

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	54
Stage 13: Exporting Protein Data	61
Exporting Protein Data to Pathways Tool(s)	62
Stage 14: Exporting identifications for submission to PRIDE	65
Stage 15: Protein Statistics	67
Stage 16: Reporting	68
Creating an Inclusion list	70
Appendix 1: Stage 1 Data Import and QC review of LC-MS data set	72
Appendix 2: Stage 1 Processing failures	76
Appendix 3: Stage 1 Data QC review and addition of exclusion areas	78
Appendix 4: Licensing runs (Stage 3)	79
Appendix 5: Manual assistance of Alignment	80
Appendix 6: Within-subject Design	84
Appendix 7: Power Analysis (Peptide Ion Stats)	86
Appendix 8: Resolve Conflicts	87
Appendix 9: Using Clip Gallery to Save and Export Pictures and Data	91
Appendix 10: Exporting protein identities to IPA	93
Appendix 11: Exporting protein identities to MetaCore	96
Appendix 12A Search engine parameters for Mascot	98
Appendix 12B Use Additional Export Fields in Mascot for PRIDE	99
References	99

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 72) 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 support@nonlinear.com 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.

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
•	•	•	•	•	•	0	•	•	0		



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 QIP_Tutorial DDA.Progenesis QIP Archive** file using the **Open** button and press Open.

File				
xperiments				nonline
•				A Waters Comp.
erform analysis Combine analysed fr	actions			New to Progenesis QI for
ecent experiments		Search	Q	Here are some resources to help you
QIP Open Experiment			—	get started with Progenesis QI for proteomics:
😋 🔍 🛡 🐌 « Progenesis 🕨 Progenesis	QI.p Tutorial for DDA 🔹	 ✓ Search Progenesis QI. 	p Tutoria 🔎	<u>The Progenesis QI for proteomics</u> workflow
Organize 🔻 New folder		1		User guide and tutorial data
🐚 ShareFile	 Name 	*	Date modifi	<u>Frequently-asked questions</u>
BitTorrent Sync	🚮 Progenesis QI	p Tutorial for DDA.Progen.	18/04/2016	Quickly so to an ion man location
Deskton	E			Want to quickly validate your sample
Libraries				running by zooming to a known ion?
Documents				500
👌 Music				and 1
Pictures				Go To Location
Videos				
Andy Borthwick				augusta fagas materialment ateria
Network				50
AVSEESVA	III		•	「日本配修長になった。「後方法
File name: Progenesis	OLn Tutorial for DDA.Progenesis(Experiments and Archi 	ives (*.Prc 🔻	Jump to a specific m/z and RT using the
	.,		<u>.</u>	Go To Location tool in the top-left
		Open 🔻	Cancel	corner of the ion maps.
				Latest blog posts
				<u>Missing values: the Progenesis co-</u>
her experiments				detection solution
🖄 New 🛛 🖉 Open			2 0 5070 00	Identification scoring in Progenesis OI
			3.0.5978.29	 Missing values: the bard truths

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.

QIP 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 existin	g experiment								
Experiment to repla	ace: Progenesis QI.p Tutorial for DDA 🔹								
Oreate a new expension	eriment								
Experiment name:	Progenesis QI.p Tutorial for DDA								
Save to folder:	and Demo Suites\Progenesis QI.p Tutorial for DDA Browse								
	Import Cancel								

Then press Im	port.
---------------	-------

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).

OP Progenesis O	p Tutorial for DD	A - Progenesis OI f	or proteomi	s							
File Import Data	Review Alignment Filte	Experimen ering Design Setu	t Review P Ip Pickin	eak Peptide Ion g Statistics	Identify Peptides	QC Metrics	Refine Identifications	Review Proteins	Protein Statistics	Report	A Waters Company
Import Data			A1				<< Acti	ons 🔹 >			🕢 Help 🕶
1 Select yo Select one click the II Format: About 2 Perform a Click the b processing possible, b	ur run data to f the available mport button: mzXML files this data format automatic proce putton below to si p. This will analys before suggesting Start processing can b ing.	data formats ther Import Download of assing tart automatic e your data as far the next step. automatic process be started while ru	thers 82	612.5	825.0 1037	5 1250.0	1462.5	1675.0 1887. m/z *	About t • MS • MS/ • Tota • Mas	his run peak count: 1,, MS count: 8,98 al ion intensity: <u>ked areas</u> : nor	109,718 13 2.188e+009 1e
Imported rups	earn more about	automatic proces	ssing G		n 🤺 🖻	M.					
States			80			÷					
A1	A2	A3									
۹ <i>۹</i> ۴.			100	n Time (min)							
C1	C2	C3	•	 Retention 			Zoom:	0.9.2		Sec	tion Complete)

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 72). 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 78). This is not required for this data set.

<<	Act	ions 🔹 💽 >>		
1500		Mask areas for	peak picking	
	×	Remove run	Delete	
	_		Masked areas: none	886+009

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

Now start the Automatic Processing.



Stage 2A: Automatic Processing of your data

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

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



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

- 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.

QIP Start automatic processing	
Select an alignment reference To compensate for drifts in retention time, all runs in the 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	
Use the most suitable run from candidates that I select	
Use this run:	
A1 -	
For information on choosing the alignment reference, and select your own candidates, please see the <u>online guidanc</u>	why you might want to <u>e</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 72 for more details on using the other options).

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

GIP 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 the peaks in your samples.	their isotopic
After the automatic alignment is finished, do you want to start automatic Perform peak picking	s 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	×	٩	P Peak Picking Parameters	—
Runs for peak picking Peak picking lim	its Maximum charge Retention time limits		Runs for peak picking Peak picking lin	its Maximum charge Retention time limits
Choose runs for peak picking			Retention time limits	
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 stil 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 minimum and maximum retention time for peak picking. Tons that elute before or after these values will be ignored.	 ✓ Ignore ions before 10 minutes ✓ Ignore ions after 75 minutes
	OK Cancel			OK Cancel

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.

QIP Start automatic pro	ocessing		- • ×
Experiment design Experiment designs experimental condit	n allow you to group a ions.	nd compare your samples a	ccording to their
By defining an exper calculated automatic	iment design in advand ally.	e, statistical measures such a	as ANOVA can be
📝 Set up an expe	riment design		
Enter a name fo	r the experiment desig	n:	
AC			~
Load the criteria	for grouping runs from	m this file:	
			Browse
Group runs by:	<no groups="" valid=""></no>		Ŧ
What file formats an	re supported?		
		< 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	Α	02/03/2011	Fridge A
A2	А	02/03/2011	Fridge A
A3	Α	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.

Experiment designs allow you to group and compare your samples according to their experimental conditions. By defining an experiment design in advance, statistical measures such as ANOVA can be					
calcula S	et up an expe	any. riment desian			
Er	nter a name fo	r the experiment design:			
A	AC		~		
Lo	ad the criteria	a for grouping runs from this file:			
E	:\Customer Da	ata\Progenesis QIp_4.2 Tutorials and Demo Suites\Prog	Browse		
G	roup runs by:	Conditions	-		
		Conditions			
	Date of Collection				
-					
What f	ile formats a	re supported?			

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.

📭 Start automatic processing 📃 📼 💌	QIP 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 sample 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 mean calibrant abundance measured across runs 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	Relative Quantitation using all peptides • Uses all peptides identified as part of a protein • Allows comparison of a single protein across runs
protein s. 	Absolute Quantitation for HCP using Hi-N • Requires a calibrant protein to calculate absolute amounts Uses calibrant abundance measured per run • Uses the most abundant N peptides • Calculate the amount of HCPs are run in find and po

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 A2 as the alignment reference

Automatic processing	
Current step: Choosing an alignment reference	
Importing runs: 6 of 6 processed Selecting reference: Choosing an alignment reference Aligning runs: Pending Peak picking: Pending Protein quantitation: Pending Crancel Cancel Once Alignment completes Peak Picking comment PP Automatic Processing (51%)	UP Automatic Processing Automatic processing Current step: Calculating automatic alignment Importing runs: 6 of 6 processed Selecting reference: A2 Aligning runs: 0 of 5 processed Peak picking: Pending Creating design: Pending Protein quantitation: Pending Cancel Concel
Automatic processing	
Current step: Analysing	
Automatic processing Current step: Analysing ✓ Importing runs: 6 of 6 processed ✓ Selecting reference: A2 ✓ Aligning runs: 5 of 5 processed ● Peak picking: Picking	QIP Processing Complete Automatic processing complete. Time taken: 2 minutes 30 seconds

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	A Waters Company
Identify P	eptides			MS/MS	5pectra							

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 76)

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

Stage 2B: After Automatic Processing

When Processing completes, depending on what stages you selected to perform, the Automatic Processing Dialog displays 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.

Progenesis QI.p Tutorial for DDA - Progenesis QI for proteomics				
File Review Experiment Review Pe Import Data Licensing Alignment Filtering Design Setup Picking	ik Peptide Ion Identify Statistics Peptides QC Met	Refine Review ics Identifications Proteins	Protein Statistics Report	nonlinear
	• • •	0 0 0	0.05	A Waters Company
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	borthwick\Documents\Customer Data\P	ogenesis QI.p v3.0 Tutorials and De	mo Suites\Proge Unlicensed	J V
lease licence code which can be obtained from C:\Users\andy	borthwick\Documents\Customer Data\P	ogenesis QI.p v3.0 Tutorials and De	mo Suites\Proge Unlicensed	J V
a sales representative. C:\Users\andy	borthwick\Documents\Customer Data\P	ogenesis QI.p v3.0 Tutorials and De	mo Suites\Proge Unlicensed	
Once licensed, your runs can be analysed on C:\Users\andy	borthwick\Documents\Customer Data\P	ogenesis QI.p v3.0 Tutorials and De	mo Suites\Proge Unlicensed	
any installation of the software. The licence is C:\Users\andy	borthwick\Documents\Customer Data\P	ogenesis QI.p v3.0 Tutorials and De	mo Suites\Proge Unlicensed	
automatically included when archiving an C:\Users\andy experiment.	borthwick\Documents\Customer Data\P	ogenesis QI.p v3.0 Tutorials and De	mo 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.	able to save experiment You cannot save the analysis with analysis will be lost.	out a valid license. If you close no	W your	
If you have just installed a dongle, <u>click here</u> ,		OK Car	ncel	
	Run licence	ode:	Use Lic	ence Code
1			Secti	on 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 79)

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 (A2).

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:



• In the Run table click on Run C1 to make it current. You will now be looking at the alignment of C1 to A2 in the unaligned view. Now drag out an area to review on the **lon 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, the Ion 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 Ion Intensity Map as coloured squares, acts as a guide drawing your attention to areas of the alignment. These range from Good (Green) through OK (Yellow) to Needs review (Red). 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 80).



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.

Poviow Alignment	
Sample ions are aligned to compensate for drifts in	
retention time between runs.	QIP Protected from editing
Learn about the visualisations shown here	Delete existing analysis?
For maximum reproducibility, the software can automatically align your runs.	If you change the alignment, it will invalidate the current analysis including peptide ion pattern and IDs, editing, and tags.
Align runs automatically	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 	
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 80)

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 C1, 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

QP Peak Picking Parameters	×	QIP Peak Picking Parameters
Runs for peak picking Peak picking lim	its Maximum charge Retention time limits	Runs for peak picking Peak picking limits Maximum charge Retention time limits
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 run which is left un-ticked will not affect the peptide ion outlines, it 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	 Sensitivity You can adjust the sensitivity of the peak picking algorithm using 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. Absolute ion intensity % Base Peak The automatic sensitivity method uses a noise estimation algorithm to determine the noise levels in the data. The higher the sensitivity value, the more peptide ions will be detected. fewer default more as minimum peak width, any ion that has eluted over a shorter period will be rejected.
	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	Apply a minimum peak width				
a minimum peak width, any ion that has eluted over a shorter	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.

QP 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 You can set the maximum charge of ions to be detected. Ions with a charge greater than this value will be rejected. Maximum ion charge:	Retention time limits You can set the minimum and maximum retention time for peak picking. Ions that elute before or after these values will be ignored. Image: Construction of the peak picking. Ions that elute before or after these values will be ignored. Image: Construction of the peak picking. Ions that elute before or after these values will be ignored.
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 14473 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 (814 peptide ions)				
Charge 2 (5929 peptide ions)				
Charge 3 (5341 peptide ions)				
Charge 4 (1708 peptide ions)				
Charge 5 (465 peptide ions)				
Charge 6 (79 peptide ions)				
Charge 7 (37 peptide ions)				
Charge 8 (21 peptide ions)				
Charge 9 (16 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 914 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 normalis	ation

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.

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'.

QP Progenesis	IP Progenesis QL p Tutorial for DDA - Progenesis QI for proteomics														
File	Boviour		Experiment	Review Reak	Pontido Ion	Identify		Dofine	Bouieur	Brotoin					
Import Data	Alignment	Filtering	Design Setup	Picking	Statistics	Peptides	QC Metrics	Identifications	Proteins	Statistics	Report	nonlinear			
											0-10	A Waters Company			
Review normalisation Normalisation Graphs Normalisation Method															
Normalisatio different sar	on is required to nple runs.	allow comp	arisons across	Normalise to	Normalise to all proteins										
By assuming are unaffect use the fact to normalise	g that a significa ed by experime or by which the back to its refe	nt number o ntal conditio sample as a rrence.*	of peptide ions ns, we can whole varies	Normalise to all proteins Normalise to a set of housekeeping proteins Don't use any normalisation											

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.

Report		A Waters Com
Before	During	After
x x1	x2	X3
Y Y1	Y2	Y3
Z Z1	Z2	Z3
	Y Y1 Z Z1	Y Y1 Y2 Z Z1 Z2 Y Y1 Y2 Y1 Y2 Y2 Y1 Y2 Y2 Y2 Y2 Y2 Y2 Y2 Y2 Y2 Y2

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 84

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.

IP Create New Experiment Design
Enter a name for the experiment design:
AC
How do you want to group the runs?
Or a construction of the second se
Copy an existing design:
O Import criteria from a file: Browse
Group runs by: <pre><no groups="" valid=""></no></pre>
What file formats are supported? Create design Cancel

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

OP Progenesi	s QI.p Tutorial f	or DDA - Pro	genesis QI for p	roteomics								- • •
File	Review	Filteries	Experiment	Review Peak	Peptide Ion	Identify	05.1111	Refine	Review	Protein	12	nonlinear
	Alignment	ritering	Design Setup	Picking	Statistics	Peptides	QUMetrics		Proteins	Statistics	Report	A Waters Company
			•					- /	100		1.12	A waters company
ACIA	New											C neip
Setup cor	nditions			Runs	Add Selected F	luns to Cond	ition 🔻 Se	earch	ر	Q		
Setup the c (e.g., contro your sample	onditions that (ol, drug A, etc) es to the correc	you want to , and then a ct condition.	compare below assign each of	′ 🗖	Add to nev	v condition		C2			C3	
A			Delete		A			Sec.			- 1946 -	
			A1 <u>Remove</u>									
			A2 Remove									
			A3 <u>Remove</u>									
Add condit	tion											
											Secti	on Complete 🏵

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**.

QIP Create New Experiment Design												
Enter a name for the experiment design:												
AC-2												
How do you want to group the runs?												
Oroup the runs manually												
Copy an existing design:	AC	Ŧ										
Import criteria from a file:	DDA_v3.0 Tutorial Groups.csv	Browse										
Group runs by:	Conditions	•										
What file formats are support	Conditions Date of Collection 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 for DDA - Progenesis QI for pro	oteomics	
File Review Experiment	Review Peak Peptide Ion Identify Refine	Review Protein poplingor
Import Data Alignment Filtering Design Setup	Picking Statistics Peptides QC Metrics Identifications	Proteins Statistics Report
	\bullet \bullet \bullet \bullet	A Waters Company
AC AC-2 I X New		Ø Help ▼
Setup conditions	Runs Add Selected Runs to Condition Search	٩
Setup the conditions that you want to compare below (e.g., control, drug A, etc), and then assign each of your samples to the correct condition.	C3	
A Delete	23/2	
A1 <u>Remove</u>		
A2 <u>Remove</u>		
A3 <u>Remove</u>		
C Delete		
C1 Remove		
C2 <u>Remove</u>		
Add condition		
		Section 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.

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.

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

•••

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

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







Run Expression profile

ArcSinh Normalised Abundance

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 1005.4 m/z and 31.7 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.

eview Peak Picking						
Vo filter applied						
#	Anova (p)	q Value	Fold	Tag 💌	Not	
.2255	0.343	0.07	41.8		A	
7734	0.343	0.07	1.2		10	
5755	0.343	0.07	1.17		10.	
184	0.344	0.0701	1.13			
7804	0.344	0.0701	1.82		12	
3392	0.344	0.0702	1.33		12	
5048	0.344	0.0702	1.25		10.	
7334	0.344	0.0702	1.56		10	
5482	0.344	0.0702	1.28			
3195	0.345	0.0703	6.91		10.	
-	!!				- P	

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.

