak	Peptide Ion	Identify		Refine	Review	Protein			nlinear
Import Data	Alignment	Filtering	Design Setup	Picking	Statistics	Peptides	QC Metrics	Identifications	Proteins	Statistics	Report	ΰŲ	ninnear
		•					•	•	9/			A Wa	ters Company
Protein report	Peptide report												
Report on y	our proteins	5				Structur	e the repo	rt					
Generate a rep	ort containing t	the proteins	s of interest ir	n your experir	nent.	Enter a re	port title:						
4 Choose v	vhat to report	upon				Tutorial N	lethiodine oxid	lised proteins					
Using the	list below, filter		to show only	the proteins	you want to	Select the	sections to in	clude in the rep	ort:				
include in	the report.		② Learn	about tagging	and filtering	🔍 🔳 Ov	erview run						
O Customia	e the report					💎 🔳 Da	ta processing i	methods					
📕 Enter a tit	le for your repo		ct the section	s you want to	include in it	~							
using the	controls to the	right.					periment desig	n					
	er applied					\sim	otein report						
T protein:	s may be hidden	1			Edit		_	protein abundar	ices and pept	ides identifie	d for each prot	ein	
Accession	Anova (p)	Fold	Tag 💌	Description		_	rotein table eptide tables						
gi 209571234	7.76E-07	56.6	4	cell wall pro	tein V [[Clostr	_							
gi 126699971	9E-06	11.7	٩	thioredoxin	2 (Trx2) [Clost	\sim	otein Details				_		
gi 126697690	1.58E-05	5.9	٩	ferredoxin/f	lavodoxin oxic	· · _		of every protein	which match	es your curre	ent filter		
gi 260682215	(+1) 2.03E-05	37.0	٩	hemaggluti	nin/adhesin [C	V T	-						
gi 254973900	(+9) 3.49E-05	3.8	٩	flagellin sub	unit [Clostridi	V E	xpression profi	le					
gi 254976387	(+5) 6.5E-05	2.4	٩	cell surface	protein (S-lay	👻 🔲 Pe	ptide ion table						
gi 126700790	(+1) 8.44E-05	2.8	•	enolase (2-p	ohosphoglyce	マ 🔳 Pe	ptide ion detai	ls					
gi 255101963	(+2) 9.9E-05	2.6	٨	cell surface	protein (S-lay		_						
gi 126700129	0.000148	2.4	٩	translation i	nhibitor endo	Create re	port						
gi 126697654	0.00023	2.6	٩	30S riboson	nal protein S8								
gi 126700372	0.0037	5.4	۹	PTS system	HPr protein [(
gi 54781347	0.00402	1.9	4	2-hydroxyis	ocaproate-Co								
gi 126697631	0.00934	1.5	4	50S riboson	nal protein L7/								
gi 126700078	0.0107	1.3	•		haperone Dna								
gi 126697969	(+1) 0.0222	1.9	۷	Beta-subun	it of electron t								
•					Þ								
Experiment de	sign: AC		•										
	-												

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

Tutorial Methiodine oxidised proteins																			
Experiment:	Progen	esi	s QI	.p Tutorial	for D	DA													
Report created: 29/06/2017 16:42:47																			
Proteins																			
Proteins																			
Protein building options																			
Protein grouping Group similar proteins																			
Protein quant	-					g Hi-3													
														1					
Accession	Peptides	So	ore	Anova (p) *	Fold	Tags	Descr	ription						Average Abunda	Normalised				
														٨	c				-
gi 254976387	63 (18)	70	7	gi 254973	900_	_				_			_						
gi 255101963	61 (16)	67	_	51 234773	700														
			fla	agellin subun	it [Clo	stridiu	um diff	icile Q	CD-66c2	26]									
<u>gi 209571234</u>	25 (12)	-	Ŭ	o peptides															
<u>gi 260682215</u>	24 (11)	21	³ s	equence			P	Peptide on	Score	Hits	Mass	Charge	Tags	Conflicts	Modifications	ln guantitatio	Average Abundan	Normalised	
gi 254973900	15 (15)	19	4					011								quantitatio		C	
gi 126697969	10 (10)	9	2 4	ADDAAGLAISEK				148	100.52	5	1230.6087	2		0		yes	3.11e+00	5 9.63e+005	
<u>gij 120077707</u>	10 (10)	2	2	DTDVASEMVNLSK	<			1839	98.96	4	1407.6560	2		0		no	1.28e+00	5 3.05e+004	
<u>gi 126700790</u>	7 (7)	6	3	DTDVASE <mark>M</mark> VNLSH	<			3330	84.71	4	1423.6495	2		0	[8] Oxidation (M)	no	7.34e+004	1.37e+004	
				ADELLQLK				277	((10	r	1011 1011	2	-	-	()		2.0.1 00	0.45 004	
gi 126697631	7 (7)	6		ADELLQLKDEVER			Acces	ssion s	gi 254	973	900 (+9)								
gi 126700078	6 (6)	5	8	ADELLQLKDEVE															
				RDTDVASEMVNL					-	sub	unit [Clost	ridium	diffic	ile QCD-	66c26]				
				RDTDVASEMVNL				tides 1											
				RDTDVASE <mark>M</mark> VNL	SK				945.78	_									
				RDTDVASE <mark>M</mark> VNL	SK				.49e-00	5									
				SSSIEFNGK				Fold 3		vali	ie ≤ 0.05								
				LESTQNNLNNTIE	NVTAAF	SR		-			nge ≥ 2								
				ESTQNNLNNTIE				-			with Oxid	ation (M)						
				- NILVQASQS <mark>M</mark> LA				- "	Journea	cion	With OXIC	acion (i	.,						
											А					с			
							175	ET T			A					C			
					17.5 T 17.0 T 17.0 T 16.5 T 15.5 T 15.5 T 14.5 T														
						1	17.0												
						AL	16.5	1			E								
]	9 16.0	-											
							E 15.5												
							15.0	3								101			
						442	LD.0												
						V													
							14.0	-										•	

Note: if you scroll down on the second page of the report you can locate expression profiles for each 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.

Creating an Inclusion list

Inclusion lists can be used to try to increase the number of identified peptides you have in your experiment. They are used to control your mass spectrometer to try and concentrate the collection of MS/MS scans from specified m/z – retention time positions. Runs made using the inclusion list can then be imported into Progenesis QI for proteomics and the extra MS/MS scans added to the experiment.

As an example we are going to create an inclusion list for all the peptide ions that show a Significant difference between groups A and C (Anova p<0.05) and have **no** MS/MS spectra.

Import Data	Review Alignment	Filtering	Experiment Design Setup	Peptide Ion Statistics	Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report

First return to **Review Peak Picking** using the Workflow icons.

With no filters applied right click on a peptide ion in the
table, select Quick tags and click on No MS/MS data.

Y	No filter ap	plied			
#	Anova (p)	Fold Tag 💌 Notes	Highest Mea	n Lowest Mear	
3601	3.38E-13	Most abundant			
3477	6.11E-13	Anova p-value ≤ 0.05	с	A	
141	6.73E-12	Max fold change ≥ 2	с	A	
1914	8.17E-12	Significantly up in C	с	A	
11993	1.45E-11	Up regulated in C	А	С	
3789	1.74E-11	Up regulated in A	С	А	
7708	1.8E-11	Poorly Correlated	А	С	
2219	2.48E-11	New tag	с	А	
4407	2.59E-11	Quick Tags 🔹 🕨	Anova p-value		
7473	2.87E-11	P Edit tags	Max fold chang	je	
4441	3.03E-11	Infinity 📒 🔜	Modification		
8109	3.19E-11	Infinity 😑 🔜	No MS/MS dat	a	
284	3.82E-11	Infinity	No protein ID		

QP Filter the peptide ions	×						
Create a filter Show or hide peptide ions based on a selection of their tags. Move tags to the appropriate boxes to create the filter. For more guidance, please see the <u>online reference</u> .							
Available tags:	Show peptide ions that have all of these tags:						
Most abundant (2330 peptide ions) Anova p-value ≤ 0.05 (7943 peptide ions) Max fold change ≥ 2 (9888 peptide ions) Significantly up in C (3937 peptide ions) Up regulated in C (3937 peptide ions) Up regulated in A (4006 peptide ions) Poorly Correlated (3 peptide ions)	No MS/MS data (8428 peptide ions) Show peptide ions that have at least one of these tags: Hide peptide ions that have any of these tags:						
Clear the filter	OK Cancel						

Filter the table so that it is only showing peptide ions with **No MS/MS data**. To do this click **Create** and drag the tag on to the Show peptide ions that have all these tags and click OK.

The table will now only be displaying peptide ions that have no ms/ms.

	v Peak Pi Tag filter ap	-													ſ		
Т	peptide ions	; may be h	idden												l	Edit	
#	Anova (p)	q Value	Fold	Tag 💌	Notes	Highest Mean	Lowest Mean	m/z	z	Mass	RT (mins)	RT window (mins)	Abundance	Intensity	Max CV (%)	MS/MS	F
141	7.34E-12	4.9E-09	Infinity			С	А	901.2218	2	1800.429	39.344	1.13	4.03E+06	1.43E+07	2.92	0	4
3063	1.05E-11	4.9E-09	Infinity			С	А	533.9859	3	1598.936	30.105	0.347	1.75E+04	3.32E+05	2.07	0	
7723	1.82E-11	5.62E-09	Infinity			А	С	847.7614	3	2540.262	55.142	0.202	8.85E+03	7.15E+04	2.22	0	
3790	2.03E-11	5.62E-09	2.03E+07			С	А	1441.1479	5	7200.703	39.344	0.484	2.49E+05	2.96E+05	173	0	
9422	2.54E-11	5.62E-09	Infinity			С	А	811.4650	2	1620.915	31.979	0.293	6.93E+03	7.57E+04	2.37	0	
7487	2.98E-11	5.63E-09	Infinity		1	А	С	1207.0705	2	2412.127	34.551	0.566	3.91E+04	8.4E+04	2.92	0	
8125	3.14E-11	5.63E-09	Infinity		1	С	А	1002.0215	2	2002.028	32.829	0.21	8.56E+03	9.91E+04	2.56	0	
11692	4.03E-11	5.79E-09	Infinity			С	Α	719.4194	2	1436.824	23.204	0.246	2.92E+03	4.46E+04	2.42	0	
12321	4.07E-11	5.79E-09	Infinity			С	Α	859.1351	4	3432.511	33.684	0.337	4.62E+03	5.65E+04	2.57	0	
11001	4.23E-11	5.79E-09	Infinity		2	С	Α	1011.0473	2	2020.080	41.307	0.22	5.4E+03	5.16E+04	2.64	0	
10047	4.72E-11	5.84E-09	Infinity		2	С	Α	1157.2779	3	3468.812	53.605	0.368	2.01E+04	9.39E+04	3.08	0	
7970	8.29E-11	8.99E-09	Infinity		2	А	С	822.3761	2	1642.738	39.598	0.468	1.43E+04	7.11E+04	3.39	0	-
۰																	P

Now select Export Inclusion List... from the file menu

Then select the appropriate format.

QIP Export	inclusion list
Select yo	our machine type and export the inclusion list:
Format:	Thermo Finnigan inclusion list About this data format Download others
Save to:	C:\Users\andy.borthwick\Documents\Customer Data\Progenes
	Export Cancel

Finally export the file to an appropriate location

Note: with certain MS machines it is possible to widen the retention time windows being used, this can be controlled using the following dialog.

QIP P	rogenesis QI.p Tutorial for DDA - Progenesis QI for pr	oteomi
	File	
F	Save	Revi Pi
	Close	
	Export peptide ion data	
	Export deconvoluted peptide ion data	
	Import peptide ion numbers as tag	
	Export all identifications	
	Export to Proteolabels	
	Export inclusion list	
	Experiment properties Show Clip Gallery	
× 10	Exit 950 3.87E-11 Infinity 📲 📃	
Inc	clusion list retention time windows	
D	o you want to widen the retention time windo	ws
b	y 0.00 🚔 minutes?	
	Yes No	

If you require further information on the inclusion list file formats then click the link **About this data format** in the Export Inclusion List dialog.

Note: The new LC-MS runs can then be added to the original experiment to increase the MS/MS coverage using the Add files facility at the Data Import Stage.

Congratulations!

This document has taken you through a complete analysis using Progenesis QI for proteomics, from Alignment through the analysis workflow to generate lists of proteins exhibiting expression changes across biological conditions using powerful Multivariate Statistical analysis of the data.

Hopefully our example has shown you how this unique technology can deliver significant benefits with

- Speed
- Objectivity
- Statistical Power

If you would like to see the benefits of running Progenesis QI for proteomics using your own data and explore the Progenesis QI for proteomics workflow please go to Appendix 4: Licensing Runs (page 78).

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, AB SCIEX, Agilent, Bruker and Thermo) 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.

QP Create New LC-MS Experiment	
Create a new label-free experiment named:	
Progenesis QI.p_Tutorial for DDA	
Data type	
Profile data	
Centroided data	
Resolution (full width at half maximum) 50000	
Machine true	
Machine type	
High resolution mass spectrometer	High resolution mass spectrometer
	e.g. Waters SYNAPT G2/G2-S, AB SCIEX TripleTOF, Agilent QTOF, Bruker Maxis, Thermo LTQ Orbitrap
Experiment folder	Thermo LTQ Iontrap in Enhanced mode.
Save experiment in the same folder as the run data	
Choose an experiment folder	Low resolution ion trap
	e.g. Bruker HCT, Bruker HCT Ultra, Thermo LTQ XL
	Thermo FT-ICR
Create experiment	Cancel

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 mzXML files

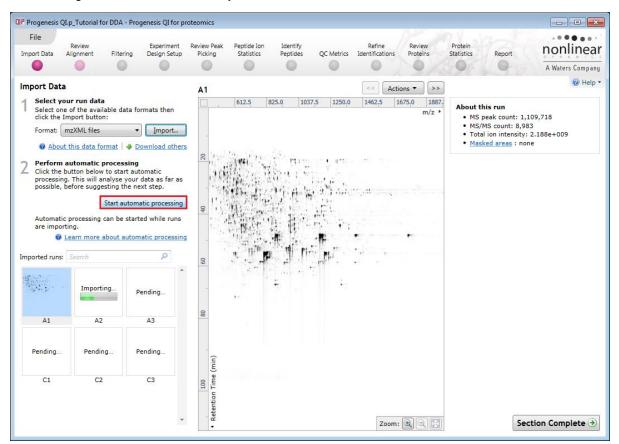
Then locate your data files using Import...

QIP Progenesis	QI.p_Tutorial f	or DDA - Prog	genesis QI for p	roteomics								-	
File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Ion Statistics	Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report	D Y	nlinear aters Company
Import Dat	ta												🕜 Help 🔻
Select o	your run data ne of the avai Import butto	lable data fo	rmats then	_									
Format:	mzXML files	•	Import										
@ Abo	AB SCIEV ()	5942.56324	vnload others	QIP Select	files								×
2 Perform	Version: 1.0 Agilent (.d)	.5877.31834	matic	00) 🗸 🚺 🔍 Pro	genesis 🕨	Progenesis QI	p Tutorial for DI	0A_v3.0 ►	- ↓	Search Progen	esis Ql.p Tu	toria 🔎
processi possible	Version: 2.0 Bruker Dalto		ata as far as : step.	Organ	ize 🔻 Nev	v folder					8	= - E	
	Version: 1.0		c processing		vorites			^ Na	me	*		Date mo	dified
	mzML Files Version: 1.0	.5955.26268	I while runs		vorites Desktop				Progenesis (QI.p Tutorial for	DDA.Analysis	26/04/20	016 10:40
are imp	mzXML files Version: 1.0		tic processing	. 🕠	Downloads				A1.mzxml	· · · · · · · · · · · · · · · · · · ·	,		008 08:23
· ·	NetCDF files Version: 1.0			÷	Dropbox				A2.mzxml			20/05/20	008 08:27
Imported run	Thermo (.ray		9		Recent Places				A3.mzxml				008 08:29
	Version: 1.0 Thermo FT-I		A	~	ShareFile				C1.mzxml				008 11:50
	Version: 1.0	CR (.raw) .5942.37174		•	BitTorrent Syn	5			C2.mzxml				008 11:50
					esktop Libraries Documents Music			÷ (C3.mzxml	III		20/05/20	008 08:25
No	runs have bee	n imported y	et			File <u>n</u> ame:	"C3.mzxml" "A1	l.mzxml⁼ "A2.mz	xml" "A3.mz	xml" "C1 ▼ [mzXML files (*.r <u>O</u> pen ▼	nzxml) Can	▼ cel
			Ŧ										

Locate and select all the Data files (A1 to C3).

On loading the selected runs your data set will be automatically examined and the size of each file will be reduced by a 'data reduction 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: you can 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.

Additional details on Selection of Alignment Reference

1. Assess all runs in the experiment for suitability

- This method compares every run in your experiment to every other run for similarity.
- The run with the greatest similarity to all other runs is chosen 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 can take a long time.

	ou want to choos ss all runs in the				
	the most suitable	1		c+	
0.000	the most suitable	run nom candi	uates tridt i sele		
	A1				
	AL	v			

2. Use the most suitable run from candidates that I select

• This method asks you to choose a selection of reference candidates, and the automatic algorithm chooses the best reference from these runs.

OP Start automatic processing			QP Start automatic processing	I		×
Select your alignment refe To mark a run for assessmer use the button below to mar	nt as an alignment reference	e candidate, double-click it or	Select your alignment re To mark a run for assessme use the button below to ma	ent as an alignment referen	ce candidate, double-click i	it or
Mark selected runs Clea	ar all	Q	Mark selected runs	ear all	٩	
		*				*
A1	A2	A3	A1	A2	A3	
S						E
C1	C2	C3 +	<u>с1</u>	E C2	С3	•
0 of 6 runs marked as candid	dates		3 of 6 runs marked as cand	lidates		
	< Back	Next > Cancel		< Back	Next > Canc	el

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.

This method is the most appropriate when you have some prior knowledge of your runs suitability as references:

- runs from pooled samples
- runs for one of your experimental conditions will contain the largest set of common peptides.

It is also the preferred option when analysing a large number of runs

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 it's chromatography before loading further runs).

Review Chromatography

Each data file appears as a 2D representation of the run. If you created a **profile** experiment, at this stage you will be warned if any of the data files have been 'centroided' during the data acquisition and conversion process, as files must be of one format or the other.

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

A1 A2 A3 C1 C2 C3 Mask areas for peak picking... X Remove this run Del

Note: you can also multi-select runs to remove by holding down the Ctrl key.

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

QIP Progenesis QI.p_Tutorial for DDA - Progenesis QI for p	oteomics	
File Review Experiment Import Data Alignment Filtering Design Setup	Review Peak Peptide Ion Identify Refine Review Picking Statistics Peptides QC Metrics Identifications Proteins	Protein Statistics Report nonlinear
	0 0 0 0 0 0/0/	A Waters Company
Import Data	C1 << Actions - >>	Help *
 Select your run data Select one of the available data formats then click the Import button: Format: mzXML files mport About this data format Download others Perform automatic processing Click the button below to start automatic processing. This will analyse your data as far as possible, before suggesting the next step. Start automatic processing Automatic processing can be started while runs are importing. @ Learn more about automatic processing 	612.5 825.0 1037.5 1250.0 1462.5 1675.0 1887. R	About this run • MS peak count: 1,182,000 • MS/MS count: 9,542 • Total ion intensity: 3.643e+009 • <u>Masked areas</u> : none
Imported runs: Search P	60	
	8	
A1 A2 A3		
	Retention Time (min)	
	Zoom: Q C	Section Complete ④

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 the alignment automatically already.

Now move to the next stage in the workflow (page 14 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; the automatic processing dialog will prompt you to 'drop-off' at the **Review Alignment** stage on completion to investigate the problem.

QP Processing Complete								
Automatic processing complete (with warnings).								
 Importing runs: 	7 of 7 processed							
 Selecting reference: 	C1							
Aligning runs:	6 of 6 processed							
	A 1 run failed to align - continuing without it							
Peak picking:	15935 peaks found							
Creating design:	Created							
 Protein quantitation: 	Relative Quantitation using Hi-3							
	Close Identify Peptides	s 🔿						

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

QIP Processing Complete			×
Automatic process Time taken: 3 minutes 57	• • •	h warnings).	
A Importing runs:	7 of 7 processed 1 failed to import		
 Selecting reference: 	C1	•	
 Aligning runs: 	5 of 5 processed		
 Peak picking: 	15935 peaks found		
 Creating design: 	Created		
 Protein quantitation: 	Relative Quantitation	using Hi-3	
		<u>C</u> lose	Identify Peptides 🌖

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, where in importing, you specified selection of the alignment reference 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).