Review Peak Picking														
Y	Create													
#	Anova (p)	q Value	Fold	Tag 👻 Not										
12255	0.343	0.07	41.8	A										
7734	0.343	0.07	1.2	10.										
5755	0.343	0.07	1.17											
184	0.344	0.0701	1.13											
7804	0.344	0.0701	1.82											
3392	0.344	0.0702	1.33	1										
5048	0.344	0.0702	1.25	1										
7334	0.344	0.0702	1.56	10.										
5482	0.344	0.0702	1.28											
8195	0.345	0.0703	6.91	· ·										
×	11			÷										

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

QIP Pro	genesis	IP Progenesis QLp Tutorial for DDA - Progenesis QI for proteomics																	- • •	
Fil	e t Data	Revi Alignr	ew ment	Filtering	Experime Design Se	ent Re etup	view Peak Picking	Peptide Ion Statistics	Identify Peptides	QC Metrics	Iden	Refine tifications	Review Proteins	Protein Statistics	Report					A Waters Company
Revi	ew Pe	ak Pic	king																	
-	No fil	lter app	lied																	Create
#	Ano	va (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 Peptide Scor	e Peptide Mod
1	0.41	4 (0.0823	1.05		Add				805.4410		2413.301	55.594		1.24E+08	1.08E+08				A
17	0.18	3 (0.0412	1.1		10.		С	Α	1207.6552	2	2413.296	55.594	4.64	9.37E+07	5.04E+07	7.62	102		
39	5.49	E-08 4	4.05E-07	1.16E+05		12		с	Α	1100.5857	3	3298.735	45.749	3.15	8.53E+07	7.87E+07	15.9	19		
61	2.34	E-06 5	5.8E-06	1.6E+03		12		Α	С	1176.2272	3	3525.660	48.790	8.8	6.83E+07	2.22E+07	23.5	55		
9	3.43	E-08	3.05E-07	1.35E+03		10.		Α	С	656.8613	2	1311.708	44.596	4.21	6.1E+07	1.17E+08	8.98	39		
19	1.87	E-07 9	9.36E-07	805		12		С	А	988.9848	2	1975.955	51.484	4.82	5.3E+07	9.19E+07	14.5	48		
10	6.3E-	-08 4	4.29E-07	134		10.		С	Α	663.8693	2	1325.724	47.557	4.77	5.18E+07	1.69E+08	6.45	67		
23	3.4E-	-07 :	1.41E-06	6.37E+03		12		С	А	900.9713	2	1799.928	40.004	2.87	4.71E+07	8.16E+07	24.6	15		
54	3.63	E-06 8	B.06E-06	3.25E+03		12		Α	С	1061.0071	2	2120.000	54.334	6.1	4.26E+07	2.52E+07	26.6	64		
38	5.04	E-06 1	1.05E-05	1.28E+03		10		Α	С	997.4478	2	1992.881	32.637	2.53	4.19E+07	3.81E+07	34.1	29		-
۲ 📃																				Þ
		Delete	1 selecte	d nentide id	n															
							Run Exp	ression profile												
IDD	isplay	2D Dis	play 3D	Display			Run:		804	80	5	806	80	07 808	809			612.5 82	5.0 1037.5 1250.0 1	462.5 1675.0 188
							Aggregat									m/z •				m/z •
							k∂ Select	Edit			-	11 11		E H H		1 1	5	constant in	Same and the	••••
							🖻 Split	Merge	<u> </u>		×						100	h al a	William	and the second sec
							📮 Add	🗖 Delete			R.		H 1 -			٢F		t said	i i a na	
							🍤 Undo	陀 Redo	25								3	1.5		
							Charge sta	ite: 🚺 1 🛑 2	, iii		+	-			19		uiu .	1.1	4 ²	
								3 4	ue eo		4	G 11 1					8 E			
									D T		-					a a a a a a a a a a a a a a a a a a a	- 5			
									65 Enti		+			di di	tt b	L	100 tenti			
									- Re								Re			
											1							1		
																			Secti	on Complete 🏵

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

QIP Prog	enesis QI.p 1	Futorial for	DDA - Prog	jenesis QI f	for prote	omics											
File		aviau		Eveneric	ment	Review Beak	Reptide Ion	Identify		Patra	Paulau	Protein					
Import	Data Alig	gnment	Filtering	Design :	Setup	Picking	Statistics	Peptides	QC Metrics	Identifications	Proteins	Statistics	Report				nonlinear
													•				A Waters Company
Revie	w Peak P	icking															
7	No filter a	pplied															Create
#	Anova (p)	q Value	Fold	Tag 💌	Notes		Highest Mean	Lowest Me	an m/z	z Mass	RT (mins)	RT window (mins)	Abundance	Intensity	Max CV (%)	MS/MS Protein	Peptide Score P
2286					Ad												*
3496					Ad												
6490					Ad												
5200					Ad												
1330	0.00682	0.00316	2.05		Ad	ld a note	С	А	No tags t	to assign	42.014	0.935	1E+05	8.97E+05	21.5	10	
1439	0.0267	0.00888	2.1		la.		A	с	New tag	-	39.518	1.05	9.99E+04	4.53E+05	26.2	9	
4175	0.0169	0.00634	392		10.		С	A	Ouick Ta	as 🕨	35.643	0.595	9.98E+04	2.69E+05	173	2	
1244	0.00648	0.00304	2.42		ha.		Α	с	Glittaar	<u></u>	42.725	0.817	9.98E+04	5.23E+05	22.4	7	
4973	0.0536	0.015	126		19.		С	A L	Edit tags		54.032	0.831	9.98E+04	1.96E+05	87.7	0	
2085	0.0684	0.0181	3.29		12		С	Α	943.1200	3 2826.338	56.463	1.35	9.97E+04	4.53E+05	60.8	0	Ψ.
-																	4
	Delete	2333 sele	cted peptid	e ions		Due E											
1D Dis	play 2D F	Display 3	D Display			Run:	pression profile										
		inspiraly se	o onspiraly			Aggrega	te	8	04 80	5 806	807	808	(809.7327		612.5	825.0 1037.5 1250	0.0 1462.5 1675.0 188
							+ Da Edit	5		a			m/:	•			m/z ►
								825						2	Sec. Section	Ser. Sec.	
						spirt	Merge	<u> </u>			M n H n H	37				and the second	
						L-‡ Add	Delete		JAT Do			808.9486 m/z, 50	0.5752 min,	z=2	Wiger.	建钢油的方	
						🔊 Unde	o 🤇 🍋 Redo	55					. """	99	2	and the second second	
						Charge s	ate: 📑 1 📑 :	2		TH				-	E.	1.1.	
								1 9 U						80	in a		
								- uo						1 g -	uoi		
								65 stent	1			1 <u>11</u> 11		10	stent		
								* *							• *		
Experim	nent design	: AC		•												Sec	tion Complete $igleto$

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

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**.

Revie	Review Peak Picking								
Y	No filter ap	plied							
#	Anova (p)	q Value	Fold	Tag 💌	Note	s Highest Mea	n		
3496	0.162	0.0363	1.64		la.	С			
6490	0.024	0.0082	11.7		a.	с			
5200	0.0414	0.0124	148		a.	А			
1330	0.00682	0.00	Most abur	ndant	A	<u>dd a note</u> C			
1439	0.0267	0.00	New tag		-þ	А			
4175	0.0169	0.00	Ouick Tag	s		Anova n-value	ŀ.		
1244	0.00648	0.00	Edit			Max fold change	1		
4973	0.0536	0.01	Luit tags		-	Modification	L		
2085	0.0684	0.0181	3.29		1	No MS/MS data			
1381	3.15E-06	7.27E-06	7.84			No protein 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
-------------------	--------

Review	w Peak Pi No filter ap	cking plied		Crea	te			
#	Anova (p)	q Value	Fold	Tag 💌	Nc			
3496	0.162	0.0363	1.64	-	a *			
6490	0.024	0.0082	11.7	-	1			
5200	0.0414	0.0124	148	-				
1330	0.00682	0.00316	2.05	-	1			
1439	0.0267	0.00888	2.1		14			
4175	0.0169	0.00634	392		10.			
1244	0.00648	0.00304	2.42		10			
4973	0.0536	0.015	126		14			
2085	0.0684	0.0181	3.29		10.			
1381	3.15E-06	7.27E-06	7.84		·			
×	"				+			
Delete 2333 selected peptide ions								

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

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

Review Peak Picking								
7	No filter ap	Crea	ate					
#	Anova (p)	q Value	Fold	Tag 💌	Not			
1613	1.43E-10	1.59E-08	Infinity		·			
927	0.284	0.0583	1.25					
489	7.52E-05	9.24E-05	11.5		la.			
2050	0.0491	0.014	8.71					
1067	0.000316	0.000285	7.79		10			
1743	0.0144	0.00562	3.29		la.			
403	0.865	0.152	1.02		la.			
2529	3.94E-05	5.52E-05	155		la.			
426	0.0489	0.014	1.86		la			
469	0.0331	0.0105	1.48		ha. 🗸			
•					Þ			

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

QIP 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: Most abundant (2333 peptide ions) Anova p-value ≤ 0.05 (8174 peptide ions) Max fold change ≥ 2 (9822 peptide ions)	Show peptide ions that have all of these tags: 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			

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
387	2.63E-05	3.99E-05	21.4	-	🔜 Add a note	С
12372	0.0337	0.0106			🔜 Add a note	С
2946	0.00336	0.00181	5.4	-	🔜 Add a note	с
2953	0.00923	0.004	1.02E+03		🔜 Add a note	с
12376	0.0107	0.00449			🔜 Add a note	с
9122	0.0127	0.00512	354		10	А
14212	0.000258	0.000244	Infinity		No.	А
9120	0.0251	0.00849	86.6		14	А
14407	0.0266	0.00886	2.77		14	А
9307	0.0062	0.00293	6.75		14.	А

Create a tag for them called **Significantly up in C**, tagging 3946 peptide ions.

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

	QP Filter the peptide ions	
Tag filter applied peptide ions may be hidden	Create a filter Show or hide peptide ions based on a selection of create the filter. For more guidance, please see t	f their tags. Move tags to the appropriate boxes to he <u>online reference</u> .
Make sure that only the tag for the Most abundant peptide ions is shown and press OK.	Available tags: Anova p-value ≤ 0.05 (8174 peptide ions) Max fold change ≥ 2 (9822 peptide ions) Significantly up in C (3946 peptide ions)	Show peptide ions that have all of these tags: Most abundant (2333 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.	Clear the filter	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 86)

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 \leq 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 8174 peptide ions.

OP Filter the peptide ions Create a filter Show or hide peptide ions based on a selection to create the filter. For more guidance, please se	of their tags. Move tags to the appropriate boxes ee the <u>online reference</u> .
Available tags:	Show peptide ions that have all of these tags:
Max fold change ≥ 2 (9822 peptide ions) Significantly up in C (3946 peptide ions) Most abundant (2333 peptide ions)	Anova p-value ≤ 0.05 (8174 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

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	Cance	

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.

QIP Filter the peptide ions Create a filter Show or hide peptide ions based on a selection o create the filter. For more guidance, please see t	f their tags. Move tags to the appropriate boxes to he <u>online reference</u> .
Available tags:	Show peptide ions that have all of these tags:
Most abundant (2333 peptide ions) Max fold change ≥ 2 (9822 peptide ions) Up regulated in A (4228 peptide ions) Significantly up in C (3946 peptide ions) Up regulated in C (3946 peptide ions)	Anova p-value ≤ 0.05 (8174 peptide ions) Show peptide ions that have at least one of these tags: Hide peptide ions that have any of these tags:
<u>C</u> lear 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.



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 of 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 **37276** 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 37276 ms/ms spectra								
MSMS Preprocessing								
Limit fragment ion count								
Deisotoping and charge deconvolution								
Import search results								
Clear identifications 🔹								

File Beiner	QIP Progenesis QI.p Tutorial for DDA - Prog	enesis QI for proteo	omics										×
Honting top searches by exporting peak list free to protein distribution program. DS/DS Spectal	File Review Import Data Alignment Filtering	Experiment R Design Setup	Picking Sta	ide Ion Iden tistics Pept	tify des QC Metrics	Refine Identification	Review s Proteins	Protein Statistics Report			ŗ	online A Waters Comp	ar
75 90 0 10000 € less than Protein description contains 35 84 0 10 score ess than Protein description contains 37 74 0 10 <t< td=""><td>Identify Peptides Run ma/ms ion searches by exporting p to a protein identification program. Image: searches by exporting p Image: searches by exporting p</td><td>eak list files</td><td>MS/MS Spectra Batch inclusic Peptide i Cr Scan nu Exp</td><td>Rank greater th on ≠ less than harge less than mber less than orted equal to</td><td>ating export list of an v 5 v v</td><td>ms/ms spectra</td><td>Peptide ion Precursor Precursor F Peptide Protein</td><td>intensity less than intensity less than (%) less than (%) less than Run name contains sequence contains accession contains</td><td></td><td></td><td></td><td></td><td></td></t<>	Identify Peptides Run ma/ms ion searches by exporting p to a protein identification program. Image: searches by exporting p Image: searches by exporting p	eak list files	MS/MS Spectra Batch inclusic Peptide i Cr Scan nu Exp	Rank greater th on ≠ less than harge less than mber less than orted equal to	ating export list of an v 5 v v	ms/ms spectra	Peptide ion Precursor Precursor F Peptide Protein	intensity less than intensity less than (%) less than (%) less than Run name contains sequence contains accession contains					
44 70 0 10000 <td>75 90 0 47 88 0 35 84 0 27 80 0 8 78 0 4 75 0 73 74 0</td> <td>ſ</td> <td>ID : ID : ID :</td> <td>score less than</td> <td>n Scan number</td> <td>Exported Pe</td> <td>Protein de I sptide ion intensity 1 5 a + 005</td> <td>escription contains nclude in export</td> <td> clude from exp (%) Char 65.2 </td> <td>Clear all filter</td> <td>s Isotope</td> <td>Id score F</td> <td>P</td>	75 90 0 47 88 0 35 84 0 27 80 0 8 78 0 4 75 0 73 74 0	ſ	ID : ID : ID :	score less than	n Scan number	Exported Pe	Protein de I sptide ion intensity 1 5 a + 005	escription contains nclude in export	 clude from exp (%) Char 65.2 	Clear all filter	s Isotope	Id score F	P
34 65 0 54 64 0 54 64 0 65 0 64 0 7 7 7 7 8 0 8 0 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000 10000 1000	44 70 0 116 69 0 7 68 0 10 67 0			5 14444 A3 6 1 C3 6 2 A2	6324 7200 2890	No No No III	1.2e+005 1.2e+005 1.1e+008 7.7e+007	7.4e+003 7.4e+004 7.9e+007 6.6e+007	61.0 71.7 85.5	2 656.8617 3 805.4420 2 539.3012	1 1 1 1	•	· · · · · · · · · · · · · · · · · · ·
Import search results Section Complete 3	34 65 0 54 64 0 Clear the search program you're using: Mascot Help Export 18131 ms/ms spectra MSMS Preprocessing Limit fragment ion count 1000 Disstopping and charge deconvol Import search results Clear identifications •	tion	Peptide ion n Run:A1 Scan 29272- 001 31.399- 011 33.529-	humber 2, m number 2686	/z 539.3004, r 1 2 39 1 1 1 5 42	etention tim 5000 - 5000 - 5000 - 0 -	e 30.518 min, c	tharge +2	600 m/z	4 Lul L 800	1000 Section	Complete	*

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.

ſ	Export	ort Rank # Run Scan number		Exported	Peptide ion intensity	Precursor intensity	(%)	Charge	Precursor m/z	Isotope	Id s				
ſ	V	Ę	0.01	C2		NI-	E 7 00E	1 0 005	777	- 	<u> 604 0775</u>	1		*	
Ĩ	1	',	404	AD			z.je+000	3.4e+003	a matche 15.9	a to the sa	or cerositation (1997)	1			
		6	1800	A3	4618	No	2.9e+005	1.1e+005	39.4	2	783.8677	1			
ľ		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 **18131**

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 18131 spectra to export, as described above:

- 1. Select appropriate search engine i.e. Mascot
- 2. Click 'Export 18131 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 12A (page 98) 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 12A (page 98)**

Select the search program you're using:								
Mascot 👻								
Mascot								
PLGS (*.xml)								
SEQUEST (dta & out files)								
Phenyx								
Spectral Library Search								
Ion Accounting								
Deisotoping and charge deconvolution								
Import search results								
Clear identifications 👻								
Select the search program you're using:								
Select the search program you're using: Mascot								
Select the search program you're using: Mascot Help								
Select the search program you're using: Mascot Help Export 18131 ms/ms spectra								
Select the search program you're using: Mascot Help Export 18131 ms/ms spectra MSMS Preprocessing								
Select the search program you're using: Mascot Help Export 18131 ms/ms spectra MSMS Preprocessing Limit fragment ion count 40								
Select the search program you're using: Mascot Help Export 18131 ms/ms spectra MSMS Preprocessing Limit fragment ion count 40 Deisotoping and charge deconvolution								
Select the search program you're using: Mascot V Help Export 18131 ms/ms spectra MSMS Preprocessing Limit fragment ion count 40 Deisotoping and charge deconvolution Import search results								

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 Ion 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.