QIP Processing Complete		×
Automatic process	ing failed.	
Time taken: 55 seconds		
Importing runs:	7 of 7 processed	
	A 1 failed to import	
X Selecting reference:	All reference candidates failed to import	
Aligning runs:	Unable to start.	
Peak picking:	Unable to start.	
Creating design:	Unable to start.	
Protein quantitation:	Unable to start.	
	lose	Import Data 🏈

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

Appendix 3: Stage 1 Data QC review and addition of exclusion areas

During the process of Data QC you may identify areas of the raw data for a particular run that appear 'noisy' yet still have identifiable 'isotopic patterns'.

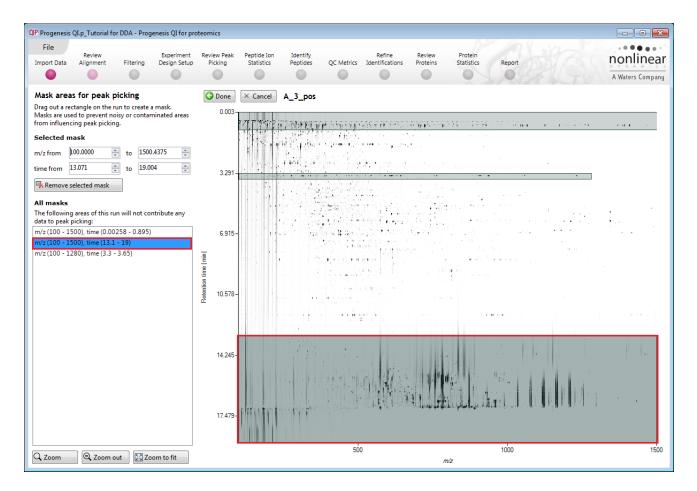
For example if the run is part of a 'replicate set' of runs it is possible to exclude such areas on the noisy run by applying a mask to the area. By doing so this area is excluded during the initial part of the detection process in order that it does not 'interfere' with the detection of the peptide ions in the replicate group.

To do this select **Mask areas for peak picking** from the **Actions** menu on the top left of the screen.

<<	Acti	ons 🔻 >>	
5 1250.0		Mask areas for peak picking	
	×	Remove run	,182,000 542
		<u>Masked areas</u> :	

Drag out an area over the noisy part of the run to create the mask.

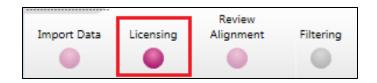
Note: Click Done to return to the Import Data view where you can zoom into the masked where you will see the isotopic peptide ions in the noise.



Note: if the level of noise is high and affecting many of your runs a preferred approach would be to reoptimise the chromatography to improve the levels of noise in your data.

Appendix 4: 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.

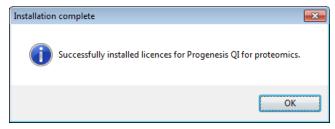
The runs in your experiment will be listed as shown below.

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

nport Data Licensing Alignment Filtering Design Se				Protein Statistics	Report	A Waters Comp
Dongle License Runs						
This installation is currently restricted to analyse licensed runs only.	Run name				Licence state	License this run
To license your runs, you need an evaluation or	C:\Users\andy.borthwick\Documents\Cus	tomer Data\Progenesis QI.p	v3.0 Tutorials and [Demo Suites\Pro	ge Unlicensed	
	C:\Users\andy.borthwick\Documents\Cus	tomer Data\Progenesis QI.p	v3.0 Tutorials and [Demo Suites\Pro	ge Unlicensed	
a sales representative.	C:\Users\andy.borthwick\Documents\Cus	tomer Data\Progenesis QI.p	v3.0 Tutorials and [Demo Suites\Pro	ge Unlicensed	V
Once licensed, your runs can be analysed on	C:\Users\andy.borthwick\Documents\Cus	tomer Data\Progenesis QI.p	v3.0 Tutorials and [Demo Suites\Pro	ge Unlicensed	
any installation of the software. The licence is	C:\Users\andy.borthwick\Documents\Cus	tomer Data\Progenesis QI.p	v3.0 Tutorials and [Demo Suites\Pro	ge Unlicensed	V
automatically included when archiving an experiment.	C:\Users\andy.borthwick\Documents\Cus	tomer Data\Progenesis QI.p	v3.0 Tutorials and [Demo Suites\Pro	oge Unlicensed	V
If your runs have been licensed on another computer, <u>click here</u> to make the licences available on this computer.						
If you have one, you can <u>open a licence file</u> to install.						
If you have just installed a dongle, click here.						
		Run licence code: xxx-xxx	0(-3000(-300)		Use Lice	ence 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 5: Manual assistance of Alignment

Approach to alignment

To place manual alignment vectors on a run (C2 in this example):

- 1. Click on Run C2 in the **Runs** panel, this will be highlighted in green and the reference run (C1) will be highlighted in magenta.
- 2. You will need approximately 5 alignment vectors evenly distributed from top to bottom of the whole run.
- 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

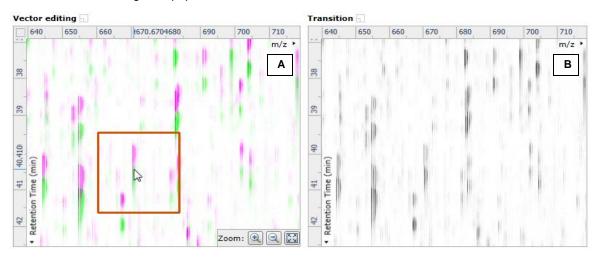


Note: the peptide ions moving back and forwards between the 2 runs in the **Transition** window (B) indicates the misalignment of the two runs.

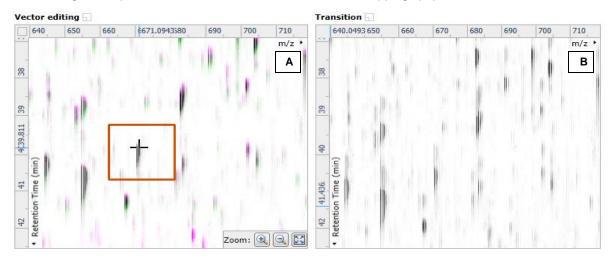
Note: the lon Intensity Map gives you a colour metric, visually scoring the current alignment. The overall score is placed next to the Vectors column in the table. With each additional vector added this score will update to reflect the 'changing' overall quality of the alignment. The colour coding on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

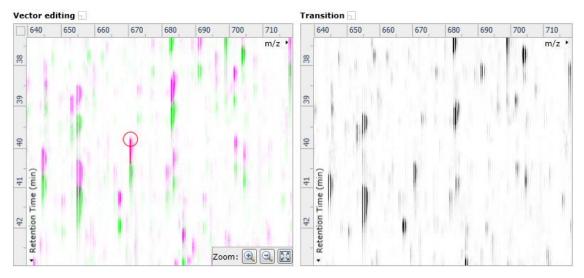
4. Click and hold on a green peptide ion in Window A as shown below



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.



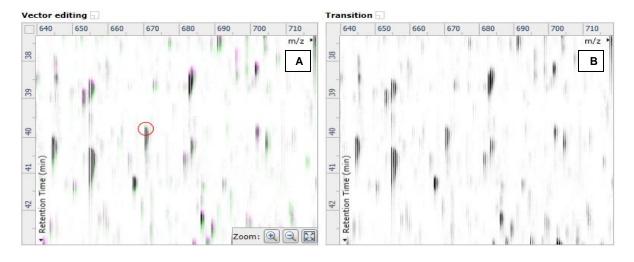
Tip: while holding down the mouse button hold down the **Alt** key. This will allow smooth movement of the cursor as the **Alt** key allows you to override the 'automatic alignment' performed as you depress the mouse button.



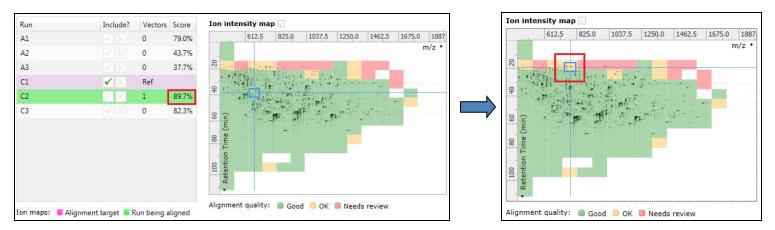
6. On releasing the left mouse button the view will 'bounce' back and a red vector, starting in the green peptide ion and finishing in the magenta peptide ion will appear.

Note: an incorrectly placed vector is removed by right clicking on it in the Vector Editing window

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.

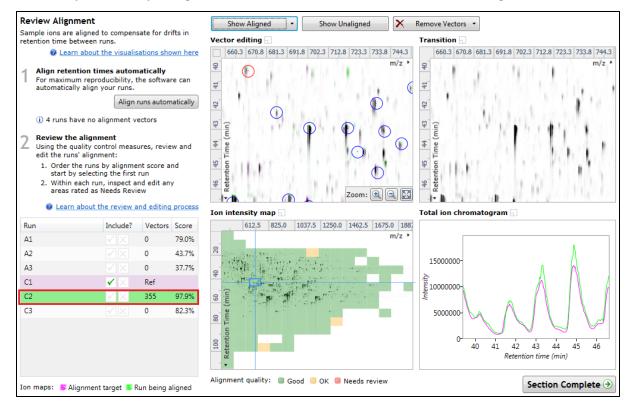


- 10. The shift in the Retention Time (RT) is as a result of incorrect running of the chromatography. In many of these cases if the Automatic Alignment fails to generate a good alignment then removing all the alignment vectors for this run and placing a **single** manual vector to act as a 'seed' for the Automatic Alignment algorithm maybe all that is required to generate a good alignment.
- 11. In the case of the example shown above placing 2 vectors from the top to the bottom of the run is sufficient to markedly improve the alignment (**Note**: the improvement in the score with the addition of only 2 vectors).
- 12. At this point you would redo the automatic alignment of this image by selecting **Align runs automatically**. **Note**: if you are focusing only on the alignment of one run, then un-tick the other runs in the alignment dialog.

Automatic Al	5	<u>-</u> ×
Run	ch runs to automatically align: Notes	Vectors
🔲 A1	This run has not been automatically aligned	0
A2	This run has not been automatically aligned	0
A3	This run has not been automatically aligned	0
🗹 C2	run has user vectors	2
🔲 C3	This run has not been automatically aligned	0
		OK Cancel

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.

To review the vectors automatic and manual, return to page 15.