Metrics

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

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

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

Interpretation and use

Waters

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

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

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

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



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

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

Reporting

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

Report all metrics:

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

Report flagged metrics:

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

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

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

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 Beview Description Descr	QP Progenesis Q	I.p Tutorial for D	DA - Prog	jenesis QI for pro	teomics														
Review Description Description Description Packing Statics Percent Statics Percent 1 First back decision circles inference and objects 1 Score act of decision circles decise. 1 Score act of decision circles decise. 1 Score act of decision circles decise. Score act of decision circles decise. 1 Score act of decision circles decise. Score act of decise.	File																		
Paper Las Agence Preces Que Merce Description Report All Refine Identifications If your peptide identifications include unwanted or relevant results, you can ennow them hares. Secure as at of identifications you want to ident. Secure as at of identifications you want you want you want you want you you want you you you want you	Invest Date	Review	Citized e e	Experiment	Review P	eak	Peptide Ion	Ide	ntify	00.00	R	efine	Review	Protein	Derest.				nonlin
Ar With Section Sectin Sectin Section Section Section Section Section Secti	Import Data	Alignment	riitering	Design Setup	PICKIN	9	Statistics	Pep	ltides	QC Metrics	Ident	Cations	Proteins	Statistics	Report				DYNAM
Refine Identifications Secure Length (estimations) include unwanted or include unwanted unwa																	SC.	8 18	A Waters Co
if your peptide identifications include unwanted or intrevent result, you can ennow the harker. 1 Score lists than • 40 Sequence Length lists than • 1 1 Score lists than • - <td>Refine Iden</td> <td>tifications</td> <td></td> <td></td> <td>Bat</td> <td>ch del</td> <td>etion criteria</td> <td></td> <td>_</td>	Refine Iden	tifications			Bat	ch del	etion criteria												_
Image: Intermediation criteria in the property usant to detect. His less than • Charge less than • 2 Delete the unwanted identifications you want to detect. Accession contains • 2 Delete the unwanted identifications you don't wart, click. etter: • Description contains • • Delete Matching Search Result, to delete the identifications you don't wart, click. etter: • Description contains • • Delete Matching Search Result, to delete the invalue for a start again To specify search Result, to delete the IDs that are not highlighted • Retention Time less than • • 7 Description contains • Description contains • • Description contains • 7 Search Result, to delete the IDs that are not highlighted 159 137.53 745.91 51.25 2 148.98 0.66 FAELLGEV • 0 0 0 0 0 FAELLGEV • 0 <td>If your peptide irrelevant resu</td> <td>identifications i Its, you can rem</td> <td>include u nove ther</td> <td>nwanted or n here.</td> <td></td> <td></td> <td></td> <td>Score</td> <td>less than</td> <td>•</td> <td>40</td> <td></td> <td></td> <td></td> <td>Sequence Lengt</td> <td>h less than</td> <td>•</td> <td></td> <td></td>	If your peptide irrelevant resu	identifications i Its, you can rem	include u nove ther	nwanted or n here.				Score	less than	•	40				Sequence Lengt	h less than	•		
In the backh deletion criteria, enter the property values for a set of identifications you want to delete, i.e. was to identifications you want to the the identifications you want to the interval identifications is a set of identification is interval	Specify a	set of deletio	n criteri	a				Hits	less than	•					Charg	e less than	•		
 Absolute mass error (ppm) less than v Accession locatians v Modifications contains v Modifications contains v Modifications contains v Better the IDs that are not highlighted IDs Sarch Results, to delete the IDs that are not highlighted IDs Sarch Results, to delete the IDs that are not highlighted IDs Sore Hits m/z RT(mins) Charge Mass Mass error Sequence Accession Modifications to delete forterine soft the not matching search results Sore Hits m/z RT(mins) Charge Mass Mass error Sequence Accession Modifications Delete matching search results Delete matching search results Delete matching search results Delete matching search results Total Hits m/z RT(mins) Charge Mass Mass error Sequence Accession Modifications Delete matching search results Total Hits m/z RT (min Charge Tag 10 0004410 5559 3 12 05 5983214 1149 2 13 05 50593 1 12 053304 3052 2 13 03 05 82569 4575 4 32987; 071 9 EUTPASA 9 glp25205 14 05 064337 2734 2 16 12 3930 5 82569 4575 4 32987; 071 9 EUTPASA 9 glp25205 12 05 13303 1 12 05 05303 2 12 01 12 05 1303 2 12 01 12 05 100 2 13 05 66868 4460 2 14 00 12 0 12 5 11700 15873 2 13 05 66868 4460 2 14 00 12 0 12 5 11700 15873 2 13 05 66868 4460 2 14 00 12 0 12 5 11700 15873 2 13 05 66868 4460 2 14 05 073320 2 4239 2 14 05 073320 2 4239 2 15 06433 5 45227 4222 2 10 00 05 9 glp2732954 5-layer protein partial [[Costridium diffice 10] 15 06433 140 2 16 05 0733 20 23311 032 10 00 05 9 g	In the bat	ch deletion crite	eria, ente	r the property				Mass	less than	•					Sequenc	e contains	•		
2 Delete the unwanted identifications of deter the identifications you don't want, click either: Image: the identifications you don't want, click either identifications you don't want,	delete.	a bec of identifi	, ,	ou mane to	Abs	solute	mass error	(ppm)	less than	•					Accessio	n contains	•		
Z To delete the identifications you don't wart, click either: • Delete Matching Search Results, to delete the highlighed IDs • Delete Non-matching Search Results, to delete the IDs that are not highlighted • Delete Non-matching Search Results, to delete this Us that are not highlighted • Delete Non-matching Search Results, to delete this Us that are not highlighted 3 Reset the criteria to start again rospecify another batch of identifications to step 1 above. • Total Hits m/z RT (min Charge Tag Mass Source FAELLGEW. • gill266991 • Unbergythin [Clostridium difficle 530] • Total Hits m/z RT (min Charge Tag 10 80.35 745.91 51.25 2 1489.80 0.06 FAELLGEW. • gill266991 • Unbergythin [Clostridium difficle 530] • Total Hits m/z RT (min Charge Tag 10 80.35 745.91 51.25 2 1489.80 0.06 FAELLGEW. • gill266991 • Unbergythin [Clostridium difficle 530] • Total Hits m/z RT (min Charge Tag 10 80.35 745.91 51.25 2 1489.80 0.06 FAELLGEW. • gill25491 Rubergythin [Clostridium difficle 530] • Total Hits m/z RT (min Charge Tag 2 3) Delete th	e unwanted id	dentifica	tions				m/z	less than	•					Descriptio	n contains	•		
Poleite Matching Search Results, to delete the inplifyhed IDs Intermediation Center of the criteria to start again To specify another batch of identifications to delete, cick Reset the Criteria and then return to step 1 above. # Score Hits m/z RT(mind) Charge Mass err Sequence Accession Modifications Description Image: Construction of the poly of the construction of the constructin the construction of the construction of the construction of the	To delete either:	the identification	ns you de	on't want, click			Detection		1 11							-			
Implified US Delete matching search results Delete matching search results Delete mon-matching search results	Delet	e Matching Sear	rch Resul	ts, to delete the			Recention	i nine	less than	•					Modification	contains	•		
delete the IDs that are not highlighted Bestet the criteria to start again To specify another batch of identifications to delete, click Reset the Criteria and then return to step 1 above. Score Hits m/z RT(mink) Charge Mass Mass Mass Mass Sequence Accession Modifications Description Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then return to step 1 above. Image: Click Reset the Criteria and then re	• Delet	ignted IDs e Non-matching	Search	Results, to										📃 Delete	matching search r	esults 🗌 Del	lete non-n	natching search resu	ults Reset the crite
Reset the criteria to start again Vic Nord Number of the criteria to start again Note of the criteria to start again <td>delet</td> <td>e the IDs that a</td> <td>re not hi</td> <td>ghlighted</td> <td></td> <td></td> <td>Correct</td> <td>1.044</td> <td></td> <td>DT/mins</td> <td></td> <td>Maria</td> <td>Marria</td> <td></td> <td>A</td> <td>Madifferentia</td> <td></td> <td></td> <td>Description</td>	delet	e the IDs that a	re not hi	ghlighted			Correct	1.044		DT/mins		Maria	Marria		A	Madifferentia			Description
To specify another batch of identifications to delete, dick Reset the Criteria and then return to step 1 above. 1 199 13753 5 74591 5125 2 1489.8 (0.6 FAELIGEV 9 gij208732 rubrerythrin [Costridium cellulovaras 7 rubrerythrin [Costridi	2 Reset the	e criteria to sta	art agaiı	n		150	127.52	HITS	7/15 Q1	51.25) Charge	1/180.80	Mass err	Sequence	Accession	Modifications	rubrandi	arin [Clostridium dif	Description
above. istep 1 above	To specify delete cli	another batch	of identif	ications to		150	127.52	5	745.91	51.25	2	1/180.8(0.06		gij2200331		Rubrend	hrin [Deculfitobacte	rium dichloroalimina
No filter applied Create 157 3 74591 51.25 2 1489.8 0.06 FAELLGEV. 9 gip25820 putative mberghini [Clostridium diffici # Total Hits m/z RT (min Charge Tag 1 10 805.4410 55.59 3 6 22.669 45.75 4 3298.7: 0.71 6 ELITPASA 9 gip2589056 s-layer protein, partial [[Clostridium gp. Lv/III] 2 60 5393.004 305.2 2 439.85 43298.7: 0.71 6 ELITPASA 9 gip2580178 cell surface protein Apartial [[Clostridium gp. Lv/III] 3 55 598.3214 31.49 2 6 161 39.30 5 825.69 45.75 4 3298.7: 0.71 ELITPASA 9 gip2520956 -s-layer protein, partial [[Clostridium] 6 3 55 598.3214 31.49 2 6 10 16.2 91.25 5 1170.00 58.73 2 238.11 0.32 LDNLG	step 1 abo	ove.		i then return to		159	137.53	5	745.91	51.25	2	1489.80	0.06	S FAFLIGEVV	gij3743023		rubrendi	nin (Clostridium cel	lulovorans 743Bl
Vertice applied Create Vertice Create Vertice Partice Partic Partice Partice						159	137.53	5	745.91	51.25	2	1489.80	0.06		gij3020752		nutative	ruben/thrin [Clostric	dium difficile OCD-33
Image Total Hits m/z RT (min Charge Tag) Image 10 205.410 55.59 3 6 25.69 45.75 4 3298.7: 0.71 © ELITPASA © gij87239956 s-layer protein, partial [[Clostridium] diff 2 60 5393.300 305.2 2 6 39.30 5 825.69 45.75 4 3298.7: 0.71 © ELITPASA © gij87239956 s-layer protein, partial [[Clostridium] diff 3 55 598.3214 31.49 2 161 39.30 5 825.69 45.75 4 3298.7: 0.71 © ELITPASA © gij87239954 s-layer protein, partial [[Clostridium] diff 4 10 624.2913 28.46 2 162 91.25 5 1170.07 58.73 2 2338.11 0.32 © DNLEDG © gij87239954 s-layer protein, partial [[Clostridium] diff [Clostridium] diff 162 91.25 5 1170.07 58.73 2 2338.11 0.32 DNLEDG © gij87239952 s-layer protein, partial [[Clostridium] diff [Clostridium] diff 162 91.25	No filte	er applied		Create		159	137.53	5	745.91	51.25	2	1489.80	0.06	S FAFLI GEVV	gij3503200		Rubrervt	hrin [Clostridium sp	DI -VIII
# Total Hits m/z RT (min Charge Tag Image: Total Hits m/z RT (min Charge Tag 1 10 8540 5559 3 1 912549763 cell surface protein (5-layer precursor protein (5-layer precursor protein (5-layer						161	39.30	5	825.69	45.75	4	3298.73	0.71	S ELIITPASA	ail87239956		s-laver p	rotein, partial [[Clos	tridium] difficile]
1 10 805.4410 55.59 3 2 2 60 539.304 30.52 2 2 6 539.304 30.52 2 2 6 393.04 30.52 2 2 6 393.04 30.52 2 2 6 393.04 30.52 2 2 3 5 598.3214 31.49 2 2 38.75 4 3298.77 0.71 9 EUTPASA 9 9/3230669 s-layer protein, A partial [[Clostridium] off 3 55 598.3214 31.49 2 2 2338.11 0.32 2 DNLGDG 9 9/j87239954 s-layer protein, A partial [[Clostridium] of 6 5 35 6043377 27.34 2 2 100.07 58.73 2 2338.11 0.32 DNLGDG 9 9/j87239952 s-layer protein (S-layer protein (S-layer protein (S-layer protein partial [[Clostridium] of 6 7 10 462.2705 35.00 2 2 1170.07 58.73 2 2338.11 0.32 DNLGDG 9 9/	# Total H	lits m/z	RT (min	Charge Tag		161	39.30	5	825.69	45.75	4	3298.73	0.71	S ELIITPASA	gi 2549763		cell surfa	ce protein (S-layer)	precursor protein) [C
2 60 539.3004 3052 2 2 161 39.30 5 823.69 45.75 4 3298.75 0.71 © ELITPASA © gij2380869 s-layer protein, partial [[Clostridium] diff 3 55 598.3214 31.49 2 2 1100.07 58.73 2 2338.11 0.32 © DNUEDGC © gij2739954 s-layer protein, partial [[Clostridium] diff 4 10 624.2913 24.46 2 2 2338.11 0.32 © DNUEDGC © gij2739954 s-layer protein, partial [[Clostridium] diff 6 15 624.317 28.46 2 2 2338.11 0.32 © DNUEDGC © gij255.019 cell surface protein (clostridium] diff 7 10 462.2705 35.00 2 2 1170.07 58.73 2 2338.11 0.32 © DNUEDGC © gij255.019 cell surface protein (clostridium] diff 7 10 462.2705 35.00 2 2 2338.11 0.32 © DNUEDG © gij255.019 cell surface protein, partial [[Clostridium] diff 9 30	1 10	805.4410	55.59	3 📮		161	39.30	5	825.69	45.75	4	3298.73	0.71	S ELIITPASA	gi 2067250		S-layer p	rotein A, partial [[Cl	ostridium] difficile]
3 55 598.3214 31.49 2 2 1170.00 58.73 2 2338.1(0.32 © LDNLGDG © gij87239954 s-layer protein, partial [[Clostridium difficie] 5 5 604.3377 27.34 2 2 1170.00 58.73 2 2338.1(0.32 © LDNLGDG © gij87239954 s-layer protein, partial [[Clostridium difficie] 5 35 604.3377 27.34 2 2 2 2338.1(0.32 © LDNLGDG © gij87239954 s-layer protein, [Clostridium difficie] 6 15 626.3146 35.6 2 162 91.25 5 1170.0(58.73 2 2338.1(0.32 © LDNLGDG © gij87239954 s-layer protein, [Clostridium difficie] 7 10 462.2705 35.00 2 If 2 91.25 5 1170.0(58.73 2 2338.1(0.32 © LDNLGDG © gij87239954 s-layer protein, partial [[Clostridium difficie] 8 12 753.8284 31.44 2 2 91.25 5 1170.0(58.73 2	2 60	539.3004	30.52	2		161	39.30	5	825.69	45.75	4	3298.73	0.71	😒 ELIITPASA	gi 92380869		s-layer p	rotein, partial [[Clos	tridium] difficile]
4 10 6242913 2846 2 2 9125 5 1170.0f 58.73 2 2338.1f 0.32 © LDNLGDG © gig12102505 S-layer protein [Clostridium difficie] 5 35 604.3377 27.34 2 2 912702505 S-layer protein [Clostridium difficie] 6 15 626.3146 39.56 2 2 1170.0f 58.73 2 2338.1f 0.32 © LDNLGDG © gij251019 cell surface protein [Clostridium difficie] 7 10 4622705 55.00 2 2 12 5 1170.0f 58.73 2 2338.1f 0.32 © LDNLGDG © gij7132944 S-layer protein [Clostridium difficie] 8 12 753.8284 31.44 2 If 162 91.25 5 1170.0f 58.73 2 2338.1f 0.32 © LDNLGDG © gij71732944 S-layer protein, partial [[Clostridium difficie] 9 30 656.8613 44.60 2 If 162 91.25 5 1170.0f 58.73 2 2338.1f 0.32	3 55	598.3214	31.49	2		162	91.25	5	1170.0	6 58.73	2	2338.1(0.32	LDNLGDG	🌒 gi 87239954		s-layer p	rotein, partial [[Clos	tridium] difficile]
5 35 6043377 27.34 2 2 162 91.25 5 1170.07 58.73 2 2338.11 0.32 © LDNLGDG © gij225952 s-layer protein (S-layer precursor pr 6 15 6263.146 95.56 2 II 162 91.25 5 1170.07 58.73 2 2338.11 0.32 © LDNLGDG © gij12239952 s-layer protein partial [[Clostridium] diff 7 10 462.2705 35.00 2 II 162 91.25 5 1170.07 58.73 2 2338.11 0.32 © LDNLGDG © gij1732944 S-layer protein partial [[Clostridium] diff 9 30 656.6613 44.60 2 II 162 91.25 5 1170.07 58.73 2 2338.11 0.32 © LDNLGDG © gij1463259 S-layer protein partial [[Clostridium] diff 16 9 30 656.8613 44.60 2 II 165 64.35 5 452.27 42.22 902.53 -0.54 © AFIGLINR © gij1266991 ruberythrin (Dostifidium difficide 30.10)	4 10	624.2913	28.46	2		162	91.25	5	1170.0	6 58.73	2	2338.1(0.32	🕥 LDNLGDG	🎯 gi 21702505		S-layer p	rotein [Clostridium	difficile]
6 15 6263146 3956 2 Image: Constraint of the const	5 35	604.3377	27.34	2		162	91.25	5	1170.0	6 58.73	2	2338.1(0.32	🔇 LDNLGDG	🎯 gi 2551019		cell surfa	ice protein (S-layer j	precursor protein) [C
7 10 462.2705 35.00 2 Image: Constraint of the con	6 15	626.3146	39.56	2		162	91.25	5	1170.0	6 58.73	2	2338.1(0.32	🔇 LDNLGDG	🎯 gi 87239952		s-layer p	rotein, partial [[Clos	tridium] difficile]
8 12 753.828 31.44 2 Image: Constraint of the cons	7 10	462.2705	35.00	2		162	91.25	5	1170.0	6 58.73	2	2338.1(0.32	🔮 LDNLGDG	🎯 gi 71732944		S-layer p	rotein [Clostridium	difficile]
9 30 656.8613 44.60 2 Image: Constraint of the con	8 12	753.8284	31.44	2		162	91.25	5	1170.0	6 58.73	2	2338.1(0.32	🔮 LDNLGDG	🎯 gi 1463259		S-layer p	rotein, partial [[Clos	tridium] difficile]
10 18 663.8693 47.56 2 Image: Constraint of the image: Constraint	9 30	656.8613	44.60	2	1	165	64.35	5	452.27	42.22	2	902.53	-0.54	AFLGLLNR	🎯 gi 3743823		Rubreryt	hrin (Desulfitobacte	rium dichloroelimina
11 20 595.3191 37.47 2 2 165 64.35 5 452.27 42.22 2 902.53 -0.54 The second	10 18	663.8693	47.56	2 🧯		165	64.35	5	452.27	42.22	2	902.53	-0.54	AFLGLLNR	🎯 gi 1266991		rubreryt	nrin (Clostridium dif	ficile 630]
12 18 573.8033 25.39 2 Image: Constraint of the state o	11 20	595.3191	37.47	2 🧃		165	64.35	5	452.27	42.22	2	902.53	-0.54	AFLGLLNR	🎯 gi 3545597		Rubreryt	hrin [Desulfitobacte	rium metallireducen
14 60 573.3242 42.39 2 🕡 🗹 165 64.35 5 452.27 42.22 2 902.53 -0.54 🔮 AFLGLURR 🍨 gij2196705 rubrerythrin [Desulfitobacterium hafnier	12 18	573.8033	25.39	2		165	64.35	5	452.27	42.22	2	902.53	-0.54	AFLGLLNR	🎯 gi 1266990		rubreryt	nrin (Rr) [Clostridium	n difficile 630]
	14 60	573.3242	42.39	2		165	64.35	5	452.27	42.22	2	902.53	-0.54	AFLGLLNR	🎯 gi 2196705		rubreryt	nrin (Desulfitobacter	rium hafniense DCB-
	*			•	4	4.00	64.5F	-	100.07	42.00	-	000.50							
Resolve conflicts >> 2903 search results. 512 matching batch delete options. Section Co		Resolve conf	flicts >>		2903	searc	h results. 512	? matchi	ng batch d	lelete opti	ons.							S	ection Complet

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 (512 matching out of 2903)

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

Click Yes.

QIP Delete 512 search results?	23
Are you sure you want to permanently delete 512 pepti	ide search results?
Yes	<u>N</u> o

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.

OF Progenesis OI n Tutorial for DDA - Progenesis OI for proteor	mics		
File			
Review Experiment Re	eview Peak Peptide Ion Identify	Refine Review Protein	nonlinoar
Import Data Alignment Filtering Design Setup	Picking Statistics Peptides QC Metrics I	entifications Proteins Statistics Report	IIOIIIIIeai
			A Waters Company
Refine Identifications	Batch deletion criteria		
If your peptide identifications include unwanted or irrelevant results, you can remove them here.	Score less than 🔹	Sequence Lengt	h less than 🔻
Specify a set of deletion criteria	Hits less than 🔻	Charg	e less than 🔻
In the batch deletion criteria, enter the property values for a set of identifications you want to	Mass less than 💌	Sequenc	e contains 🔹
delete.	Absolute mass error (ppm) less than	Accessio	n contains 👻
2 Delete the unwanted identifications	m/z less than	Descriptio	contains
either:	Retention Time less than	Modification	s contains
 Delete Matching Search Results, to delete the highlighted IDs 	less then		
Delete Non-matching Search Results, to		Delete matching search n	esults Delete non-matching search results Reset the criteria
delete the tos that are not highlighted	# So DIP Delete 714 search results?	sz ession	Modifications Description *
Reset the criteria to start again	✓ 1 75	549763	cell surface protein (S-layer precursor protein) [Clos
delete, click Reset the Criteria and then return to	I 75 Are you sure you want to	permanently delete 714 peptide search results? 2551019	cell surface protein (S-layer precursor protein) [Clos
step 1 above.	2 59	2380869	s-layer protein, partial [[Clostridium] difficile]
Sea No filter applied	2 59	2551019	cell surface protein (S-layer precursor protein) [Clos
Create	2 59	463259	S-layer protein, partial [[Clostridium] difficile]
	2 59	Yes No 2549763	cell surface protein (S-layer precursor protein) [Clos
# Total Hits m/z RT (min Charge Tag	2 59	067250	S-layer protein A, partial [[Clostridium] difficile]
1 10 805.4410 55.59 3	2 59.57 5 539.30 30.52 2	1076.55 -0.36 🎯 LLFTQVDNK 🚳 gi[21702505	S-layer protein [Clostridium difficile]
2 60 539.3004 30.52 2	2 59.57 5 539.30 30.52 2	1076.5: -0.36 🔇 LLFTQVDNK 🔇 gi 87239956	s-layer protein, partial [[Clostridium] difficile]
3 55 598.3214 31.49 2	2 59.57 5 539.30 30.52 2	1076.5! -0.36 🔇 LLFTQVDNK 🔇 gil2067250	S-layer protein A, partial [[Clostridium] difficile]
4 10 624.2913 28.46 2	2 59.57 5 539.30 30.52 2	1076.5: -0.36 🔇 LLFTQVDNK 🔇 gi 87239954	s-layer protein, partial [[Clostridium] difficile]
5 35 604.3377 27.34 2	2 59.57 5 539.30 30.52 2	1076.5! -0.36 🔇 LLFTQVDNK 🔇 gi[1121811	s-layer protein, partial [[Clostridium] difficile]
6 15 626.3146 39.56 2	2 59.57 5 539.30 30.52 2	1076.5? -0.36 🎯 LLFTQVDNK 🎯 gi[71732944	S-layer protein [Clostridium difficile]
7 10 462.2705 35.00 2	2 59.57 5 539.30 30.52 2	1076.5: -0.36 🕥 LLFTQVDNK 🕥 gi 87239952	s-layer protein, partial [[Clostridium] difficile]
8 12 753.8284 31.44 2	☑ 3 51.72 5 598.32 31.49 2	1194.6: -0.19 🔇 YQVVLYPE 🔇 gi 1463259	S-layer protein, partial [[Clostridium] difficile]
9 30 656.8613 44.60 2	✓ 3 51.72 5 598.32 31.49 2	1194.6: -0.19 🔞 YQVVLYPE 🔇 gi 2067250	S-layer protein A, partial [[Clostridium] difficile]
10 18 663.8693 47.56 2 🧃	☑ 3 51.72 5 598.32 31.49 2	1194.6: -0.19 🔞 YQVVLYPE 🔇 gi 87239954	s-layer protein, partial [[Clostridium] difficile]
11 20 595.3191 37.47 2 🧃	☑ 3 51.72 5 598.32 31.49 2	1194.6: -0.19 🔞 YQVVLYPE 🚳 gi 21702505	S-layer protein [Clostridium difficile]
12 18 573.8033 25.39 2	☑ 3 51.72 5 598.32 31.49 2	1194.6: -0.19 🔞 YQVVLYPE 🔇 gi 92380869	s-layer protein, partial [[Clostridium] difficile]
14 60 573.3242 42.39 2	3 51.72 5 598.32 31.49 2	1194.6: -0.19 🔇 YQVVLYPE 🔇 gi 2067250	S-layer protein A, partial [[Clostridium] difficile]
۰ الله الله الله الله الله الله الله الل			
Resolve conflicts >>	2065 search results. 714 matching batch delete options.		Section Complete $ ightarrow$

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

Progenesis QI.p Tutorial for DDA - Progenesis QI for proteom	ics	
File		
Review Experiment Rev Import Data Alignment Filtering Design Setup	ew Peak Peptide Ion Identify Refine Review Protein Icking Statistics Peptides OC Metrics Identifications Proteins Statistics	Report nonlinear
		A Waters Company
Poline Identifications	Datab delation estado	
If your peotide identifications include unwanted or	Score lars than	Sequence Length less than
irrelevant results, you can remove them here.		
Specify a set of deletion criteria	Hits less than 🔻	Charge less than •
values for a set of identifications you want to	Mass less than 🔻	Sequence contains
delete.	Absolute mass error (ppm) less than 🔻	Accession contains
2 Delete the unwanted identifications To delete the identifications you don't want, click	m/z less than 🔻	Description doesn't contain 🔹 clostridium difficile
either:	Retention Time less than	Modifications contains 💌
 Delete Matching Search Results, to delete the highlighted IDs 		
 Delete Non-matching Search Results, to delete the IDs that are not highlighted 		ete matching search results
 Devet the withwise to struct a using 	# Sd QIP Delete 177 search results?	ession Modifications Description
3 To specify another batch of identifications to	152 90	2556544 electron transfer flavoprotein beta-subunit [Clostrid
delete, click Reset the Criteria and then return to step 1 above.	Iss 12 Are you sure you want to permanently delete 177 peptide search	results? 1095712 cell wall protein V [[Clostridium] difficile]
	▼ 158 12	266978 flagellin C. [Clostridium difficile 630]
Create	✓ 158 12	i668937 flagellin [[Clostridium] difficile]
* Tatalilite m/s DT (size Channel Tatal	▼ 158 12 <u>Y</u> es	No 2549739 flagellin subunit [Clostridium difficile QCD-66c26]
Iotal Hits M/2 Ki (min Charge Tag	✓ 158 12	.0281485 flagellin subunit FliC [[Clostridium] difficile]
2 60 5393004 30.52 2	☑ 158 124.65 5 1159.5€ 39.52 2 2317.1: 0.21	🦉 gi[2618637 flagellin [[Clostridium] difficile]
3 55 598 3214 31.49 2	▼ 158 124.65 5 1159.5€ 39.52 2 2317.11 0.21 ♥ LESTQNNL	
4 10 624,2913 28,46 2	▼ 159 137.53 5 745.91 51.25 2 1489.81 0.06 ▼ FAELLGEV	rubrerythrin [Clostridium sp. JC122]
5 35 604.3377 27.34 2	▼ 159 137.53 5 745.91 51.25 2 1489.8(0.06 ● FAELGEV	/ in intervention (Costriction and the obj
6 15 626.3146 39.56 2	▼ 159 137.53 5 745.91 51.25 2 1489.8(0.06 S FAELLGEV	/ gil2196705 rubrervthrin [Desulfitobacterium hafniense DCB-2]
7 10 462.2705 35.00 2	✓ 159 137.53 5 745.91 51.25 2 1489.8(0.06 S FAELLGEV	/ 🕥 gi[1500162 rubrerythrin [Clostridium beijerinckii NCIMB 8052]
8 12 753.8284 31.44 2	✓ 159 137.53 5 745.91 51.25 2 1489.8(0.06 FAELLGEV	/ 🌒 gi[1266991 rubrerythrin [Clostridium difficile 630]
9 30 656.8613 44.60 2 📲	☑ 159 137.53 5 745.91 51.25 2 1489.8(0.06 S FAELLGEV)	/ 🎱 gi]3743823 Rubrerythrin [Desulfitobacterium dichloroeliminans
10 18 663.8693 47.56 2	☑ 159 137.53 5 745.91 51.25 2 1489.8(0.06	/ 🎯 gi[3028732 rubrerythrin [Clostridium cellulovorans 743B]
11 20 595.3191 37.47 2	☑ 159 137.53 5 745.91 51.25 2 1489.8(0.06	/ 🕥 gi]3594118 Rubrerythrin [Clostridium sp. DL-VIII]
12 18 573.8033 25.39 2	✓ 162 91.25 5 1170.0(58.73 2 2338.1(0.32) LDNLGDG	🕥 gi 21702505 S-layer protein [Clostridium difficile]
14 60 573.3242 42.39 2	☑ 162 91.25 5 1170.0€ 58.73 2 2338.1€ 0.32 S LDNLGDG	🌒 gi[2551019 cell surface protein (S-layer precursor protein) [Clos
< III +		
Resolve conflicts >>	1351 search results. 177 matching batch delete options.	Section Complete $ e e e e e e e e e e e e e $

Having applied all the filters there will be 1174 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 87)

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
Use protein grouping of another proteins.
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:	AC			•													
Proteins 🕎 🕨	o filter applied		C	reate	Pep	tide i	ions	of g	gi 25497	73900			Vo No	filter applied		Cre	eate
Accession	Peptides	Unique	Conflict	Score *		#	Σ	몓	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	•
🎯 gi 260682215 (+1) 24	11	18	2.081	V	67	٠	2	63.9	5	0.965	1669.889	-0.0445	42.9	3		
🎯 gi 126700407	16	16	0	1.641	1	121	٠	1	103	5	0.963	1669.888	-0.562	42.9	2		
🎯 gi 254973900 (+9) o 15	15	0	1.95	V	148	٠	2	101	5	0.921	1230.609	-0.423	24	2		
🎯 gi 126698450 (+1	gi 2549739	00 - flage	llin subunit	[Clostridi	um diffi	cile QC	D-66	c26]	125	5	0.955	2317.115	0.205	39.5	2		
gi 126697969 (+1) 10	10	0	92 -	V	172	•	2	60.9	3	0.951	2317.115	0.2	39.5	3		-
Drotoin: gil2E4	072000 8-	a allin a	ubunit ff	, Nastridi	um dif	Ficilo	000		-261	,							r
No protein col	antad	genin s	ubunit (v	Josuidi	um an	licite	QUL	J-00	czoj								
Pentide ion views	rotein resolutio	n															
Conflicting pro	toins				Donti	de ior		f.cor	oflicting	protei	n						
connecting pro					repu				nuccing	protei					a a		
Accession P	eptides Uniqu	e Conf	lict: Protei	n Score	#	2	4	9 50	core	Hits C	orrelation N	lass IV	lass error (p F	(1 (mins)	Charge I	ag 🔻	Abi
												_					
•																	
E Refine Ident	fications	Protein	options											Se	ction Cor	nplete	۲

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.

Experiment design:	AC			•															
Proteins	- No	filter appl	ied			Pep	otide i	ons	of g	gi 26068	2017				ſ	🥌 No fi	lter appl	lied	
	Υ.			Cn	eate											Y			Create
Accession	Peptides	Unique	Conflict:	Score	Tag ^		#	Σ	Ð	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	✓ Abundance	Confl ^
🎯 gi 260682017 🛛 🤇) 15	0	127	1.95E+03		V	452	0	0	93.3	5	0.961	1676.838	-1.23	35.5	2		7.45E+05	7
🌚 gi 254973900	15	0	127	1.95E+03		1	1358	0	0	99	4	0.959	1407.656	0.441	39.2	2		1.27E+05	7 ⊨
🎯 gi 5668937	14	0	123	1.9E+03		V	875	0	0	109	5	0.966	1692.833	-1.01	22	2		4.62E+05	7
🌚 gi 261863741	14	0	121	1.76E+03		1	465	0	1	47.4	5	0.976	1692.835	-0.251	21.9	3		5.57E+05	7
🌍 gi 10281485	10	0	94	1.27E+03		V	3339	0	1	84.7	4	0.869	1423.65	-0.418	23.7	2		7.64E+04	7
🎯 gi 126697810	10	0	91	1.19E+03		1	291	0	0	43.8	4	0.961	1676.838	-1.39	35.5	3		7.6E+05	7
🎯 gi 10281487	8	0	75	965		V	158	0	0	125	5	0.955	2317.115	0.205	39.5	2		5.55E+06	6
🎯 gi 21702505	26	0	69	2.5E+03		V	2758	0	0	77.6	5	0.864	1386.71	0.102	19.1	2		5.12E+04	6
🎯 gi 73745726	5	0	54	662		V	172	0	2	60.9	3	0.951	2317.115	0.2	39.5	3		3.09E+06	6
🎯 gi 260682215	24	1	32	2.08E+03	-	V	148	0	2	101	5	0.921	1230.609	-0.423	24	2		3.07E+06	6 -
•	111					•													•
1 Protein: gi 26	0682017	flagell	in subur	nit [Clostrid	lium diffi	cile (CD196	1											
Protein: gi 25	4973900	flagell	in subur	nit [Clostrid	lium diffi	cile (QCD-6	6c2	6]										
Peptide ion views	Protein reso	olution																	
Conflicting pro	oteins for	peptide	e ion 45	2		Pe	ptide	ions	of	gi 2549	73900								
Accession	Peptide	s Unique	Conflic	t: Protein Sco	ore		#	Σ	₽	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	✓ Abundance	Cor ^
🎯 gi 254973900	15	0	127	1.95E+03	v	V	452	0	0	93.3	5	0.961	1676.838	-1.23	35.5	2		7.45E+05	5
🎯 gi 260682017	o 15	0	127	1.95E+03	V	V	1358	0	0	99	4	0.959	1407.656	0.441	39.2	2		1.27E+05	7 ≡
gi 5668937	14	0	123	1.9E+03	V	1	875	0	0	109	5	0.966	1692.833	-1.01	22	2		4.62E+05	5
🎯 gi 261863741	14	0	121	1.76E+03	V	V	465	0	1	47.4	5	0.976	1692.835	-0.251	21.9	3		5.57E+05	7
🎯 gi 10281485	10	0	94	1.27E+03	V	V	3339	0	1	84.7	4	0.869	1423.65	-0.418	23.7	2		7.64E+04	5
🎯 gi 126697810	10	0	91	1.19E+03	V	V	291	0	0	43.8	4	0.961	1676.838	-1.39	35.5	3		7.6E+05	5
🎯 gi 10281487	8	0	75	965		V	158	0	0	125	5	0.955	2317.115	0.205	39.5	2		5.55E+06	E
🎯 gi 73745726	5	0	54	662	V	V	2758	0	0	77.6	5	0.864	1386.71	0.102	19.1	2		5.12E+04	e
						V	172	0	2	60.9	3	0.951	2317.115	0.2	39.5	3		3.09E+06	e
•						-		~	-		-		4000 000		~ *	-	-	2.075.05	*
																	ſ		
C Refine Iden	tification	S Prot	tein optior	15														Section Comple	te 🤿

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.