Appendix 6: 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

QP Progenesis QI.p_Tutorial for DDA - Progenesis QI f	for proteomics	;										×
File Experim Import Data Alignment Filtering Design 9			Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Rep	ort	<u>•</u>	online Waters Com	I C S
AC AC-2 New											(i) I	Help 🔻
Which experiment design type do ye	ou want to	use for this ex	(periment	?								
Between-subject Design				0-0	Within-sub							
Do samples from a given subject appear in only one condition? Then use the between-subject design.	A	Δ1	Delete Remove	subjec	ou taken samp t under differ ise the within-	ent condition	s?		Before	During	After	
To set up this design, you simply group the runs according to the condition (factor level) of the samples. The ANOVA calculation assumes that the	c	QP Create New Expe Enter a name for t Before During and	the experimen	t design:		×	use p	atient X	X1	X2	X3	
conditions are independent and therefore gives a statistical test of whether the means of the conditions are all equal.		How do you want	to group the r				P	atient Y	¥1	YZ	Y3	
	Add conditio	Copy an exist	ing design:	Г	Create design	Cancel	t	atient Z	Z1	Z2	Z3	
				L	ereate design		ŀd					
				becaus assump repeat differe reduce conditi create The wi though paired- compai	dard ANOVA is i e the data viol ottion of indepe ed measures AI nces can be eli d as a source o on differences a more powerl thin-subject de t of as an exte samples t-test rison between ed measures.	ates the ANO' ndence. With NOVA individu iminated or if between (which helps ful test). esign can be nsion of the to include	/A a al to					

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 click on 'Select Sample' in each box of the matrix and select the appropriate sample.

IP Progenesis	QI.p_Tutorial fo	r DDA - Prog	jenesis QI for pro	oteomics						
File Import Data	Review Alignment	Filtering	Experiment Design Setup		ptide Ion Identi tatistics Peptid		Refine Identifications	Review Protein Proteins Statistics	Report	nonlinear
			•		0 0	•	•	0 0	50 75	A Waters Company
AC	AC-2	Befor	e During an	d After Treatm	ent I 🗙 📄	New				🔞 Help 🕶
Setup the co experiment of of your samp in the grid.	oles to the corr	bjects for y ght, and th ect subject/	en assign each		Before	During	C After	Add Condition		
2. Add a row		ct.	rrect location	Patient A	A1	A2	A3	_		
	C	88		Subject 1	C1	Select Sample C2 C3	Sel <u>ect Sample</u>	•		
	C			Add Subject				_		
									Sect	ion Complete Э

You can create additional Experimental Designs using the New tab

All of these Experimental Designs are available at the later stages of the workflow with the exception of **Identify Peptides**, and **Refine Identifications** (including Resolve Conflicts).

Appendix 7: Power Analysis (Peptide Ion 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 Ion Statistics 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 Ion Statistics section. A selection of 3 tools will appear.

v	Principal Components Analysis Are there any outliers in my data? Does my data cluster according to my experimental conditions?
ጨ	Correlation Analysis Group my peptide ions according to how similar their expression profiles are.
	Power Analysis How many replicates should I run? What is the power of my experiment?

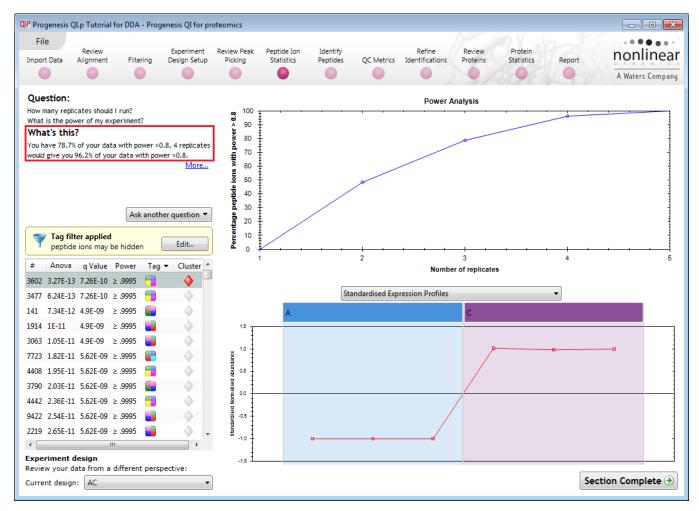
Select the option

'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 (7943), as an example, view the power analysis.



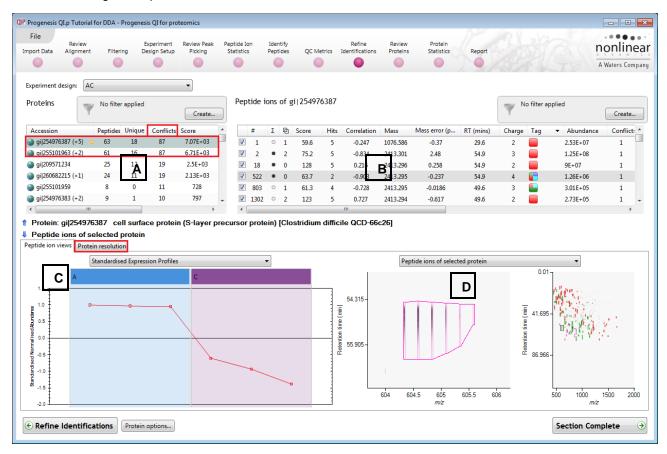
This is displayed graphically showing that 78.7% of the 7943 peptide ions have a power of 80% or that 4 replicates would give you 96.2% of your data with power > 0.8.

- 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 8: Resolve Conflicts

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

The **Resolve Conflicts** stage (now accessed at the bottom left of the Refine Identifications stage) provides a number of interrelated graphical and tabular views to assist you in the manual validation of the peptides that have been assigned to proteins and also to review the relevance of the data returned from the search.



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

QP Protein quantitation options
Quantitation method:
Relative Quantitation using Hi-N
Number of peptides to measure per protein (N): 3
Employ protein grouping, i.e. hide proteins whose peptides are a subset of another protein s.
OK Cancel

Note: manual conflict resolution is not essential to performing quantitation however it can lead to more quantifiable proteins.

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 go to page 49 and Protein Quantitation go to page 51

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.

Make 'cell surface protein (S-layer precursor etc' the current protein by clicking on it in Window A (a circular orange symbol indicates current protein). This protein has 63 peptides assigned (window B) which have a total of 87 conflicts. To view the conflicting assignments click on the **Protein Resolution** tab (window C previous page).

File Review port Data Alignment		Experiment Design Setup	Review Peak Picking		atistics		Ident Peptic		C Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report			Ca (r	onlin
•												•		St.	EL	181	A Waters Com
xperiment design: AC			•														
roteins 🕎 No	o filter applied		Create		Pept	ide i	ons c	of gi 25	4976387					₩ ^N	o filter a	pplied	Create
Accession Per	tides Unique	Conflicts	s Score	-		#	Σ	由 Scor	e Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	▼ Abundance	Conflict
🌒 gi 254976387 (+5) 🜼 6	3 18	87	7.07E+03	_	V	1	0	1 59	.6 5	-0.247	1076.586	-0.37	29.6	2		2.53E+07	1
gi 255101963 (+2) 6	1 16	87	6.71E+03		V	2	٠	2 75	.2 5	-0.834	2413 301	2.48	54.9	3		1.25E+08	1
	5 12	¹⁹	2.5E+03		V	18	٠	0 12	8 5	0.216	2413 296	B 0.258	54.9	2		9E+07	1
gi 260682215 (+1) 2		¹⁹ A				522		0 63	.7 2	-0.903	2413.295	-0.237	54.9	4		1.26E+06	1
	3 0	11	728		1	803	0	1 61	.3 4	-0.728	2413.295	-0.0186	49.6	3		3.01E+05	1
gi[255101959 i gi[254976383 (+2) ! Protein: gi]254976387 Protein: gi]255101963 sptide ion views Protein res) 1 cell surfa			precu	Irsor p	-	in) [(ium diffic	0.727 III QCD-66 Cile QCD-63		-0.617	49.6	2		2.73E+05	1
gi 254976383 (+2) 9 Protein: gi 254976387 Protein: gi 255101963	cell surfa	ace prote ace prote	in (S-layer	precu	Irsor p	prote prote	ein) ((ein) ((Clostrid	ium diffic	III QCD-666 cile QCD-636	:26]	-0.617	49.6	2		2.73E+05	
gij254976383 (+2) 9 Protein: gij254976387 Protein: gij255101963 ptide ion views Protein res Conflicting proteins for	cell surfa	ace prote ace prote on 522	in (S-layer	precu precu	rsor p rsor p Pep	prote prote	ein) ((ein) ((ions	Clostrid	ium diffic ium diffic 55101963	m cile QCD-666 cile QCD-636	:26]	-0.617 Mass error (p	49.6 RT (mins)	2 Charge	Tag	2.73E+05	•
gij254976383 (+2) 9 Protein: gij254976387 Protein: gij255101963 sptide ion views Protein res Conflicting proteins for Accession Pr	2 1 2 cell surfa 3 cell surfa olution r peptide i	ace prote ace prote on 522	in (S-layer in (S-layer	precu precu	v rsor p rsor p Pep	prote prote	ein) (ein) (ions Σ	Clostrid Clostrid of gi 2	ium diffic ium diffic 55101963 e Hits	m cile QCD-666 cile QCD-636	c26] q42]				Tag		•
gij254976383 (+2) 9	cell surfa cell surfa olution r peptide i eptides Uniqu	ace prote ace prote on 522 Je Conflic	in (S-layer in (S-layer it: Protein Sco	precu precu	virsor p irsor p Pep	prote prote tide i #	ein) [(ein) [(ions Σ	Clostrid Clostrid of gi 25	ium diffic ium diffic 55101963 e Hits .7 2	m cile QCD-666 cile QCD-636	c 26] q 42] Mass	Mass error (p	RT (mins)	Charge		 Abundance 	Conflic
gij254976383 (+2) 9	Cell surfa cell surfa olution r peptide i eptides Uniqu 63 18	ace prote ace prote on 522 Je Conflic 87 87	in (S-layer in (S-layer in (S-layer t: Protein Sco 7.07E+03	precu precu	Pep'	prote prote tide i # 522	ein) [(ein) [(ions (Σ ο	Clostrid Clostrid of gi 2 Cor 0 63	ium diffic ium diffic 55101963 e Hits .7 2 .3 4	Correlation 0.413	C26] q42] Mass 2413.295	Mass error (p -0.237	RT (mins) 54.9	Charge 4		 Abundance 1.26E+06 	Conflic
gij254976383 (+2) 9	cell surfa cell surfa olution r peptide i eptides Uniqu 63 18 61 16	ace prote ace prote on 522 Je Conflic 87 87	in (S-layer in (S-layer in (S-layer t: Protein Sco 7.07E+03	precu precu	Pep	prote prote tide i # 522 803	ein) [(ein) [(ions Σ	Clostrid Clostrid of gi 2: Clostrid of gi 2:	ium diffic ium diffic 55101963 e Hits .7 2 .3 4 23 5	ile QCD-666 cile QCD-636 cile QCD-636 correlation 0.413 0.380	Mass 2413.295	Mass error (p -0.237 -0.0186	RT (mins) 54.9 49.6	Charge 4 3		 Abundance 1.26E+06 3.01E+05 	Conflic 1
gij254976383 (+2) 9	cell surfa cell surfa olution r peptide i eptides Uniqu 63 18 61 16	ace prote ace prote on 522 ue Conflic 87 87 9eptides	in (S-layer in (S-layer tt: Protein Sco 7.07E+03 6.71E+03	precu precu	Pepi	prote prote tide i # 522 803 1302	in) [(in) [(ions Σ	Clostrid Clostrid Of gi 2: Clostrid Clostr	ium diffic ium diffic 55101963 e Hits 77 2 .3 4 23 5 22 5	Correlation 0.413 0.380 -0.220	Mass 2413.295 2413.294	Mass error (p -0.237 -0.0186 -0.617 -0.607 0206	RT (mins) 54.9 49.6 49.6	Charge 4 3 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 	Conflic 1 1
9 gi[254976383 (+2) 9 Protein: gi[254976387 Protein: gi[255101963 ptide ion views Protein res Conflicting proteins for Accession Pri 9 gi[254976387 (+5) 9	cell surfa cell surfa olution r peptide i eptides Uniqu 63 18 61 16	ace prote ace prote on 522 Je Conflic 87 87	in (S-layer in (S-layer tt: Protein Sco 7.07E+03 6.71E+03	precu precu	Pep'	prote prote # 522 803 1302 2097 3 4	in) [(in) [(ions Ω Ω	Clostrid Clostrid Clostrid © 63 1 61 2 12 0 12 1 51 6 89	ium diffic ium diffic 55101963 e Hits .7 2 .3 4 23 5 22 5 .7 5 .8 5	Correlation O.413 O.220 O.782 O.530 O.630	2413.295 2413.295 2413.294 2413.294 2413.294 2413.294 1194.62 1246.56	Mass error (p -0.237 -0.0186 -0.617 -0.607	RT (mins) 54.9 49.6 49.6 52.1 30.5 27.6	Charge 4 3 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 	Conflic 1 1 1 1
gij254976383 (+2) 9	cell surfa cell surfa cell surfa olution r peptide i eptides Uniqu 63 18 61 16	ace prote ace prote on 522 ue Conflic 87 87 9eptides	in (S-layer in (S-layer tt: Protein Sco 7.07E+03 6.71E+03	precu precu	V rrsor f rrsor f V V V V V	prote prote tide i # 522 803 1302 2097 3 4 5	in) [(in) [(ions (ο ο ο ο	Clostrid Clostrid Clostrid © 63 1 61 2 12 0 12 1 51 6 89 1 76	ium diffic ium diffic 55101963 e Hits 7 2 .3 4 23 5 22 5 .7 5 .8 5 .1 5	Correlation 0.413 0.380 -0.220 -0.782 0.530 0.630 0.607	Mass 2413.295 2413.295 2413.295 2413.294 2413.294 1194.62 1246.56 1206.661	Mass error (p -0.237 -0.0186 -0.617 -0.607 -0.206 F 0.254 0.0153	RT (mins) 54.9 49.6 52.1 30.5 27.6 26.3	Charge 4 3 2 2 2 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 2.31E+07 	• • • • • • • • • • • • • • • • • • •
gij254976383 (+2) 9	cell surfa cell surfa cell surfa olution r peptide i eptides Uniqu 63 18 61 16	ace prote ace prote on 522 ue Conflic 87 87 9peptides	in (S-layer in (S-layer tt: Protein Sco 7.07E+03 6.71E+03	precu precu	V Irsor f Irsor f V V V V V V	prote prote # 522 803 1302 2097 3 4 5 6	in) [(ions Σ Ο Ο Ο Ο	Clostrid Clostrid Clostrid	ium diffic 55101963 e Hits 7 2 3 4 13 5 12 5 .7 5 .8 5 .1 5 .9 5	Image: constraint of the second sec	Mass 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 1194.62 1246.56 1206.661 1250.615	Mass error (p -0.237 -0.0186 -0.617 -0.607 F 0206 F 0206 0.554 0.00153 0.208	RT (mins) 54.9 49.6 52.1 30.5 27.6 26.3 38.9	Charge 4 3 2 2 2 2 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 2.31E+07 1.51E+07 	, Conflic 1 1 1 1 1 1 1 1 1 1
gij254976383 (+2) 9	cell surfa cell surfa cell surfa olution r peptide i eptides Uniqu 63 18 61 16	ace prote ace prote on 522 ue Conflic 87 87 9peptides	in (S-layer in (S-layer tt: Protein Sco 7.07E+03 6.71E+03	precu precu	V Issor f Issor f V V V V V V V	prote prote # 522 803 1302 2097 3 4 5 6 8	in) [(ions (Σ Ο Ο Ο Ο Ο Ο	Clostrid Clostri	ium diffic ium diffic 55101963 c Hits 7 2 3 4 3 5 22 5 7 5 8 5 1 5 9 5 4 5	Image: Control of the second	26] 42] Mass 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 1194.62 1246.56 1206.661 1250.615 922.526	Mass error (p -0.237 -0.0186 -0.617 -0.607 -0.206 -0.206 -0.254 0.0153 0.208 -1.2	RT (mins) 54.9 49.6 52.1 30.5 27.6 26.3 38.9 34.3	Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 3.01E+07 2.76E+07 2.31E+07 1.51E+07 1.91E+07 	, Conflic 1 1 1 1 1 1 1 1 1 1 1
gij254976383 (+2) 9	cell surfa cell surfa cell surfa olution r peptide i eptides Uniqu 63 18 61 16	ace prote ace prote on 522 ue Conflic 87 87 9peptides	in (S-layer in (S-layer tt: Protein Sco 7.07E+03 6.71E+03	precu precu	V Irsor f Irsor f V V V V V V	prote prote # 522 803 1302 2097 3 4 5 6	in) [(ions (Σ Ο Ο Ο Ο Ο Ο	Clostrid Clostrid Clostrid	ium diffic ium diffic 55101963 c Hits 7 2 3 4 3 5 22 5 7 5 8 5 1 5 9 5 4 5	Image: control of the state of the	Mass 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 1194.62 1246.56 1206.661 1250.615	Mass error (p -0.237 -0.0186 -0.617 -0.607 F 0206 F 0206 0.554 0.00153 0.208	RT (mins) 54.9 49.6 52.1 30.5 27.6 26.3 38.9	Charge 4 3 2 2 2 2 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 2.31E+07 1.51E+07 	Conflic 1 1 1 1 1 1 1 1 1 1

In this case the conflicting peptide assignments are with '**The same protein**' (from a different strain) which also contains 87 conflicts. A simple resolution to these conflicts is to right click on the conflicting protein and turn off all its peptides (based on lower number of peptides and score).