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

Now set the Protein Options back to Use protein grouping.

×

Protein Quantitation options

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

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

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

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

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

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

P Protein quantitation options							
Quantitation method:							
Relative Quantitation using Hi-N							
Absolute Quantitation using Hi-N Requires a calibrant protein to calculate absolute amou Uses mean calibrant abundance measured across runs Uses the most abundant N peptides Allows comparison between proteins within a run 	ınts						
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 ic • Allows comparison of a single protein across runs	dentifications						
Relative Quantitation using all peptides • Uses all peptides identified as part of a protein • Allows comparison of a single protein across runs							
Absolute Quantitation for HCP using Hi-N • Requires a calibrant protein to calculate absolute amou • Uses calibrant abundance measured per run • Uses the most abundant N peptides • Calculates the amount of HCPs per run in fmol and ng	ınts						

QP 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 (i.e. for enolase) and amount for the calibrant.

		UP 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): 3	N	Calibrant accession:	gi 1267007	
Calibrant accession:	\square	▲ The calibrant protein can't be found. How will measurements be calculated?	gi 126700790 enolase (2-phosphogly	cerate dehvdratase) (2-phospho-D-g
Amount (fmol):		Amount (fmol):	gi 126700794	
Use protein grouping i.e. hide proteins whose pentides are a subset		Anodate (Intoly)	glyceraldehyde-3-phos	phate dehydrogenase [Clostridium d
of another protein's.		Employ protein grouping, i.e. hide proteins whose subset of another protein's.	e peptides are a	
OK Cancel		0	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.		W No filter a	pplied Creat	te	h P		@ H	lelp 🔻
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 quantitation		8.6	Α	С	transketolase, central and C-terminal (Sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycolaldehy	20.1	2.34	
use of Hi-N, protein grouping and more.		3.54 C A fructose-1,6-bisphosphate aldol	fructose-1,6-bisphosphate aldolase [Clostridium difficile 630]	2.13	7.54	_		
		Protein options 11.1 A C thioredoxin 2 (Trx2) [Clostridium difficile 630]		thioredoxin 2 (Trx2) [Clostridium difficile 630]	482	43.3		
		4.97	Α	С	transcription elongation factor GreA [Clostridium difficile 630]	90.1	18.1	*

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.

To generate values for **Absolute Quantitation for HCP** (Host Cell Proteins) select the 5th option on the drop down at the beginning of this section.

Note: there are 2 differences between this HCP-specific option and the Absolute Quantitation option:

- The mass of protein present is also reported in ng in addition to the amount in fmol
- The calculation of protein amounts and masses present is based on measurements of the calibrant protein present in the same run, and not a pooled measurement over all the runs. It does still assume a specified amount of a calibrant in every run. However, it does not pool information across the runs to derive its relationship between the calibrant and observed abundance. The amount of a contaminant is instead calculated directly by relation to the observed amount of the calibrant in the same run.

Review Proteins Using this screen, you can find the proteins of interest in your experiment.	Vo filter applied Create				🔞 Hel
Set the quantitation options	Description	Amount (fmol) - A	Amount (fmol) - C	Amount (ng) - A	Amount (ng) - C
If you've not already done so, choose between relative and absolute quantitation.	formatetetrahydrofolate ligase [Clostridium difficile 630]	9.45	11.2	0.57	0.674
use of Hi-N, protein grouping and more.	fructose-1,6-bisphosphate aldolase [Clostridium difficile 630]	4.22		0.14	0.168
Protein options	gamma-aminobutyrate metabolism dehydratase/isomerase [Clostridium difficile 630]	31.9	14.4	1.78	0.802
	glyceraldehyde-3-phosphate dehydrogenase [Clostridium difficile 630]	6.56	14.4	0.238	0.522

As 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 for HCP using Hi-N		Absolute Quantitation for HCP using Hi-N
Number of peptides to measure per protein (N): 3		Number of peptides to measure per protein (N):
Calibrant accession:		Calibrant accession: gi 126700790
Amount (fmol):		Amount (fmol): 50
Use protein grouping i.e. hide proteins whose peptides are a subset of another protein's.	, , , , , , , , , , , , , , , , , , ,	Use protein grouping i.e. hide proteins whose peptides are a subset of another protein's.
OK Cancel		OK 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.

How the values for the calibrant behave across runs comparing Absolute Quant and Absolute Quant for HCP using the same calibrant

Description	Amount (fmol) - A	Amount (fmol) - C		
elongation factor Ts [Clostridium difficile 630]	46.5	5.06		
enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) [Clostridium difficile 630]	25.4	74.6		
F0F1 ATP synthase subunit beta [Clostridium difficile 630]	3.31	11.4		
Description	Amount (fmol) - A	Amount (fmol) - C	Amount (ng) - A	Amount (ng) -
Description elongation factor Ts [Clostridium difficile 630]	Amount (fmol) - A 91.9	Amount (fmol) - C 3.4	Amount (ng) - A 3.06	Amount (ng) - 0.113
Description elongation factor Ts [Clostridium difficile 630] enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase) [Clostridium difficile 630]	Amount (fmol) - A 91.9 50	Amount (fmol) - C 3.4 50	Amount (ng) - A 3.06 2.31	Amount (ng) - 0 0.113 2.31

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 of	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.

QP 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...**

UP Progenesis QLp Tutorial for DDA - Progenesis QI for pr	oteomics												- D ×
File Review Experiment Import Data Alignment Filtering Design Setup	Review Peak Peptide Ic Picking Statistics	on Ide ; Pep	entify ptides	QC Metri	Refine cs Identifications	Re Pro	view teins	Prot Statis	ein tics	Report			nonlinear
		(A Waters Company
Review Proteins Using this screen, you can find the proteins of interest in your experiment.	W No filter applie	ed Cre	eate	Search		٩							@ Help
Set the quantitation options	Accession	Peptides	Unique p	eptides	Confidence score	Anov	a (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, use of Hi-N.	🎯 gi 126699603	3	3		236	1.23E	-06	3.26E-06	۵	5.37	A	С	decarboxylase *
protein grouping and more.	🎯 gi 126701103	1	1		51.7	5.06E	-06	1E-05	۵	6.17	Α	С	ribose-5-phose =
Protein options	🎯 gi 126699939	3	3		236	5.3E-0	06	1E-05	۵	8.6	Α	С	transketolase,
	gi 126697972	2	2		130	6.61E	-06	1.09E-05	۵	3.54	С	Α	fructose-1,6-bi
2 Create a shortlist to review In the table, sort and filter the proteins based	🔇 gi 126699971	5	5	Ano	va p-value ≤ 0.05		-06	1.33E-05	4	11.1	А	С	thioredoxin 2 (
on their measurements, to generate a shortlist	🎯 gi 126701179	3	3	Max	fold change ≥ 2	-	-05	1.33E-05		4.97	A	С	transcription el
for further review.	🎯 gi 254974053	1	1	Neu	-		05	1.41E-05	۵	898	С	A	cell surface pro
To cort the table by a given value, simply click	🎯 gi 126699756	3	3	Oui	r lag	-	05	1 415 05		0.10	Α	С	elongation fact
the relevant column header.	🎯 gi 260682215 (+1)	24	11	Qui	ck rags		-	Anova p-v	aiue		С	Α	hemagglutinin,
Review the proteins	🎯 gi 254973900 (+9)	15	15	T Edit	tags	-	-	Modificati	nange	_	Α	С	flagellin subun
3 For each protein of interest, review its peptide	🎯 gi 126697690	6	6	Add	to Clip Gallery			Sequence	011		Α	С	ferredoxin/flav
measurements and correlations:	🎯 gi 126700634	3	3		298	2.99		Pentide ta	 as contain		Α	С	PTS system ma
View peptide measurements	🎯 gi 126699940	1	1		73.3	3.46e	-05	2.095-00		7.10	Α	с	transketolase, I
You can also double-click to review a protein.	🎯 gi 126700297	2	2		164	3.67E	-05	2.7E-05	۵	2.29	Α	С	propanediol ut
A Export data for further processing	🎯 gi 126697583	1	1		79.3	4.33E	-05	3E-05	۵	3.78	Α	С	DNA binding p
4 By exporting your data to external tools, there's polimit to your applying.	ail126700790 (+1)	8	8		744	4.75E	-05	3E-05		2.94	С	Α	enolase (2-pho *
Export to pathways tool Export protein measurements	Selected prote	in: thi	oredox	in 2 (Trx2) [Clostr	idiu	m d	lifficile	630]				,
Export pentide measurements				А									
Export peptide ine measurements	inh Nomalised Abundance			H	¢						1 ‡ 1		
	274 4 -												
Experiment design Review your data from a different perspective:	4												÷
Current design: AC •	Quantifiable proteir	ıs display	yed: 145									Sectio	n Complete Э

As an example, let us explore Thioredoxin 2. The table indicates that this protein is most highly expressed in Condition A by 11.1 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.



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.



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.

	La Tabadal Kaal													
ur Progenesis QI.	i.p Tutorial for	UUA - Proge	nesis Qi for proti	eomics										
File	Review		Experiment	Review Peak Peptid	e Ion Io	lentify	Refine	Review	Pro	tein				
Import Data	Alignment	Filtering	Design Setup	Picking Statis	stics Pe	eptides QC Metri	cs Identifications	Proteins	Stati	stics	Report			nonimear
											0			A Waters Company
Review Prot	eins													🕜 Help 🔻
Using this scree interest in your	en, you can fir r experiment.	nd the prote	ins of	Wo filter ap	Cr	eate Search		Q						
🚽 Set the qu	uantitation o	ptions		Accession	Peptide	s Unique peptides	Confidence score	Anova (p)	q Value	Tag 🔹	Max fold change	Highest Mean	Lowest Mean	Description
If you've n	not already do elative and ab	ne so, choos solute quan	se titation	🎯 gi 126700198	1	1	46.6	0.000104	4.9E-05	۹	9.14	Α	С	phosphatase, 2C 🔺
use of Hi-N	N, protein gro	uping and m	nore.	🎯 gi 126700129	3	3	267	0.000125	5.68E-05	٠	2.32	Α	С	translation inhibi
		Protein	options	🎯 gi 254974605	4	4	311	0.000154	6.79E-05	۲	2.78	С	A	electron transfer
				🧐 gi 254973854 (-	+4) 5	5	428	0.000166	7.08E-05	۵	3.03	С	A	60 kDa chaperon
2 Create a s	shortlist to r	eview	ine	🎯 gi 126698915	ail254973	854 - 60 kDa chape	ronin [Clostridium d	difficile OCD	-66c261		_	с	А	30S ribosomal pr
based on t	their measure	ments, to ge	enerate a	🎯 gi 126698631 (-	gi 126697	767 - 60 kDa chape	ronin (Protein Cpn6	0) (GroEL p	rotein) [Clo	stridium d	ifficile 630]	С	A	cell wall-binding
shortlist fo	or further revie	ew.		🔇 gi 255101963 (-	gi 255305	5190 - 60 kDa chape	ronin [Clostridium o	difficile ATO	C 43255]			A	С	cell surface prote
Mow How	are the meas	surements c	alculated?	gi 126699972	gi 717328	373 - heat shock prote	tein GroEL [[Clostrid	dium] diffici	le]			Α	с	thioredoxin redu
To sort the click the re	e table by a gr elevant colum	ven value, s n header.	simply =	gi 126699583	2	2	197	0.000219	8.03E-05	•	5.1	С	A	RNA-binding pro
				gi 126697654	3	3	215	0.00023	8.22E-05		2.45	Α	с	30S ribosomal pr
3 For each p	ne proteins protein of inter	est, review	its	gi 126699299	2	2	311	0.000248	8.44E-05		3.34	A	с	dinitrogenase iro +
peptide me	easurements a	and correlat	ions:	٠			1	1	1					•
	View pept	tide measure	ments	Selected pro	tein: 60	kDa chapero	onin [Clostrie	dium di	fficile	QCD-66	c26]			
You can als	so double-clic	k to review	a protein.	View peptide measu	irements									
Export da	ata for furthe	er processi	ng				A							
By exporting there's no	ing your data f limit to your a	to external t analvsis.	cools,	· 7.0 -										
	· ·			E 6.0) (
	Export	to pathways	tool	Y FO			.							
	Export pro	tein measure	ements 💡	CSin CSin			•							
Experiment de	esign			₹ 4.0 =										•
Review your dat	ata from a diffe	erent perspe	ective:											,
Current design:	: AC		•	Quantifiable prot	eins displa	yed: 145							Secti	on Complete 🏵

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.

QIP Progene	esis QI.p Tutorial for	DDA - Pro	ogenesis QI for prot	teomics												
File Import Dat	Review ta Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide I Statistic	on Identify s Peptides	QC Metric	Refine s Identifications	Review Proteins	Protein Statistics	Report				non A Waters	inear Company
🛞 Return	n to list of proteins	Ŷ	Previous protein	🕂 Next prot	ein									n la hi		Help •
Review Review the validate th	selected protein e selected protein's heir expression pati	in s identifie terns.	d peptides and	Accession: Description	gi 2549 n: 60 kDa	073854 🌒 chaperonin [Clos	tridium diffi	cile QCD-66c26]					Y	No filte	r applied	eate
4 Choo	ose the level of de	etail		Σ #	Score	Correlation	Anova (p)	Max Fold Change	Highest Mean	Lowest Mean	Tag 💌	Abundance	m/z	Charge	Retention Time	(mins) N
View	the properties and	expressi	on profiles of	• 770	66.390	0.971	0.00273	2.07	С		-	1.784E+05	619.8642		32.549	
eithei	Show: Peptides Peptide Peptide <th< td=""><td>3.52</td><td>с</td><td>A</td><td>-</td><td>2.952E+05</td><td>766.9192</td><td>2</td><td>40.786</td><td></td></th<>							3.52	с	A	-	2.952E+05	766.9192	2	40.786	
Tint y	you can also double	a-click a r	pentide to select	• 978	77.730	0.988	0.000371	2.99	С	Α	-	2.918E+05	696.7028	3	54.573	0
and v	Tip: you can also double-click a peptide to select and view its component ions. • 2005 106.050 0.99						6.88E-05	3.83	С	Α	-	2.153E+05	1044.5499	2	54.624	-
) Com	pare expression p	profiles		0 1185	68.460	0.997	1.17E-06	51.3	С	A		1.44E+05	535.8258	2	35.849	
Selec	t peptide ions in th	e table t	o show their	• 13282	50.250	0.755	0.0187	10.6	С	A		4074	717.079	3	72.049	
capit	Select all peptide	ions	Select ions of 🔻													
Corre	elation values for th	ne expre	gi 2549738	54	lium diffic	ile.	L									
also t	be seen in the table	e.	QCD-66c26]	aronni [eioach		are										
3 Reso	olve any quantitat any peptide ions wh	tive ou	gi 12669776	7				Standardised Exp	ression Profiles			•				
an ou	utlier for this protei	n.	60 kDa chape protein) [Close	eronin (Protein stridium diffici	Cpn60) (e 630]	GroEL				С						
You c detail	an outlier for this protein.											_				
			gi 8886080 heat shock pr difficile]	rotein GroEL [Clostridiu	m]									~	
Experime Review yo	ent design our data from a diffe	erent pe	gi 71732873 heat shock pr difficile]	rotein GroEL [Clostridiu	m]										
Current de	esign: AC	_			promoo.	Contributes to	protein me	asurements 🔳 Do	oes not contribu	ute to protein r	measureme	ents				

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.

QIP Proge	enesis QI.p Tutorial fo	or DDA - Proge	enesis QI for pro	teomics												
File Import	Review Data Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Ion Statistics	Iden Pepti	tify des QC Metric	Refine Identifications	Review Proteins	Prote Statist	ein tics	R	eport		nonlin	near
_												1	STAR A		A Waters Co	mpany
Revie Using t interest	w Proteins this screen, you can t in your experiment	find the prote	eins of	Y No	filter applied	Creat	e gi 12670	0790	×						v	neip •
1 Se	et the quantitation	options	~	Accession	_	Peptides	Unique peptides	Confidence score	Anova (p)	q Value	Tag	•	Max fold change	Highest Mean	Lowest Mean	Descr
I If be us	you've not already o etween relative and se of Hi-N, protein g	tone so, choo absolute quar couping and n	ntitation, nore.	🔮 gi 1267	00790 (+1)	8	8	744	4.75E-05	3.3E-05			2.94	С	A	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 2 of the poorly correlated right click on the selection and Create a New tag, **Poorly Correlated**.

QIP Progenesis QI.p Tutorial for DDA - Progenesis QI for prot	teomic	s										E	- • •
File Review Experiment Import Data Alignment Filtering Design Setup	Revie Pic	w Peak king	Peptide Id Statistic	on Identify s Peptides	QC Metric	Refine Identifications	Review Proteins	Protein Statistics	Report			nç	nlinea
							•			18 1	1.8	AV	laters Company
	[↓ N	lext prot	ein										1 Help
Review selected protein	Acc	ession:	ail 1267	00790 🚳									
Review the selected protein's identified peptides and validate their expression patterns.	Des	cription	enolase 630]	(2-phosphoglyce	rate dehyd	ratase) (2-phospho	-D-glycerate h	ydro-lyase) [Cl	lostridium	difficile	No filter a	applied	Create
Choose the level of detail	Σ	#	Score	Correlation	Anova (p)	Max Fold Change	Highest Mean	Lowest Mean	Tag 💌	Abundance	m/z	Charge	Retention
View the properties and expression profiles of either peptides or individual peptide ions:	٠	774	88.370	0.995	0.00061	2.95	с	A	-	2.13E+05	664.8871	2	63.377 ^
Show: O Peptides @ Peptide ions	0	2729	105.600	0.977	0.00201	13.4	С	Α	-	1.099E+05	1024.0464	2	66.666
Tip: you can also double-click a peptide to select	0	849	57.950	0.962	0.000319	2.78	с	Α	-	1.052E+05	525.3028	2	39.086
and view its component ions.	•	3750		0.959	0.000485	2.57	С	Α	-	4.77E+04	730.37	3	53.054
7 Compare expression profiles	•	975	92.430	0.956	7.59E-05	2.95	C	Α	-	2.842E+05	860.4494	2	50.1
Select peptide ions in the table to show their expression profiles in the chart below.	0	4780	54.030	0.930	2.05E-05	5.12	С	А	-	1.996E+04	729.9083	2	41.666 ≡
Select all peptide ions Select ions of 🔻	•	2332	106.390	0.915	0.000278	3.05	С	Α	-	1.753E+05	1095.0501	2	53.054
Correlation values for the expression profiles can	0	5894	87.980	0.907	0.00123	4.23	с	Α	-	3.459E+04	974.4527	2	37.638
also be seen in the table.	0	13944		0.871	0.00657	16.8	с	Α	-	2844	649.9705	3	37.554
Resolve any quantitative outliers	0	9722	45.310	0.826	0.0436	8.07	С	A		8924	683.0336	3	66.666
5 Tag any peptide ions whose expression profile is	0	4625	105.850	0.675	0.193	Most abunda	nt		-	3.77E+04	974.0087		61.338
an outlier for this protein. Q Learn about tagging and filtering	0	10040		0.567	0.529	Anova p-valu	e ≤ 0.05	A		4287	649.6748	3	61.338 -
You can then review their identifications in more	- 6					Max fold char	ige≥2			_			,
detail at the Refine Identifications step.					St	Significantly u	ip in C			•			
			A			Up regulated i	in C	Contra				0	
		3.0 T				Up regulated i	in A	Create ne	ewiag			l	
		2.5			[New tag		- P	oorly Correlat	ed			
	Jance	2.0				Quick Tags	•					Canaal	
	-unqv	15				🚰 Edit tags						Cancer	
	allsed					Add to Clip G	allery		/				
	LI NOU	1											
	tandise							/	/				
	stand	0.0		a								0	
Experiment design Review your data from a different perspective:		-0.5		0			-	ē				6	
Current design: AC •	Pept	tide ion	profiles:	Contributes to	protein me	asurements 🔲 Do	oes not contribu	ute to protein r	neasureme	ents			

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

QP Progenesis QI.p Tutorial for DDA - Progenesis QI for proteom	25		
File Review Experiment Revi	w Peak Peptide Ion Identify Statistics Peptides QC Metrics Identifications Prot	view Protein teins Statistics Report	nonlinear A Waters Company
Refine Identifications If your peptide identifications include unwanted or irrelevant results, you can remove them here. Specify a set of deletion criteria In the batch deletion criteria, enter the property values for a set of identifications you want to delete.	Batch deletion criteria	their tags. Move tags to the appropriate boxes to	
 2 Delete the unwanted identifications To delete the identifications you don't want, click either: Delete Matching Search Results, to delete the highlighted IDS Delete Non-matching Search Results, to delete the IDs that are not highlighted 	Absolute mass Available tags: Rete Most abundant (2333 peptide ions) Anova p-value ≤ 0.05 (6174 peptide ions) Max fold change ≥ 2 (9822 peptide ions) Up regulated in A (428 peptide ions)	Show peptide ions that have all of these tags: Poorly Correlated (2 peptide ions) Show peptide ions that have at least one of these tags:	g search results) Reset the criteria
3 Reset the criteria to start again To specify another batch of identifications to delete, click Reset the Criteria and then return to step 1 above.	# S Significantly up in C (3946 peptide ions) Ø 4215 Up regulated in C (3946 peptide ions) Ø 4215 Up regulated in C (3946 peptide ions) Ø 4215 Ø 4216 66	Hide peptide ions that have any of these tags:	[* cell surface protein (S-layer precurs cell surface protein (S-layer precurs DNA-directed RNA polymerase sub
▼ Tag filter applied peptide ions may be hidden Edit # Total Hits m/z RT (min Charge Tag ▼ 4625 10 974.000 61.34 2 ■ 1004(649.674 61.34 3 ■	 ✓ 4218 66 ✓ 4229 ··· ✓ 4236 63 ✓ 4288 86 ✓ 4346 49 ✓ 4346 47 ✓ 4346 47 	OK Cancel	DNA-directed RNA polymerase sub cell wall protein V [[Clostridium] dif 30S ribosomal protein S13 [Clostrid fructose-1,6-bisphosphate aldolase rubrerythrin [Clostridium difficile 62 pentidyl-ong/u (s-trans isomerase

Waters

Then highlight the two 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)

QIP Progenesis	QI.p Tutorial for D	DA - Proge	nesis QI for pro	eomics													
File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review F Pickin	eak Pe g S	ptide Ion tatistics	Identif Peptide	y is QC	Metrics	Refine Identificatio	ons	Review Proteins	Protein Statistics	Report			A Waters Company
Refine Iden	ntifications de identifications	include unv	vanted or	В	atch delet	ion criteria	Score le	ss than	•				Sequen	ice Length (less than 🔻		
1 Specify In the ba values for delete.	a set of deletion atch deletion crit or a set of identif	on criteria eria, enter ications you	the property u want to		hoolute m		Hits le Mass le	ss than ss than	•					Charge (Sequence (less than contains		
2 Delete t To delete either: • Dele high	the unwanted id e the identificatio ete Matching Sea hlighted IDs	dentificati ns you don rch Results	ons 't want, click , to delete the		bsolute ii	Retentior	m/z le Time le	ss than ss than	•			Dalat	D Mo	difications	contains		Parat the mitoria
• Dele the	ete Non-matching IDs that are <i>not</i>	g Search Re highlighted	sults, to delet									Delet	e matering search	Tesuits	_ Delete non-matchin	ig search results	Reset the citteria
7 Reset th	he criteria to st	art again															
To specific click Resabove.	ify another batch set the Criteria ar	of identificand then ret	ations to delete urn to step 1	, E	# 4492 4521 4593	Score 53.44	Hits 4	m/z 652.97 600.30 700.87	RT(mins) (41.13 47.32 18.38	Charge M 3 19 3 17 2 13	Mass M 955.89 797.88 399.73	Mass err -0.07 1.12 -0.70	Sequence SNSDLNTVS FIEEIGYYN IAEELGNRE	Accessi gi 1267 gi 1266 gi 2549	on Modifications 7004 5988 1763	cell wall binding 30S ribosomal pr cell surface prote	[* protein [Clostridiu otein S16 [Clostrid in [Clostridium dif
J To specif click Res above.	ify another batch set the Criteria ar ilter applied de ions may be hic	of identificand then retuined	ations to delete urn to step 1 Edit		# 4492 4521 4593 4593 4625	Score 53.44 53.44 105.85	Hits 4 5	m/z 652.97 600.30 700.87 700.87 974.01	RT(mins) 0 41.13 47.32 18.38 18.38 61.34	Charge N 3 19 3 17 2 13 2 13 2 19	Mass N 955.89 797.88 399.75 399.75 946.00	Mass err -0.07 1.12 -0.70 -0.70 -0.04	Sequence NSDLNTVS FIEEIGYYN JAEELGNRE IAEELGNRE VNQIGTITE	Accessi gil1267 gil1266 gil2549 gil3575 gil2964	on Modifications 7004 763 5510 1520	cell wall binding 30S ribosomal pr cell surface prote ErfK/YbiS/YcfS/Y enolase [Clostrid	[* protein [Clostridiu otein S16 [Clostrid in [Clostridium dif hhG [Clostridium d um difficile NAP08
To specific click Resabove.	ify another batch set the Criteria au ilter applied de ions may be hic I Hits m/z R1 974.001 63 600 67	dden T (min Cha 1.34 2	Edit		# 4492 4521 4593 4593 4593 4625 4625 4634 4642	Score 53.44 53.44 105.85 105.85 60.64	Hits 4 4 5 5 5 5	m/z 652.97 600.30 700.87 700.87 974.01 974.01 1116.5: 863.41	RT(mins) (41.13 47.32 18.38 18.38 61.34 61.34 61.34 30.02 22.56	Charge N 3 19 3 17 2 13 2 13 2 19 2 19 2 19 2 19 2 19 2 19 2 19 2 19 2 19 2 19 2 19 2 19	Mass N 955.85 797.88 399.72 399.72 946.00 046.00 046.00 046.00 046.00 046.00 046.00 046.00 046.00	Vlass err -0.07 1.12 -0.70 -0.70 -0.04 elete selec -0.26	Sequence NSDLNTVS FIEEIGYYN AEELGNRE VNQIGTITE VNQIGTITE.ted peptide(s) QADREGYP	Accessi gil267 gil266 gil2549 gil2549 gil2964 gil2964 gil266 gil266	on Modifications 1004 19988 1550 15520 1007 1001 19991	cell wall binding 30S ribosomal pr cell surface prote ErfK/YbiS/YcfS/Y enolase [Clostrid enolase (2-phosp translation inhibi rubrerythrin [Clo	[* protein [Clostridiur otein S16 [Clostrid in [Clostridium dif shG [Clostridium d un difficile NAP08 hoglycerate dehys tor endoribonuclei stridium difficile 62

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 Methionine 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 296449145	4	4	450	0.086	0.00967		1.65	Α	С
🎯 gi 255101963 (+2)	61	16	6.71E+03	0.000201	8.78E-05		2.52	Α	С
🔇 gi 254976387 (+5)	63	18	7.07E+03	7.51E-05	4.33E-05	۵	2.52		А
🕥 gi 1267004	Anova p-va	alue ≤ 0.05	81.7	0.000762	0.000203	-	1.72	С	А
🕥 gi 1267004 🔘	Max fold cl	nange≥2	222	0.0108	0.00161	•	1.61	Α	С
🔮 gi 2551019	New tag		641						
🕥 gi 2551019 🛛 🔅	Quick Tags	•	Anova p-valu	e	3E-05		44	Α	С
🔇 gi 2549740 🕋	Edit tags		Max fold char	nge)137		898	С	Α
🔇 gi 2549763	Add to Clip	Gallery	Modification.)023		2	С	Α
🔇 gi 254975791	2	2	Sequence		319		1.14	Α	С
🕥 gi 254976383 (+2)	9	1	Peptide tags o	ontain	7E-05		482	С	Α
🕥 gi 126699073	2	2	122	0.0143	0.00201		2.23	С	Α
🚳 gi 126700407	16	16	1.64E+03	0.0498	0.00599	-	1.61	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 a protein has	
Modification with: Oxidation M]
Can I use wildcards?	
Apply the following tag:	
Modification with Oxidation M	
	_
Create tag Cancel	J

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	
Create a filter Show or hide proteins based on a selection of t the filter. For more guidance, please see the on	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 (109 proteins) Max fold change ≥ 2 (86 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.

QP Progenesis OLp Tutorial for DDA - Progenesis OI for prote-	omics										
File Review Experiment I Import Data Alignment Filtering Design Setup	Review Peak Peptide Io Picking Statistics	n Ide Pep	ntify tides QC Metri	Refine ics Identifications	Review Proteins	Prof Stati	tein stics	Report			nonlineal A Waters Company
Review Proteins Using this screen, you can find the proteins of interest in your experiment.	Tag filter appli proteins may be	ied e hidden (Edit	Search	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, use of Hi-N	🔇 gi 126699971						٨	11.1			thioredoxin 2 (
protein grouping and more.	🎯 gi 260682215 (+1)	24	11	2.08E+03	1.77E-05	2.1E-05	۲	37.1	С	A	hemagglutinin,
Protein options	🎯 gi 254973900 (+9)	15	15	1.95E+03	2.3E-05	2.51E-05	۹	3.58	Α	С	flagellin subun
	🎯 gi 126697690	6	6	578	2.86E-05	2.8E-05	۲	5.64	Α	С	ferredoxin/flav
7 Create a shortlist to review	🕥 gi 126700790 (+1)	7	7	638	4.75E-05	3.3E-05	۲	2.94	с	A	enolase (2-pho
on their measurements, to generate a shortlist	🎯 gi 209571234	25	12	2.5E+03	4.78E-05	3.3E-05	۲	5.99	A	С	cell wall proteir =
for further review.	gi 254976387 (+5)	63	18	7.07E+03	7.51E-05	4.33E-05	۲	2.52	С	A	cell surface pro
W How are the measurements calculated?	gi[126700129	3	3	267	0.000125	6.8E-05	٩	2.32	Α	с	translation inhi
To sort the table by a given value, simply click the relevant column header.	gi 255101963 (+2)	61	16	6.71E+03	0.000201	8.78E-05	٩	2.52	Α	с	cell surface pro
	🕥 gi 126697654	3	3	215	0.00023	9.14E-05	٩	2.45	Α	с	30S ribosomal
3 For each protein of interest, review its peptide	gi[126700372	2	2	126	0.00386	0.000743	٩	4.95	Α	с	PTS system HP
measurements and correlations:	gi[54781347	6	6	544	0.00517	0.000903		1.9	Α	с	2-hydroxyisoca
View peptide measurements	gi[126697631	7	7	626	0.00981	0.0015		1.41	A	с	50S ribosomal 👻
You can also double-click to review a protein.	+		1	m		1					Þ
4 Export data for further processing By exporting your data to external tools, there's no limit to your analysis. Export to pathways tool Export protein measurements Export peptide measurements	Selected prote View peptide measurem	in: thio nents	oredoxin 2(A	Trx2) [Clostr	idium d	difficile	e 630]		С	-	
Experiment design Review your data from a different perspective:	Arcsinh Ar								1 4 1	Sectio	n 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**.

how or hide proteins based on a selection of he filter. For more guidance, please see the <u>o</u>	their tags. Move tags to the appropriate boxes to creat <u>line reference</u> .
vailable tags:	Show proteins that have all of these tags:
 Anova p-value ≤ 0.05 (109 proteins) Max fold change ≥ 2 (86 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:

4 Expo By ex no lim	Export data for further processing 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								

QIP F	Progenesis QI.p Tutorial for DDA - Progenesis QI for prot	eomics
Г	File	
	Save	Review Peak Picking
	Close	
	Export peptide ion measurements	🤝 Tag
	Export peptide measurements	1 prc
	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
	E.c.a	🎯 gi 126
L^		🔇 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.

port protein measurements		
noose properties to be included ir	n exported file	
Accession		
Peptide count		
Unique peptides		
Confidence score		
🖊 Anova (p)		
🛿 q Value		
Max fold change		
Power		
Highest mean condition		
Lowest mean condition		
Mass		
Description		
Nomalized abundance		
Raw abundance		
Amount (fmol)		
Tags		
	ОК	Cancel

QIP 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

		/	()			Normalize	ad abundar	nce	· · · · · ·		
						A			С		
Accession	Peptide count	Unique peptides	q Value	Max fold change	Description	A1	A2	A3	C1	C2	C3
gi 126699971	5	5	1.33E-05	11.14094087	thioredoxin 2 (Trx2) [Clostridium difficile 6	1154.236	1215.294	1131.041	88.54863	109.4784	116.1807
gi 260682215;gi 209570719	24	. 11	1.78E-05	37.07319764	hemagglutinin/adhesin [Clostridium diffici	11.81321	16.35164	18.26058	530.6459	528.247	662.2461
gi 254973900;gi 260682017;gi	15	, 15	2.17E-05	3.581960969	flagellin subunit [Clostridium difficile QCD	5371.804	4705.512	5579.548	1446.399	1404.639	1519.994
gi 126697690	6	, 6	2.47E-05	5.642505943	ferredoxin/flavodoxin oxidoreductase sub	407.2672	466.0251	354.417	70.15977	74.2685	73.15401
gi 126700790;gi 296452046	7	7	3.00E-05	2.944231317	enolase (2-phosphoglycerate dehydratase)	62.67638	56.74192	64.69636	195.2659	177.3942	169.416
gi 209571234	25	, 12	3.00E-05	5.993442061	cell wall protein V [[Clostridium] difficile]	1917.599	1987.075	2256.879	369.1209	288.9127	370.0155
gi 254976387;gi 296452394;gi	63	, 18	3.86E-05	2.521368327	cell surface protein (S-layer precursor prote	\$ 25956.46	27703.39	25118.69	66993.88	60513.36	71122.46
gi 126700129	3	, 3	5.68E-05	2.318926178	translation inhibitor endoribonuclease [Clc	1468.906	1439.99	1571.524	672.1819	580.3695	679.5581
gi 255101963;gi 21702505;gi	61	. 16	7.81E-05	2.521286161	cell surface protein (S-layer precursor prote	€ 69028.69	73480.05	61068.81	28145.33	24515.63	28082.57
gi 126697654	3	, 3	8.22E-05	2.451280888	30S ribosomal protein S8 [Clostridium diffir	256.4914	288.947	294.6438	119.5426	121.4714	101.6975
gi 126700372	2	. 2	0.000676	4.953785887	PTS system HPr protein [Clostridium diffici	26.49133	34.09065	24.56343	3.318793	7.936991	5.932161
gi 54781347	6	6	0.000887	1.901229315	2-hydroxyisocaproate-CoA transferase [Clo	399.6587	323.1495	309.6788	179.5084	155.1933	208.3611
gi 126697631	7	7	0.00149	1.412713841	50S ribosomal protein L7/L12 [Clostridium /	607.4534	520.8522	662.8525	440.9649	412.827	414.0928
gi 126700078	6	, 6	0.001572	1.398255682	molecular chaperone DnaK [Clostridium dif	130.7325	118.3408	103.9304	173.6002	165.4416	154.5476
gi 126697969;gi 255654423	10	10	0.0036	1.733236988	Beta-subunit of electron transfer flavoprot	. 1696.664	1586.885	2411.808	1149.335	897.5683	1239.062

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.