File Review port Data Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Statisti		Ident Peptie		QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report				Waters Con
periment design: AC			•													
roteins	No filter app	ied	Create		otide	ions o	of gi 2!	4976387					₹ ^N	lo filter	applied	Create.
Accession	Peptides Uni	que Conflicts	Score	-	#	Σ	④ Sco	re Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Conflict
) gi 254976387 (+5) 🜼	63 6	3 0	7.07E+03		1	0	1 5	0.6 5	-0.247	1076.586	-0.37	29.6	2		2.53E+07	0
gi 255101963 (+2)	0 (0	0	V	2	•	2 7	5.2 5	-0.834	2413.301	2.48	54.9	3		1.25E+08	0
gi 209571234	25 1	2 19	2.5E+03	V	18	•	0 1	28 5	0.216	2413.296	0.258	54.9	2		9E+07	0
gi 260682215 (+1)	24 1	l 19	2.13E+03		522	٠	0 6	3.7 2	-0.903	2413.295	-0.237	54.9	4		1.26E+06	0
gi 255101959	8 0	11	728	V	803	0	1 6	.3 4	-0.728	2413.295	-0.0186	49.6	3		3.01E+05	0
													-			
gi 254976383 (+2) Protein: gi 2549763 Protein: gi 2551019	963 cell su					ein) (Clostri				-0.617	49.6	2	-	2.73E+05	•
gi 254976383 (+2) "" Protein: gi 2549763 Protein: gi 2551019 ptide ion views Protein	387 cell su 963 cell su n resolution	rface protei	n (S-layer p	∢ precurso precurso	prot prot	ein) (ein) (Clostri	lium diffic	iile QCD-66 cile QCD-63	c26]	-0.617	49.6	2	-	2.73E+05	•
) gi 254976383 (+2)	387 cell su 963 cell su n resolution	rface protei rface protei e ion 522 iique Conflict	n (S-layer p	erecurso precurso Pe re	prot prot ptide	ein) (ein) (ions	Clostri	lium diffic lium diffic 55101963	m cile QCD-66 cile QCD-63	c26]	-0.617 Mass error (p	49.6 RT (mins)	2 Charge		2.73E+05 ▼ Abundance	•
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	Precurso Precurso Pe	prot prot ptide #	ein) [ein) [ions Σ	Clostric Clostric of gi 2 Scc 0 6	lium diffic lium diffic 55101963 re Hits 3.7 2	ile QCD-664 cile QCD-634 correlation	c26] q42] Mass 2413.295	Mass error (p -0.237	RT (mins) 54.9	Charge 4		 Abundance 1.26E+06 	Conflic
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict	in (S-layer p in (S-layer p t: Protein Score	Precurso Precurso	prote prote # 522 803	ein) [ein) [ions Σ	Clostric Clostric of gi 2 © Scc 0 6 1 6	lium diffic lium diffic 55101963 re Hits 3.7 2 1.3 4	ile QCD-666 cile QCD-636 cile QCD-636 cile QCD-636 cile QCD-636 cile QCD-636 cile QCD-666 cile CILE Cile Cile Cile Cile Cile Cile Cile Cile	Mass 2413.295	Mass error (p -0.237 -0.0186	RT (mins) 54.9 49.6	Charge 4 3		 Abundance 1.26E+06 3.01E+05 	Conflic 0
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	precurso precurso	prot prot # 522 803 1302	ein) [ein) [ions Σ	Clostric Clostric of gi 2 Clostric of gi 2 Clostric of gi 2 Clostric of gi 2 Clostric	lium diffic lium diffic 55101963 re Hits 3.7 2 1.3 4 23 5	ile QCD-666 cile QCD-636 cile QCD-636 correlation	Mass 2413.295 2413.295 2413.294	Mass error (p -0.237 -0.0186 -0.617	RT (mins) 54.9 49.6 49.6	Charge 4 3 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 	Conflic 0 0
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	Precurso precurso	prot prot # 522 803 1302 2097	ein) [ein) [ions Σ ο	Clostrie Clostrie of gi 2	lium diffic lium diffic 55101963 re Hits 3.7 2 1.3 4 23 5 22 5	Correlation	Mass 2413.295 2413.294 2413.294	Mass error (p -0.237 -0.0186 -0.617 -0.607	RT (mins) 54.9 49.6 49.6 52.1	Charge 4 3 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 	Conflic 0 0 0 0
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	re Pe	prot prot # 522 803 1302 2097	ein) [ein) [ions Σ ο ο	Clostric Clostric of gi 2 Clostric of gi 2 Clostric	lium diffic dium diffic 55101963 re Hits 3.7 2 1.3 4 23 5 22 5 1.7 5	Correlation	c26] q42] Mass 2413.295 2413.295 2413.295 2413.294 2413.294 1194.628	Mass error (p -0.237 -0.0186 -0.617 -0.607 -0.206	RT (mins) 54.9 49.6 49.6 52.1 30.5	Charge 4 3 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 	Conflic 0 0 0 0 0
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	re Pe	prot prot # 522 803 1302 2097 3 4	ein) [ein) [ions Σ	Clostric Clostric of gi 2 Clostric	lium diffic dium diffic 55101963 re Hits 3.7 2 1.3 4 23 5 22 5 1.7 5 9.8 5	Correlation	Mass 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 1194.628 1246.568	Mass error (p -0.237 -0.0186 -0.617 -0.607 -0.206 0.254	RT (mins) 54.9 49.6 49.6 52.1 30.5 27.6	Charge 4 3 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 	Conflic 0 0 0 0 0
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	re Pe	prot prot # 522 803 1302 2097	ein) [ein) [ions Σ Ο Ο Ο Ο Ο	Clostric Clostric Of gi 2 @ Scc 0 6 1 6 2 1 0 1 1 5 6 8 1 7	lium diffici 55101963 re Hits 3.7 2 1.3 4 23 5 22 5 1.7 5 9.8 5 5.1 5	Correlation	C26] Mass 2413.295 2413.295 2413.295 2413.294 2413.294 1194.628 1246.568 1206.661	Mass error (p -0.237 -0.0186 -0.617 -0.607 -0.206 0.254 0.0153	RT (mins) 54.9 49.6 49.6 52.1 30.5 27.6 26.3	Charge 4 3 2 2 2 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 2.31E+07 	Conflic 0 0 0 0 0 0 0 0
gi 254976383 (+2) Protein: gi 2549763 Protein: gi 2551015 ptide ion views Protein Conflicting proteins Accession	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	re Pe	prot prot # 522 803 1302 2097 3 4	ein) [ein) [ions Ω Ω Ω Ω	Clostric Clostric Of gi 2 P Scc 0 6 1 6 2 1 0 1 1 5 6 8 1 7 2 4	dium diffic dium diffic 55101963 re Hits 3.7 2 1.3 4 22 5 22 5 1.7 5 9.8 5 5.1 5 7.9 5	Terrelation	C26] q42] Mass 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 1194.628 1246.568 1206.661 1250.615	Mass error (p -0.237 -0.0186 -0.617 -0.607 -0.206 0.254 0.0153 0.208	RT (mins) 54.9 49.6 49.6 52.1 30.5 27.6 26.3 38.9	Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2		 ✓ Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 2.31E+07 	Conflic 0 0 0 0 0 0 0 0 0 0 0 0
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	Peere Pe	r prote r prote # 522 803 1302 2097 3 4 5 6 8	ein) [ein) [ions Σ Ο Ο Ο Ο Ο Ο Ο Ο Ο Ο	Clostric Clostric of gi 2 Costric Scc 0 6 1 6 2 1 0 1 1 5 6 8 1 7 2 4 0 4	dium diffic dium diffic 55101963 re Hits 3.7 2 1.3 4 22 5 5.22 5 5.22 5 5.1 5 5.1 5 7.9 5 5.4 5	U Correlation	Mass 2413.295 2413.295 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.295 2413.	Mass error (p -0.237 -0.0186 -0.617 -0.206 0.254 0.0153 0.208 -1.2	RT (mins) 54.9 49.6 52.1 30.5 27.6 26.3 38.9 34.3	Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		 Abundance 1.26E+06 3.01E+05 2.73E+05 3.01E+07 2.76E+07 2.31E+07 1.51E+07 1.91E+07 	Conflic 0 0 0 0 0 0 0 0 0 0 0 0 0 0
gij254976383 (+2) Protein: gij2549763 Protein: gij2551015 ptide ion views Protein Conflicting proteins Accession gij254976387 (+5)	387 cell su 963 cell su n resolution s for peptid Peptides Un 63	rface protei rface protei e ion 522 iique Conflict 63 0	in (S-layer p in (S-layer p t: Protein Scor 7.07E+03	re Pe	prot prot # 522 803 1302 2097 3 4	ein) [ein) [ions Ω Ω Ω Ω	Clostric Clostric of gi 2 Costric Scc 0 6 1 6 2 1 0 1 1 5 6 8 1 7 2 4 0 4	dium diffic dium diffic 55101963 re Hits 3.7 2 1.3 4 22 5 22 5 1.7 5 9.8 5 5.1 5 7.9 5	Correlation Correlati	Mass 2413.295 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 1194.628 1245.568 1206.661 1250.615	Mass error (p -0.237 -0.0186 -0.617 -0.607 -0.206 0.254 0.0153 0.208	RT (mins) 54.9 49.6 49.6 52.1 30.5 27.6 26.3 38.9	Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2		 ✓ Abundance 1.26E+06 3.01E+05 2.73E+05 1.6E+05 3.01E+07 2.76E+07 2.31E+07 	Conflic 0 0 0 0 0 0 0 0 0 0 0 0

Note: as you un-assign the peptides the number of conflicts update 'on the fly' in all the windows. Waters The science of WHAT'S POSSIBLE."

In this case the conflicting peptides are unassigned from the 'precursor' protein.

In some cases you can resolve the conflicts between 2 proteins on the basis of consistent peptide expression. In the example below the proteins share 19 conflicts. For the protein showing 2 clear patterns of expression you can un-tick all the peptides with conflicts in the corresponding peptides table



This leaves the peptides with the same expression pattern assigned with the appropriate protein thus resolving the conflicts.



Note: 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.

To resolve the remaining conflicts in the example shown below 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.

Accession	Peptides	Unique	Conflicts	Score	^	#	ŧ	Σ	e :	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Conflic	t: 1
🔰 gi 260682215 (+1)	24	24	0	2.13E+03		V 3	677		1	67.4	2	0.843	1791.962	0.0309	51.7	2		1.08E+05		-
gi 255101959	0	0	0	0		V 2	886	0	0	63	4	0.979	1198.656	0.594	28.3	2		8.45E+04	1	1
🌛 gi 254976383 (+2) 🛛 🧕	9	7	2	797		V	952	•	2	110	4	0.123	1889.8	0.489	24.7	2		2.6E+05	Lo	
) gi 126699078	5	2	5	653		✓ 1	162	0	0	68.7	3	0.581	1268.66	B 0.259	45.1	2		1.11E+05	✓ ₀	
🕽 gi 10180205 (+1)	3	1	2	266		V	677	•	0	71.3	5	0.952	1319.646	-0.616	29.9	2		1.52E+05	0	
	-	•	-			•													•	
Protein: gi 254976 Protein: gi 101802 ptide ion views Prote	205 Cwp	66 [[Clo	•	•	um d	ifficile	QCD	-660	c26]	1										
Protein: gi 101802 ptide ion views Prote	205 Cwp in resolutio	n [[Clo	ostridium	•	um d					i 1018(0205									
Protein: gi 101802	205 Cwp in resolutions for pe	n	n 3677] difficile]		Pept	ide ic	ons o	of g	i 1018(0205 Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Тад	✓ Abundance	Conflic	t
Protein: gi 101802 bide ion views Prote Conflicting protein Accession	205 Cwp in resolutions for peperide	n [[Clo	n 3677	•		Pept	ide ic	ons (Σ	of g		_	Correlation 0.117	Mass 1198.656	Mass error (p 0.594	RT (mins) 28.3	Charge 2	Tag	 Abundance 8.45E+04 	Conflic 1	t
Protein: gi 101802 ptide ion views Prote Conflicting protein	205 Cwp in resolutions for peperide	n	n 3677] difficile] Protein Sco		Pept	ide ic # 2886	ons (Σ	of g	i 10180 Score	Hits					-			Conflic 1 1	t:
Protein: gi 101802 ptide ion views Prote Conflicting protein Accession @ gi 254976383 (*2)	205 Cwp in resolutions for peperide	n otide ion Unique 7	n 3677 Conflict] difficile] Protein Sco 797		Pept	ide ic # 2886 3677	ons α Σ	of g	i 1018(Score 63	Hits 4	0.117	1198.656	0.594	28.3	2		8.45E+04	1	t
Protein: gi 101802 ptide ion views Prote Conflicting protein Accession 9 gi 254976383 (-2) 9 gi 255101959	205 Cwp in resolutions for peperide	n otide ion Unique 7	n 3677 Conflict	Protein Sco 797		Pept	ide ic # 2886 3677 5201	ons α Σ	of g	i 1018(Score 63 67.4	Hits 4 2	0.117 0.517	1198.656 1791.962	0.594 0.0309 0.255	28.3 51.7	2		8.45E+04 1.08E+05	1	t

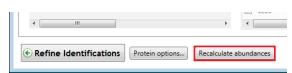
Favouring the protein with the higher score, resolve the conflict by switching off (or un-assigning) the peptide in panel F for the protein with the lower score. By doing this, the other 3 panels update to show the change in conflicts.