MetaCore: MetaCore is an integrated software suite for functional analysis of experimental data. MetaCore is based on a curated database of human protein-protein, protein-DNA interactions, transcription factors, signaling and metabolic pathways, disease and toxicity, and the effects of bioactive molecules.

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 93).

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.

Progenesis QLp Tutorial for DDA - Progenesis QI for prot	omics				
File Review Experiment Import Data Alignment Filtering Design Setup	Review Pesk Peptide Ion Identify QC Metrics Refine Review Protein Picking Statistics Peptides QC Metrics Identifications Proteins Statistics	Report			nonlinea
		0.10			A Waters Company
Review Proteins Using this screen, you can find the proteins of interest in your experiment.	Tag filter applied Edit				@ Help '
Set the quantitation options	Acces	Max fold change	Highest Mean	Lowest Mean	Description
If you've not already done so, choose between	gi QP Export Pathways Information	482	С	A	cell surface pro
protein grouping and more.	🕥 gi Select a pathways tool	17.2	С	Α	FMN-binding r ≡
Protein options	Solution of the list below. You can find out more or download new plugins using the links below.	5.45	С	A	glyceraldehyde
	🕥 gi	23.4	С	Α	ABC transporte
2 Create a shortlist to review	gi Which pathways tool do you want to use?	5.37	A	с	decarboxylase
on their measurements, to generate a shortlist	gi IMPaLA: Integrated Molecular Pathway Level Analysis	5.17	Α	С	ribose-5-phosp
for further review.	IMPaLA: Integrated Molecular Pathway Level Analysis ITPA: Incenuity Pathway Analysis	8.6	A	С	transketolase, c
Mow are the measurements calculated?	🔮 gi MetaCore	8.54	С	A	fructose-1,6-bi
the relevant column header.	s gi Panther Classification System	1.1	А	с	thioredoxin 2 (
Deview the evoteine	🕥 gi	1.97	Α	с	transcription el
3 For each protein of interest, review its peptide	🕥 gi	398	С	Α	cell surface pro
measurements and correlations:	📦 gi	9.19	A	с	elongation fact
View peptide measurements	📦 gi	37.1	С	A	hemagglutinin, +
You can also double-click to review a protein.					Þ
4 Export data for further processing By exporting your data to external tools, there's no limit to your analysis. Export to pathways tool	Sele		С		
Export protein measurements		_			
Export pentide measurements					
capore peptide measurements	NO 6 -				
Export peptide ion measurements	4 regime		3 4 1		
Experiment design Review your data from a different perspective:	4				÷.
Current design: AC	Quantifiable proteins displayed: 84			Section	n Complete Э



QIP P	rogenesis QI.p Tutorial for DDA - Progenesis QI for prot	eomic
F	File	
	Save	Revie Pic
	Close	
	Export peptide ion measurements	
	Export peptide measurements	
	Export protein measurements	Ace
	Export to pathways tool	9
	Import additional protein data	0
	Import protein accessions as tag	3
	Export mzIdentML for PRIDE submission	3
		3
	Experiment properties	9
	Show Clip Gallery	3
	5.4	9
_		9



Select either **Pathway over-representation analysis** or **Wilcoxon pathway enrichment analysis.** 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 Molecular Pat	hway Level Analysis
pathway over-representation and enrichment analysis with	expression and / or metabolite data
genes/proteins - example input for over-representation analysis - example input for enrichment analysis	- example input for over-representation analysis - example input for enrichment analysis
paste genes or proteins below	paste metabolites below
gi 1254976387 0.12 gi 1384359782 0.28 gi 132870407 0.19 gi 122670497 0.19 gi 12269951234 -1.8 gi 1226995450 1.34 gi 1226997969 -0.27 gi 1226997969 -0.27 gi 122699789 0.28 gi 122699718 -1 gi 122699718 -1 gi 1226699128 -0.12 or upload a file with genes or proteins Browse	or upload a file with metabolites Browse
optionally, provide genes/proteins background for over-representation analysis Browse	optionally, provide metabolites background for over-representation analysis Browse
Unigene 🗸	specify metabolite identifier 🗸
choose analysis type: pathway over-representation analysis Wilcoxon pathway enrichment analysis START ANALYSIS or clear the form	

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

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	
Open IMPaLA in my browser	into
< Back Copy proteins to clipboard Close	•

Although the previous page outlines the process of exporting data to pathway tools, a text file (Impala_Wilcoxon_demo.txt) has been included in the tutorial download to allow the demonstration of exporting data

to Impala for pathways analysis.

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.

Configuration and the second second		
Choose which identifications to export a	and the i	type of analysis you want to perform.
Select the type of analysis to perform:		
Wilcoxon pathway enrichment analysis		•
For Wilcoxon enrichment analysis, choo	se two (experimental conditions that you would
ike to compare.		Comparison
-		
A	• 5	• C •
if greater, gives a negative fold chang	e.	If greater, gives a positive fold change.
To perform the pathway analysis, copy the genes/proteins section of the IMPat	the prot A searc	tein data to the clipboard and paste it into h page.
Open IMPaLA in my browser		



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.

Cholesterol biosynthesis

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

0.0975

22

Reactome 4

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.

Open	IDE Inspector 2.5.2 Export Help		
	Quick Start	Try Examples	
A	Open Ident	Select mzML/mzXML/mzid/PRIDE xml Files Look Jn: PRIDE submission DDA Tutorial MSMS.mgf DDA Tutorial.mzid ML	
	Review Pro	File Name: DDA Tutorial.mzid Files of Type: [mzid or .mzidentmi or .mzmi or .mzi or .mgf or .ms2 or .pki or .dta or .api or .mziab or .cdf or .gz 🔻	More examples
	Feedback	Open Cancel	
	🖂 Give Us Your Feed	dback	
	When use PRIDE Inspector, p	Nease cite:	

Click open.

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

QIP P	rogenesis QI.p Tutorial for DDA - Progenesis QI for prot	e
	File	
	Save	F
	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**

🕾 Load spectrum files				×
🕠 mzldentMI Files				🚷 Help
🧪 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.

🔺 Pi	RIDE Inspector 2.5.2														
Ope	n <u>E</u> xport <u>H</u> elp										_				
	l.j Assays	Overview Protein Peptide	Overview Protein Peptide Spectrum Summary Charts												
	🧪 DDA Tutorial.mzid 🛛 🛄 🗙	Protein List					👤 Up	🛓 Update Protein Details 🛛 🍸 Decoy Filter 🕜 (
9		Protein Group ID	#	#PSMs		#Distinct Peptides	▼ #	PTMs		5					
~		+ ProteinGroup_1_gi_x007C_2	2 <u>qi 254976387</u>	4	176		58	7							
a		+ ProteinGroup_2_gi_x007C_2	2 <u>qi 255101963</u>	4	453		56	6							
		ProteinGroup_3_gi_x007C_2	2 <u>qi 209571234</u>	1	133		23	2							
		+ ProteinGroup_4_gi_x007C_2	2 <u>qi 260682215</u>	1	112		22	2			- 11				
		ProteinGroup_5_gi_x007C_1	1 <u>qi 126700407</u>	8	38		17	0							
		+ ProteinGroup_6_gi_x007C_2	2 <u>qi 254973900</u>	1	100		12	5			- 11				
		+ ProteinGroup_7_gi_x007C_1	1 <u>qi 126698450</u>	7	70		11	1			- 11				
		ProteinGroup_9_gi_x007C_1	1 <u>gi 126697970</u>	3	38		9	1							
		+ ProteinGroup_10_gi_x007C	<u>qi 254976383</u>	3	33		9	0			- 11				
		+ ProteinGroup_8_gi_x007C_1	1 <u>gi 126697969</u>	Ę	50		9	2							
		ProteinGroup_11_gi_x007C_	<u>qi 255101959</u>	3	31		8	0			_				
		ProteinGroup_12_gi_x007C_	<u>qi 126699128</u>	4	19		7	0							
		ProteinGroup_13_gi_x007C_	_qi 126698643	2	23		7 0								
		ProteinGroup_14_gi_x007C	_qi 126700790	3	31		7	2							
_		PSM List [gi]126697970] Modified residues: [C - 57.0215]													
	🕅 Assay Summary	Peptide	Ranking	Delta m/z	Charge	Precursor m	z Modifications	Length	Start	Stop					
	🕒 Spectra found	IAPWIELLGEGR	1	-0.0002	3	455.9413		13	24	36					
	Proteins found	IAPVVIELLGEGR	1	-0.0002	3	455.9413		13	24	36					
	• Hotelina lound	IAPVVIELLGEGR	1	-0.0002	3	455.9413		13	24	36					
	Protein Group found	IAPVVIELLGEGR	1	0.0003	2	683.409		13	24	36	- 11				
	Pentides found	IAPVVIELLGEGR	1	0.0003	2	683.409		13	24	36	<u></u>				
	Mod: LINIMOD:4		1	0 0003	2	602 400		12	24	26	•				
		Spectrum Fragmentation Ta	able Sequence												
		🛆 Protein sequence is retrieve	ed from the original f	ile					👤 Update	Protein De	etails				
	 Auto MS/MS Annotations 						Selected P	TM Fit	Euzzy Eit	Overla	an				
		Accession: 126697	970						- 1 422) 1 10		· · ·				
		38 peptides (38 mat	tched, 9 distinct), 98/	345 amino aci	ids (28.4%	6 coverage)									
		MNDTKDLSSY	KNVWTFAEOR	EGKTAP	VVTE	LLGEGRKLAK	EVDAELCATL	LGK DVDG	TAK	60					
		ELITEGADKV	YVADDALLEK	YTTDAY	TKVI	KDATDE TKPE	IMLEGATHIG	BDLAPRI	ASR	120					
		VGTGLTADCT	KLETDPEDKK	TKOTRP	AFGG	NTMATITCPN	HRPOMSTVRP	GVMDKAE	KDE	180					
		TRTGEVIALD	YKTTODDTRT	TVLETV	KTKK	DLVSLTDANV	TVSGGLGLGG	PEGFEMI	KKT	2.40					
		ADKLGGVVGS	SRAAVDAGWI	DHSHOV	GOTG	TTVKPNLYIA	CGISGAIOHL	AGMOSSD	FII	300					
		AINKNPAAPI	LEIADYGVVG	DLHEIV	PMLI	EKLDSVDDLL	EAIKA		HOLYGODE II						

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

Note: For exporting mzldentML 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 96)**

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.**

QIP Filter the proteins												
Create a filter Show or hide proteins 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 proteins that have all of these tags:												
 Anova p-value ≤ 0.05 (109 proteins) Max fold change ≥ 2 (86 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:											
<u>C</u> lear 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 QI.p T	utorial for DDA	- Progenesis QI	for pro	teomics									
File													
Re Import Data Alic	eview	Experi	ment	Review Peak	Peptide Ion	Identify Roptidor	OC Matrice	Refine	Review	Protein	Report		nonlinear
Import Data Alig		ening Design	Setup	FICKING	Statistics	replices	QCIMETICS		FIOLENIS	Statistics	Report		
												1.2	A Waters Company
Protein report Pe	otide report												
Report on you	r proteins					Stru	icture the re	port					
Generate a report	containing the	proteins of inte	erest in	your experime	ent.	Enter	r a report title:						
1 Choose wha	t to report up	on				Tuto	orial Methionine	oxidised proteins	5				
Using the list include in the	below, filter yo report.	our data to show	v only	the proteins yo	ou want to	Selec	ct the sections t	o include in the	report:				
			0 L	earn about tag	ging and filteri	ng 🕑 [Overview run						
Customise t	e report					(•)	Data process	ng methods					
Z Enter a title f	or your report	and select the s	ection	s you want to i	nclude in it usi	ng 🖓 i	Evnariment d	ecian					
the controls t	o the right.						i experiment d	esign				_	
🤝 Tag filter a	pplied						Protein repor	t					
proteins ma	ıy be hidden				Edit	Inclu	ide tables show	ing protein abun	idances and j	peptides identi	fied for each	protein	
Accession	Anova (p)	Fold Ta	q 💌	Description			Protein table						
gi 126699971	9.17E-06	11.1	4	thioredoxin 2	(Trx2) [Clostridi	ur	Peptide tabi	es					
gi 260682215 (+1)	1.75E-05	37.1	٩	hemagglutini	n/adhesin [Clost	tri 📀 [V Protein Detai	s					
gi 254973900 (+9)	2.3E-05	3.6	٩	flagellin subu	nit [Clostridium	di Repo	orts the full deta	ils of every prot	tein which ma	atches your cu	rrent filter		
gi 126697690	2.86E-05	5.6	٩	ferredoxin/fla	vodoxin oxidore	ed	Tags						
gi 126700790 (+1)	4.75E-05	2.9	٩	enolase (2-ph	osphoglycerate	d	Expression p	orofile					
gi 209571234	4.78E-05	6.0	٩	cell wall prote	in V [[Clostridiu	im 🕑 [Peptide ion tag	able					
gi 254976387 (+5)	7.51E-05	2.5	٩	cell surface p	rotein (S-layer p	re 🕥	Peptide ion d	etails					
gi 126700129	0.000125	2.3	٩	translation in	hibitor endoribo	oni C							
gi 255101963 (+2)	0.000201	2.5		cell surface p	rotein (S-layer p	re Crea	ate report						
gi 126697654	0.00023	2.5		30S ribosoma	I protein S8 [Clo	ost							
gi 126700372	0.00368	5.0	٩	PTS system H	Pr protein [Clos	tri							
gi 54781347	0.00517	1.9		2-hydroxyiso	caproate-CoA tr	an							
gi 126697631	0.00981	1.4	•	50S ribosoma	I protein L7/L12	2 [0							
gi 126700078	0.0105	1.4	•	molecular cha	aperone DnaK [(Clc							
gi 126697969 (+1)	0.0289	1.7	•	Beta-subunit	ot electron tran	ste							
•						•							
Experiment design	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

Tutoria	l Meth	nioni	ine oxio	dise	d pr	oteins											
Experiment	: Progene	esis QI	.p Tutorial	for DI	DA										-		
Report crea	ted: 01/1	10/201	9 10:08:2	9													
Proteins	Proteins																
Protein buildin	rotein building options																
Protein group	Vrotein grouping Group similar proteins																
Protein quantitation Relative Quantitation using Hi-3																	
Accession	ccession Peptides Score Anova Fold Tags Description Average Normalised																
			(P)"										Abur	C			
<u>gi 254976387</u>	63 (18)	7074.5	0 7.51e-005	2.52		cell surface	protein (S-layer pr	ecurs	sor protein)	[Clostridi	ium	3.49	e+007 8.80e+00	07		
<u>gi 255101963</u>	61 (16)	6709.7		difficile QCD-66c26]													
gi 209571234	25 (12)	2502.3	<u>gi 254</u>	254973900													
gi 260682215	24 (11)	2078.2	flogolling	ubusit	[Clea	tridium d	ifficile 0	CD 66-2	061								
gi 254973900	15 (15)	1945.7	15 peptid	es		striaium a	inneite Q	CD-00C	0]								
g <u>i 126697969</u>	10 (10)	925.4	Soguanca				Doptido	Score	Hite	Mass	Chargo	Tage	Conflicts	Modifications	In	Avorago No	brmalisod
<u>gi 126700790</u>	7 (7)	638.0	Sequence				lon	Score	TIILS	mass	Charge	lags	connicts	mounications	quantitation	Abundance	es
<u>gi 126697631</u>	7 (7)	625.7		AICEIZ			149	100.52	5	1220 6097	2					A	C
gi 126700078	6 (6)	582.2	DTDVASEM	VNLSK			1358	98.96	4	1407.6558	2		0		no	1.27e+005	3.41e+004
			DTDVASE	VNLSK			3339	84.71	4	1423.6495	2		0	[8] Oxidation	no	7.64e+004	1.48e+004
			IADELLQL	<		Acce	ssion gi	25497	3900) (+9)							
			IADELLQL	KDEVER													
			IADELLQLK	KDEVER		Descri	otion flag	ellin sub	unit	[Clostridiu	m diffici	le QC	D-66c26]				
					(Pep	tides 15 core 194	(15) 5.78									
			IRDTDVASE	EMVNLSP	、 (- A	nova 2.3	De-005									
			IRDTDVASE		<	_	Fold 3.5	8 Iva p-val	ue <	0.05							
			ISSUEENIC	ĸ		_	Max	fold ch	ange	≥ 2							
			LESTQNNL	NNTIEN	VTAAES	iR	Mod	lificatior	wit	h Oxidatior	n (M)						
			LESTQNNL	NNTIEN	VTAAES	R			_	٨				<u> </u>			
			M NILVQAS														
							-										
						16.5 P				E							
						16.0											
						N 15.5) \$ 1			
						4uiSo 14.5											
						4 14.0	1										
							4									,	

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.



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**.