Accession	Peptides	Unique	Conflicts	Score	*		#	Σ	몓	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	✓ Abundance	Conflict: ^
🎯 gi 260682215 (+1)	24	24	0	2.13E+03		V	3677		1	67.4	2	0.843	1791.962	0.0309	51.7	2		1.08E+05	0 =
🎯 gi 255101959	0	0	0	0		V	2886	0	0	63	4	0.979	1198.656	0.594	28.3	2	-	8.45E+04	0
🎯 gi 254976383 (+2) 🔇	9	9	0	797		V	952	٠	2	110	4	0.123	1889.898	0.489	24.7	2		2.6E+05	0
🔇 gi 126699078	5	2	5	653		1	1162	0	0	68.7	3	0.581	1268.666	0.259	45.1	2		1.11E+05	0
🎯 gi 10180205 (+1)	1	1	0	136	-	1	677	٠	0	71.3	5	0.952	1319.646	-0.616	29.9	2		1.52E+05	0 _
▲ []	-		-	1		•													Þ
1 Protein: gi 25497	6383 ce	ell surfac	e protei	n [Clostridi	um d	iffici	le QC	D-6	6c26	6]									
Protein: gi 101802	205 Cw	p66 [[Cl	ostridium] difficile]															
Peptide ion views Prote	in resolutio	on																	
Conflicting protein	ns for pe	ptide io	n 3677			Pe	ptide	ions	of	gi 1018	0205								
Accession	Peptide	es Unique	Conflict	Protein Sco	re		#	Σ	中	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Conflict:
🎯 gi 254976383 (+2)	<mark>o</mark> 9	9	0	797			2886	0	0	63	4		1198.656	0.594	28.3	2	-	8.45E+04	0
🎯 gi 255101959	0	0	0	0			3677	0	1	67.4	2		1791.962	0.0309	51.7	2		1.08E+05	0
🕥 gi 10180205 (+1)	1	1	0	136		V	6201	٠	2	136	4	0.999	2319.16	0.255	59.6	2		6.03E+04	0 🧃
						V	10796	٠	1			0.999	2319.16	0.243	59.5	3		1.06E+04	0
• [Þ	•		_	_		_	III							Þ

Continue with Conflict resolution until no remaining conflicts in the Proteins table.

Accession	Peptides	Unique	Conflicts	Score	^		#	Σ	中	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Conflict
gi 126699959 (+1)	2	2	0	208		V	158		0	138	5	0.936	1489.803	0.108	50.4	2		1.98E+06	0
gi 126700584	1	1	0	59.8		V	1066	٠	0	59.6	3	0.949	1489.802	-0.793	50.4	3		1.15E+05	0
gi 126699811	1	1	0	104		V	159	٠	0	64.4	5	0.928	902.533	-0.56	41.6	2		5.62E+05	0
gi 126698351	1	1	0	62.8		V	1843	0	0	129	5	0.687	2002.073	-0.137	50.2	2		1.29E+05	0
gi 126698842	2	2	0	193			693	0	0	67.6	4	0.821	2002.074	0.0508	50.2	3		2.75E+05	0
					•	4						m					_		
Protein: gi 12669 Protein: gi 12669 tide ion views Prote	9128 rut	orerythr					30]												
Protein: gi 12669 tide ion views Prote	9128 rul ein resolutio	orerythr	in [Clost			30]		ions	of	gi 126699	9128								
Protein: gi 12669 ide ion views Protei	9128 rul ein resolutio	prerythr	in [Clost			30]		ions Σ	of s	gi 126699 Score		Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Confli
rotein: gi 12669 ide ion views Protei onflicting protei Accession F	9128 rut ein resolutions for pe	prerythr	in [Clost	ridium diffi otein Score		30]	- otide i	ions Σ				Correlation 	Mass 2002.074	Mass error (p 0.0508	RT (mins) 50.2	Charge 3	Tag	 Abundance 2.75E+05 	Confli
rotein: gi 12669 iide ion views Protei onflicting protei Accession F gi 126699128	9128 rut ein resolutions for pe	prerythr	in [Clost in 693 Conflict: Pr	ridium diffi otein Score	cile 6	30]	tide i #	Σ		Score						-	Tag		
rotein: gi 12669 iide ion views Protei onflicting protei Accession F gi 126699128	9128 rub ein resolutions for pe Peptides Ur 4	ptide io	in [Clost on 693 Conflict: Pr 0 402	ridium diffi otein Score	cile 6	30] Pep	• tide i # 693	Σ Ο		Score 67.6	Hits 4		2002.074	0.0508	50.2	3	Tag	2.75E+05	0
Protein: gi 12669 tide ion views Prote	9128 rub ein resolutions for pe Peptides Ur 4	ptide io	in [Clost on 693 Conflict: Pr 0 402	ridium diffi otein Score	cile 6	30] Pep	otide i # 693 581	Σ 0		Score 67.6 66.1	Hits 4	0.764	2002.074 1638.893	0.0508	50.2 50.4	3	Tag	2.75E+05 2.23E+05	0

Note: the abundances will need to be recalculated as a result of performing Conflict resolution. This is achieved by clicking on the Recalculate abundances, which appears during Conflict resolution



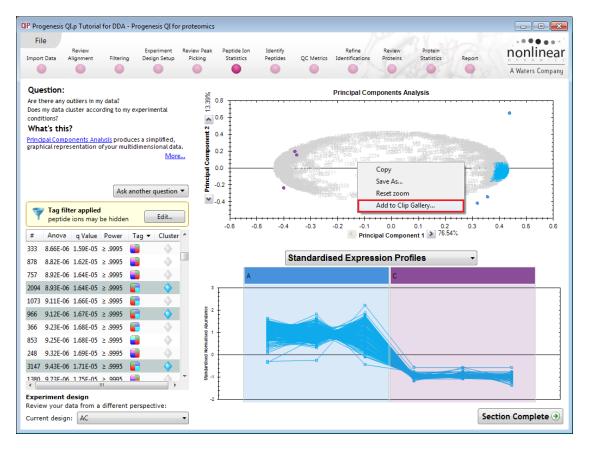
Appendix 9: Using Clip Gallery to Save and Export Pictures and Data

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

The saved images 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, displaying the PCA plot.

Right click on the Biplot View and select Add to clip gallery....

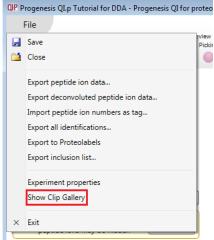


This will open a dialog displaying what is to be saved and allows you to alter the title and provides a description of the image for later reference.

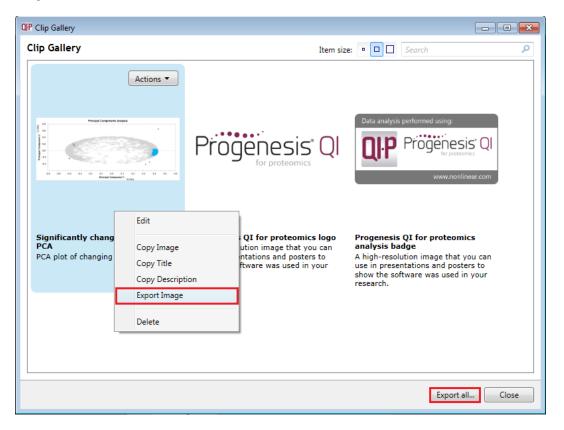
QIP Add to Clip Gallery
Enter a title and description for this clipping:
Principal Components Analysis
14
41
44
-0.5 -0.5 -0.4 -0.3 -0.2 -0.1 8.8 8.1 0.2 0.3 8.4 0.5 0.5 Principal Component 1, N.52%
Significantly changing proteins PCA
PCA plot of changing proteins
Add to Clip Gallery Cancel

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 10: Exporting protein identities to IPA

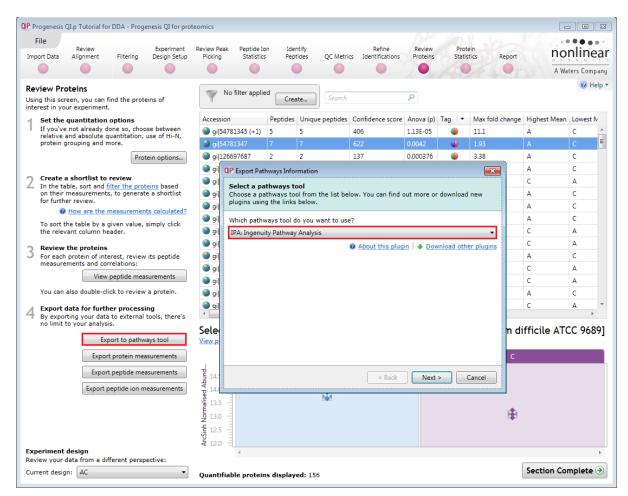
IPA (Ingenuity Pathway Analysis): is a commercial web-based software application for the analysis, integration and interpretation of data derived from any experiments that generate gene and chemical lists with expression data, such as metabolomics and proteomics.

Before exporting data to IPA, make sure that you can launch IPA using the default browser on your system. The plug-in relies on the browser to upload the Progenesis-exported data to IPA.

Note: the use of IPA requires a licensed copy, for licensing details please visit the IPA website.

First select the Protein data to export to the pathways tool using tag filtering to 'focus' the set to export.

Click Export to pathways tool and select IPA and click Next.



Select the appropriate identifier type used for proteins in the experiment.

Note: for proteins with different identifier types, filter the protein list to contain only proteins with single identifier type at a time, then perform the export operation and merge the resulting lists in IPA.

QP Export Pathways Information
Configure your export Choose which identifications to export and the type of analysis you want to perform.
Select the identifier type used for proteins in the experiment:
UniProt/Swiss-Prot Accession 🔹
UniProt/Swiss-Prot Accession
GenPept GI Number
▲ For an overrepresentation analysis, filter your data to only show proteins with significantly altered expression.
< Back Export proteins to IPA Close

Finally select the way you want to export the proteins to IPA.

If you are performing over-representation analysis or other types of analysis that do not require the protein expression data then you can make use of the **Create a list** option to export data to IPA.

Note: using this option you will be asked to filter your data to only show proteins with significantly altered expression

QIP Export Pathways Information	×
Configure your export Choose which identifications to export and the type of analysis you wa	nt to perform.
Select the identifier type used for proteins in the experiment:	
UniProt/Swiss-Prot Accession	•
Select the way you want to export the proteins to IPA:	
Upload expression dataset	•
Create a list	
Upload expression dataset	
Baseline: Comparison:	
■ A ▼ 🔄 🔲 C	•
If greater, gives a negative fold change. If greater, gives a pos	itive fold change.
< Back Export proteins to I	PA Close
a buck	

If you require to perform enrichment analysis or other types of analyses that use protein expression data, select the option to **Upload expression dataset** to IPA.

Note: You need to select the type of protein identifier used by at least one protein whose expression data you want to export. If there are proteins with different identifier types, you will be able to select additional identifier types in IPA after the dataset is uploaded.

The operation exports data from a single observation at a time - select the two experimental conditions (taken from the current **experiment design as set in Progenesis QI for proteomics** you wish to compare as that observation).

Click Export proteins to IPA .

	/apps.ingenuity.c ♀ ← 🔒 ♥ 🛲 Ingenuity – Ingenuity Login 🗙	
1.1.0.	vorites <u>T</u> ools <u>H</u> elp	
INGEI	NUITY	^
Welcome! F	Please login	Contact Customer Support
Email	smith@work.com	Customer Support Phone: 650.381.5111
Password	•••••	Hours: 6am - 5pm (PST) Monday - Friday (excluding holidays)
	Remember my password	support@ingenuity.com
		For Product and Sales related inquiries contact:
	Sign Up Forgot Password	650.381.5056 sales@ingenuity.com
		×
<		>

If you have access to a licenced copy of IPA then log in

IPA will open displaying the imported data from Progenesis QI for Proteomics.

	t <u>V</u> iew <u>W</u> indow	Пер						Provide Feedback Support Janucz Nykiel Clo
		Ger	nes and Chemicals	Diseases and Fun	ctions Pathways and	d Tox Lists		
w ×		Ente	er gene names/symbo	ols/IDs or chemical/	drug names here		<u>5</u> E/	ARCH Advanced Search 👰
aset l	Ipload - New Dataset	2016-05-20 10:29 AM	л					Dataset Upload Workflow Instructions
Select	File Format:	Flexible For	mat	▼ Mor	re Info			Data Upload Workflow
Conta	ins Column Header:	🔿 Yes 🌘	No					Use Dataset Upload to import your dataset file into IPA.
Select	Identifier Type:	UniProt/Sw	viss-Prot Accession	▼ Sper	cify the identifier type fo	ound in the dataset.		Once uploaded, many different analysis options exist including
Arrav	platform used for expe	eriments: Not specif	ied/applicable	▼ Sele	ct relevant array platfor	m as a reference set fo	or data analysis.	Biomarker Filter, Molecular Tox and Core Analyses. Review the different type of analyses and see which one best fits
1	· · · ·	· · · ·			ations. For observations,			your poods
Use ti	ie dropdown menus to	specify the column h	ames that contain id	entitiers and observa	itions. For observations,	select the appropriate	e expression value ty	
aw D	ata (512) \ Dataset Sun	nmary (1)						1. To upload a dataset file, <u>click here</u> .
	,	· · · · ·	·					* Open 🛛
	ID 💌	Observation 1 💌	Ignore 🔹	Ignore	▼ Ignore	▼ Ignore	▼ Ignore	Look jn: 🗀 Multiple Rank 👻 🔯 😫 😂
		Exp Fold Chan 💌						Batch (mult-timepoint) Affy_with_p-value_and_fold.txt
1	P34731	-1.00902155389087						LL_with_Fold and_Normalized.txt
2	C4YQR7	1.03618874042727						
3	C4YMC3	-1.1630436093987						
4	C4YQN7	-1.14248088552437						
5	P43098	-1.00167260838215						
	P43098 C4YR46	-1.00167260838215 1.11144195146305						
5	C4YR46 O13430							Fle tjune:
5 6	C4YR46	1.11144195146305						Fås ljøne: Fås of type: [Al Fås:
5 6 7	C4YR46 O13430 P46587 P82610	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931						
5 6 7 8	C4YR46 O13430 P46587 P82610 P46598	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961						Files of Type: Al Files
5 6 7 8 9 10 11	C4YR46 O13430 P46587 P82610 P46598 Q96VB9	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 -1.10613636924535	 					Files of type: Al Files
5 6 7 8 9 10 11	C4VR46 013430 P46587 P82610 P46598 Q96VB9 P41797	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 -1.10613636924535 1.50747573107542	 					Ples of type: AFRes
5 6 7 8 9 10 11	C4VR46 013430 P46587 P82610 P46598 Q96VB9 P41797 P28877	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 -1.10613636924535 1.50747573107542 -1.20624862059398	 					Files of type: Al Files
5 6 7 8 9 10 11 12 13 14	C4YR46 O13430 P46587 P82610 P46598 Q96VB9 P41797 P28877 C4YK39	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 -1.10613636924535 1.50747573107542 -1.20624862059398 1.05494661189734	 					Files of type: If Files Open Coreal 2. Select the dataset file from your computer and click the Open button.
5 6 7 8 9 10 11 12 13 14 15	C4YR46 013430 P46587 P82610 P46598 Q96VB9 P41797 P28877 C4YK39 Q59KZ1	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 -1.10613636924535 1.50747573107542 -1.20624862059398 1.05494661189734 1.05494661189734	 					Plead type: Network Open Control 2. Select the dataset file from your computer and click the Open button. 3. Select Flexible format for the file format from the
5 6 7 8 9 10 11 12 13 14 15 16	C4YR46 O13430 P46587 P82610 P46598 Q96VB9 P41797 P28877 C4YK39 Q59KZ1 C4YL05	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 -1.10613636924535 1.50747573107542 -1.20624862059398 1.05494661189734 1.0399981345519 -1.0312270752554	 					Flee of type: If Flees Open Corect Que to the dataset file from your computer and click the Open button.
5 6 7 8 9 10 11 12 13 14 15 16 17	C4YR46 O13430 P46587 P82610 P46598 Q96VE8 P41797 P28877 C4YK39 Q59KZ1 C4YK05 P46273	111144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 1.10613636924535 1.50747573107542 1.20624862059398 1.05494661189734 1.035494661189734 1.0312270752554 1.31165982609058	••• ••• ••• ••• ••• ••• ••• ••• ••• ••					Resorting Reference Select the dataset file from your computer and click the Open button. S. Select Flexible format for the file format from the dropdown menu.
5 6 7 8 9 10 11 12 13 14 15 16 17 18	C4YR46 O13430 P46587 P82610 P46598 Og96V89 P41797 P28877 C4YK39 Q59KZ1 C4YL05 P46273 C4YL05 P46273 C4YL18	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 1.1061363624535 1.50747573107542 1.20524862059398 1.05949661189734 1.03999813845519 1.03999813845519 1.31165982609058 1.32690493012736						Plead type: Alfree Copen Copen 2. Select the dataset file from your computer and click the Open button. 3. Select Flexible format for the file format from the
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	C4YR46 O13430 P46587 P82610 P46598 Q6V89 P41797 P28877 C4YK39 Q59KZ1 C4YK39 Q59KZ1 C4YL05 P46273 C4YLB O94039	1.111441951.46305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004223965961 1.10613636924535 1.50747573107542 1.20624862059398 1.05494661189734 1.031220725254 1.31155982609058 1.25690493012776						Plead type: Alfiles Come Cancel Come Cancel Select the dataset file from your computer and click the Open button. Select Flexible format for the file format from the dropdown menu. Select an Identifier Type from the dropdown menu. IPA
5 6 7 8 9 10 11 12 13 14 15 16 17 18	L4YR46 O13430 P45587 P82610 P4598 O96VB9 P41797 P28877 C4YK39 Q59KZ1 C4YL05 P46273 C4YL8 O94039 O13287	1.11144195146305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004923965961 1.1061363624535 1.50747573107542 1.20524862059398 1.05949661189734 1.03999813845519 1.03999813845519 1.31165982609058 1.32690493012736	** ** ** ** ** ** ** ** ** ** ** ** **					Prever type: Affect Open button. Select Flexible format for the file format from the dropdown menu. Select an Identifier Type from the dropdown menu. IPA supports many identifiers and symbols and will attempt to
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	C4YR46 O13430 P46587 P82610 P46598 Q6V89 P41797 P28877 C4YK39 Q59KZ1 C4YK39 Q59KZ1 C4YL05 P46273 C4YLB O94039	1.111441951.46305 1.0005161363124 1.16697510036399 1.14143675382931 1.06004223965961 1.10613636924535 1.50747573107542 1.20624862059398 1.05494661189734 1.031220725254 1.31155982609058 1.25690493012776	** ** ** ** ** ** ** ** ** ** ** ** **					Prevar type: AFRee Copen Dutton. Select Flexible format for the file format from the dropdown menu. Select an Identifier Type from the dropdown menu. IPA supports many identifiers and symbols and will attempt to guess at the type of identifier in your dataset file. To

You can now explore your protein expression data using the tools available in IPA. Instructions on how to manage and explore your imported data set are provided by IPA.

Appendix 11A Search engine parameters for Mascot

The parameters applied to the Mascot search that yielded the search results used in this user guide are shown below:

MASCOT	MS/MS Ions Search			
Your name	andy b	Email	andy.borthwick@nonlinear.com	
Search title	Progenesis QIp search			
Database(s)	Mark_Test2	Enzyme	Trypsin 🗸	
	MSDB ^ NCBInr	Allow up to	1 V missed cleavages	
	NCBInr_UserGuide	Ouantitation	None	\checkmark
-				
Taxonomy	Firmicutes (gram-positive bacter	ia)	~	
Fixed modifications	Carbamidomethyl (C)	> <	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term)	^
	Display all modifications		Amidated (Protein C-term) Ammonia-loss (N-term C)	
Variable modifications	Oxidation (M)	> <	Biotin (K) Biotin (N-term) Carbamyl (K) Carbamyl (N-term) Carboxymethyl (C)	~
Peptide tol. ±	9 ppm V # ¹³ C 0 V	MS/MS tol. ±	0.6 Da 🗸	
Peptide charge	2+ 🗸	Monoisotopic	• Average O	
Data file	C:\Users\andy.borthwick\Document	Browse		
Data format	Mascot generic V	Precursor	m/z	
Instrument	ESI-TRAP V	Error tolerant		
Decoy		Report top	AUTO V hits	
	Start Search		Reset Form	

Database : NCBInr (circa 04/16) was used with the Taxonomy restriction set to Firmicutes Fixed modifications: Carbamidomethyl (C) and variable modification Oxidation (M) Peptide Tol: 9ppm Instrument: ESI-Trap

Appendix 11B Use Additional Export Fields in Mascot for PRIDE

For exporting mzldentML of results to PRIDE from Progenesis QI for proteomics you **must select** additional Mascot Export Fields when exporting your search results from Mascot: (Protein sequence (Protein Hit Information) and Start and End (Peptide Match Information))

	Protein Hit Information	
	Score	\checkmark
	Description*	\checkmark
	Mass (Da) [*]	
	Number of queries matched	\checkmark
	Percent coverage**	
	Length in residues**	
	pI**	
	Taxonomy**	
	Taxonomy ID**	
	Protein sequence**	
	emPAI	
* Occasionally requires information to be retrieved from external utilities, which can be slow		
** Always requires information to be retrieved from external utilities, which can be slow		
	Peptide Match Information	
	Experimental Mr (Da)	
	Experimental charge	
	Calculated Mr (Da)	\checkmark
	Mass error (Da)	\checkmark
	Start	
	End	\checkmark
	Number of missed cleavages	

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