Review	Review Peak Picking													
7	No filter ap	plied					Create							
#	Anova (p)	q Value F	old	Tag 💌	Notes		Highest Mean	Lowe	st Mean					
10316	2.08E-13	4.84E-	Most ab	undant	_	note				*				
9128	1.93E-12	2.24E-	Anova p	-value ≤ 0	.05		А	С						
4026	3.08E-12	2.39E- 🧧	Max fold	d change ≥	2		С	Α						
144	5.27E-12	3.07E- 🦲	Significa	antly up in	с		С	Α						
7568	1.99E-11	9.25E- 🦲	Up regu	lated in C			А	С						
6348	2.43E-11	9.42E- 🔲	Up regu	lated in A			А	С						
1917	3.06E-11	1.02E- 🦲	Poorly C	Correlated			С	А						
9494	5.34E-11	1.29E-	New tag				Α	С						
11023	6.02E-11	1.29E-	Quick Ta	ags	•	And	va p-value							
314	6.35E-11	1.29E- 🔗	Edit tags	;		Max	fold change							
3233	6.46E-11	1.29E-08 I	nfinity	-	4	Mo	dification							
10463	6.94E-11	1.29E-08 I	nfinity		12	No	MS/MS data							
1450	7.4E-11	1.29E-08 I	nfinity		15	No	protein ID			_				



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	at have no ms/ms.
--------------------	--------------------	-------------------	-------------------

Revie	eview Peak Picking																
7	Y Tag filter applied peptide ions may be hidden															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	
10316	2.08E-13	4.84E-10	Infinity			С	A	1002.0213	2	2002.028	33.698	0.212	8.82E+03	9.91E+04	0.73	0	*
9128	1.93E-12	2.24E-09	Infinity	-		Α	С	592.2268	3	1773.659	42.473	0.294	5.93E+03	5.28E+04	1.22	0	
144	5.27E-12	3.07E-09	Infinity	-	1	С	Α	901.2218	2	1800.429	40.050	1.27	4.2E+06	1.43E+07	2.69	0	
7568	1.99E-11	9.25E-09	Infinity		🔜 Add a note	А	С	822.3762	2	1642.738	40.221	0.431	1.31E+04	7.11E+04	2.36	0	
6348	2.43E-11	9.42E-09	Infinity		2	А	С	735.3522	3	2203.035	39.378	0.444	1.79E+04	9.02E+04	2.56	0	
9494	5.34E-11	1.29E-08	Infinity		2	Α	С	680.3218	3	2037.943	42.561	0.42	9.15E+03	4.99E+04	2.95	0	
11023	6.02E-11	1.29E-08	Infinity	-	2	С	Α	1446.7307	3	4337.170	53.154	0.227	2.2E+04	9.16E+04	3.27	0	
3233	6.46E-11	1.29E-08	Infinity	-	2	С	A	533.9859	3	1598.936	31.074	0.455	1.87E+04	3.32E+05	3.28	0	
10463	6.94E-11	1.29E-08	Infinity	-	2	Α	С	1052.1714	3	3153.492	31.272	0.0772	8.33E+03	5.01E+04	3.1	0	
4636	8.72E-11	1.35E-08	Infinity	-		A	С	666.6138	3	1996.820	43.337	0.633	3.15E+04	1.47E+05	3.7	0	
5323	9.81E-11	1.43E-08	Infinity	-		Α	С	946.7999	3	2837.378	34.444	0.48	4.09E+04	1.2E+05	3.94	0	
5770	1.14E-10	1.46E-08	Infinity	-		A	с	1223.0393	2	2444.064	35.134	0.472	6.4E+04	1.07E+05	4.29	0	
14169	1.16E-10	1.46E-08	Infinity			Α	С	1012.2141	4	4044.827	29.887	0.264	3.67E+03	2.69E+04	3.22	0	-
4]							Þ	

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

Then select the appropriate format.

OP Export	inclusion list					
Select your machine type and export the inclusion list:						
Format:	Thermo Finnigan inclusion list	About this data format Download others				
Save to:	ave to: C:\Users\andy.borthwick\Documents\Customer Data\Progenes Browse					
		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 Progenesis QI.p Tutorial for DDA - Progenesis QI for proteomi						
	File					
	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 Cip Gallery					
\times	Exit					
10950 3.87E-11 Infinity						
Inclusion list retention time windows						
Do you want to widen the retention time windows						
by 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 79).

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.

QIP 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 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...

UP Progenesis QLp_Tutorial for DDA - Progenesis QI for proteomics								
File Review Experiment R Import Data Alignment Filtering Design Setup R	eview Peak Peptide Ion Identify Picking Statistics Peptides QC N	Refine Review Protei Vetrics Identifications Proteins Statisti	Report A Waters Company					
Import Data Select your run data Select one of the available data formats then click the Import button: Commat Mathematical Select Your Your Select Your Select Your Your Select Your Your Your Your Your Your You	QIP Select files Organize ▼ Progenesis ► Progen Organize ▼ New folder ★ Favorites ■ Desktop	esis QL.p Tutorial for DDA_v3.0 >	A Waters Lompany W Help ~ Search Progenesis QLp Tutoria P BE Compared to the search Progenesis QL p Tutoria P Date modified al for DDA.Analysis 26/04/2016 10:40					
No runs have been imported yet	 Downloads Dropbox Recent Places ShareFile BitTorrent Sync Desktop Libraries Documents Music 	Al.mzml Al.mzml Al.mzml Cl.mzml Cl.mzml Cl.mzml Cl.mzml Cl.mzml Al.mzml Cl.mzml Cl.mzm	20/05/2008 08:23 20/05/2008 08:27 20/05/2008 08:29 20/05/2008 11:50 20/05/2008 11:50 20/05/2008 08:25 20/05/2008 08:25 mzXML files (*.mzml) Ωpen ▼ Cancel					
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.

st suitable run fro :	om candidates	that I select	
:	in canalates	chara scient	
	-		
	*		

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.



QP Start automatic processing	9		OP Start automatic processing	J		×	
Select your alignment re To mark a run for assessme use the button below to ma	eference candidates ent as an alignment referen ark multiple runs at once.	ce candidate, double-click it or	Select your alignment reference candidates To mark a run for assessment as an alignment reference candidate, double-click it or use the button below to mark multiple runs at once.				
Mark selected runs	ear all	٩	Mark selected runs	ear all	م		
		A				A	
A1	A2	A3	A1	A2	A3		
						E	
C1	C2	C3 +	E C1	E C2	C3	•	
0 of 6 runs marked as can	didates		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 OLp Tutorial for DDA - Progenesis OI 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 Donninear
	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:	612.5 825.0 1037.5 1250.0 1462.5 1675.0 1887. m/z *	About this run • MS peak count: 1,182,000
Format: mzXML files	-	 MS/MS count: 9,542 Total ion intensity: 3.643e+009
About this data format + Download others	0	<u>Masked areas</u> : none
2 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	40	
Automatic processing can be started while runs are importing.		
Imported runs: Search	8	
	8	
A1 A2 A3		
W. W. W.	100 n Time (min)	
C1 C2 C3	etentio	
	α v Zoom: ℚ ⊆ Σ	Section Complete 🤿

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.

Automatic processing complete (with warnings). Time taken: 2 minutes 58 seconds						
 Importing runs: 	/ of / 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					
	Relative Quantitation using Hi-3					

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	ing complete (witl ' seconds	th warnings).	
A Importing runs:	7 of 7 processed	1	
	🔺 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>⊆</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).

OP Processing Complete		—X —
Automatic process Time taken: 55 seconds	ing failed.	
A Importing runs:	7 of 7 processed A 1 failed to import	
× 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 right of the screen.

<<	Acti	ions 🔻 >	
5 1250.0		Mask areas for peak picking	
	×	Remove run	.,182,000 542
L	_	I otal ion intensit <u>Masked areas</u> : r	ÿ: 3.643e+009 Ione

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 areas 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.

IP Progenesis QLp Tutorial for DDA - Progenesis QI for pro File	eomics	
Import Data Licensing Alignment Filtering Design 3	an reliefer reak Pepude ton Denity Denity Indiane Protein Statistics Report Difference Protein Statistics Report D	Waters Compar
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.bothwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge Unlicensed	V
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	V
a sales representative.	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge Unlicensed	V
Once licensed your runs can be analysed on	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge Unlicensed	
any installation of the software. The licence is	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge Unlicensed	V
automatically included when archiving an experiment	C:\Users\andy_borthwick\Documents\Customer Data\Progenesis QI.p v3.0 Tutorials and Demo Suites\Proge 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: Internet code: Use Licen	nce Code
	Section	Complete

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 (in this example is 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 Ion 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 Ion intensity Map will also update with each additional vector.

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

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



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



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



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

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

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.
- 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	ignment		×
Choose whic	h runs to automatically align:		
Run	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

P Progenesis	QI.p_Tutorial fo	r DDA - Prog	enesis QI for p	roteomics											Σ
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	Rep	ort	n A	waters Cor	ea ea
AC	AC-2	C Nev	N												Help
Which ex	operiment o	lesign typ	e do you	want to us	e for this ex	periment	?								
00 00 00	Between-s	ubject De	esign				0-0 0-0 0-0	Within-sub	oject Design	n					
Do sam appear use the	ples from a gi in only one co between-sub	ven subject ondition? T ject design	hen A	_	۵1	Delete Remove	Have y subjec Then u	ou taken sam t under differ ise the within	ples from a giv ent conditions -subject desig	/en ;? n.		Before	During	After	
To set u the runs (factor l ANOVA o	p this design, according to level) of the s calculation as	you simply the conditio amples. Th sumes that	group on e the	QIP En	Create New Expe	riment Design he experimer	nt design:		X	use p	Patient X	X1	X2	X3	
conditio therefor whether are all e	ons are indepe re gives a stat r the means of equal.	ndent and istical test of the conditi	of ions	Ho	erore During and ow do you want Group the rur	to group the instantial	runs?			F	Patient Y	Y1	Y2	Y3	
			Add	conditio	○ Copy an exist	ing design:	ſ	Cranta darian	T Cancel	F E	Patient Z	Z1	Z2	Z3	
							L	Create design	Cancel	d					
							A stand becaus assump repeatu differe reduce conditi create The wi though paired- compai repeatu	and ANOVA is e the data viol otion of indepe demeasures A nces can be el d as a source o n differences a more power thin-subject di t of as an exte samples t-test rison between ed measures.	not appropriat lates the ANOV ndence. With N NOVA individu: iminated or of between (which helps t ful test). esign can be nsion of the to include more than two	e '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.

OIP Progenesis	OLp Tutorial fo	r DDA - Proc	ienesis OI for pro	oteomics						
File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Pe Picking S	ptide Ion Identit tatistics Peptid	fy es QC Metrics	Refine Identifications	Review Protein Proteins Statistics	Report	A Waters Company
AC	AC-2	Befor	e During an	d After Treatm	ent I × 🗅	New		7. de 17. 101 d		Ø Help •
Setup con Setup the co experiment o of your samp in the grid.	ditions and su design on the r ples to the corr	subjects ibjects for y ight, and the ect subject/	our en assign each condition cell		Before	During	After	Add Condition		
1. Add a colu 2. Add a row 3. Drag each in the grid	umn for each cu v for each subje n of your sampl d. es: P	ondition. ct. es to the co	rrect location	Patient A	A1	A2	A3	_		
	C	2		Subject 1	C1	Select Sample C2 C3	Select Sample			
	C	3 		Add Subject				_		
									Sect	ion Complete 🏵

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 Anova p-value ≤ 0.05 peptide ions (8174), as an example, view the power analysis.



This is displayed graphically showing that 79.1% of the 8174 peptide ions have a power of 80% or that 4 replicates would give you 96.7% 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

QIP 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 a subset of another protein s.	are a 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 88 conflicts. To view the conflicting assignments click on the **Protein Resolution** tab (window C previous page).

File	Review		E	xperiment	Review Peak	e Pep	tide Ior	n	Ident	ify		Refine	Review	Protein	1			n	onlin
oort Data	Alignment	Filterin	ng De	sign Setup	Picking	Sta	atistics		Peptic	ies QC	Metrics		Proteins	Statistics	Report			-	Y N A M
•				-			•									81 <i>#</i> 2	1	A	waters cur
periment de	esign: AC				•														
oteins	-	No filter a	applied				Peptio	de io	ons of	gi 2549	76387					- No	filter app	lied	
	Y				Create											Υ			Create
Accession		Peptides	Unique	Conflicts	Score		+	ŧ	ΣĢ	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	- Abundance	e Conf
gi 2549763	887 (+5) 🜼	63	18	88	7.07E+03		V	1	•	2 75.2	5	-0.624	2413.301	2.48	55.6	3		1.24E+08	1
gi 2551019	963 (+2)	61	16	- 22	6.71E+03		V	17	• (128	5	0.464	2413,296	0.265	55.6	2		9.38E+07	1
gi 2095712	234	25	12	¹⁸ A	2.5E+03		V	533	• (63.7	2	-0.931	2413 295	B -0.243	55.6	4		1.25E+06	1
gi 2606822	215 (+1)	24	11	18	2.08E+03		V	761	0	L 61.3	4	-0.790	2413.295	0.0288	50.5	3		3E+05	1
	107	16	16	0	1.64E+03		1	256	0	2 123	5	0.844	2413.294	-0.699	50.5	2		2.84E+05	1
gi 1267004							_		0 1	1 1 2 2	E	0.494	2413,294	-0.603	53.1	2		1.675.05	1
gi 1267004 gi 2549739	900 (+9)	15	15	0	1.95E+03	-	2	421	· · ·	J 122	2	0.404			2212	4	_	1.0/2+03	
) gi 1267004) gi 2549739 Protein: g Protein: g ptide ion vie	000 (+9) yi 254976 yi 255101 www.protei	15 387 cel 963 cel in resolutio	15 Il surfac Il surfac	0 ce protei ce protei	n (S-layer n (S-layer	recu precu	₹ rsor p rsor p	prote	ein) (ein) (Clostridio Clostridio	um diffic um diffic	cile QCD-66	c26] q42]		5512	L	-	1.0/2+03	,
gi 1267004 gi 2549739 Protein: g Protein: g tide ion vie	000 (+9) yi 254976 yi 255101 ws Protei g protein	15 387 cel 963 cel in resolution is for per	15 Il surfac Il surfac n otide io	0 ce protei ce protei n 533	1.95£+03 ▶ n (S-layer n (S-layer	• precu precu	v rsor p rsor p rsor p	prote	ein) [(ein) [(ons o	Clostridio Clostridio	um diffic um diffic 101963	cile QCD-66	c26] q42]			2	-	1.0/2+03	
gi 1267004 gi 2549739 Protein: g Protein: g tide ion vie Conflicting Accession	000 (+9) gi 254976 gi 255101 ews Protei g protein	15 387 cel 963 cel in resolution is for pep Peptides	15 Il surface Il surface n otide io	0 ce protein ce protein n 533 conflict	195£+03 ► n (S-layer n (S-layer Protein Sco	precu precu	V 2 < rsor p rsor p Pepti	prote	ein) [ein) [ons o Σ Γ	Clostridiu Clostridiu fgi 255 ∄ Score	um diffic um diffic 101963 Hits	cile QCD-66 cile QCD-63	c26] q42]	Mass error (p	RT (mins)	Charge	Tag	Abundance	e Conf
gi[1267004 gi[2549739 Protein: g Protein: g tide ion vie Conflicting Accession gi[254976	000 (+9) mi gi[254976 gi[255101 ews Protein g protein 6387 (+5)	15 387 cel 963 cel in resolution is for pep Peptides 63	15 Il surface Il surface n otide io s Unique 18	0 ce protein ce protein n 533 : Conflict 88	1.95±+03 n (S-layer n (S-layer Contemportation Second	precu precu	V 2 < rsor p rsor p Pepti V 2	prote prote ide i	ein) [(ein) [(ons ο Σ [Clostridio Clostridio f gi 255 Score 63.7	um diffic um diffic 101963 Hits 2	Correlation 0.202	c26] q42] Mass 2413.295	Mass error (p -0.243	RT (mins) 55.6	2 Charge 4	Tag	 Abundance 1.25E+06 	e Conf
gi 267004 gi 2549739 Protein: g Protein: g Conflictine Accession gi 254976	000 (+9) ji 2554976 ji 255101 ws Protei g protein 5387 (+5) 1963 (+2)	15 387 cel 963 cel in resolution is for pep Peptides • 63	15 Il surface n otide io s Unique 18	0 ce protein ce protein n 533 : Conflict 88	1.95E+03 n (S-layer n (S-layer Protein Sco 7.07E+03	precu precu		ide i	ein) [(ein) [(ons o Σ [• (Clostridio Clostridio f gi 255 B Score 0 63.7 L 61.3	um diffic um diffic 101963 Hits 2 4	Correlation 0.202 0.015	c26] q42] Mass 2413.295 2413.295	Mass error (p -0.243 0.0288	RT (mins) 55.6 50.5	Charge 4 3	Tag	 Abundance 1.25E+06 3E+05 	e Conf
gil267004 gil2549739 Protein: g Protein: g Conflicting Accession gil254976 gil255101	000 (+9) yij254976 yij255101 wws Protei g protein 6387 (+5) 1963 (+2)	15 387 cel 963 cel in resolution is for pep Peptides • 63 Tur	15 Il surface Il surface otide io s Unique 18 n off all p	0 ce protein ce protein n 533 : Conflict 88 septides	1.95E+03 n (S-layer n (S-layer : Protein Sco 7.07E+03 E+03	precu	Pepti	ide i 533 761	ein) [4 ein) [4 ons ο Σ 4 • (4 ο 2	Clostridiu Clostridiu f gi 255 Score 63.7 61.3 2 123	um diffic um diffic 101963 Hits 2 4 5	Correlation 0.202 0.015 -0.361	c26] q42] Mass 2413.295 2413.295 2413.294	Mass error (p -0.243 0.0288 -0.699	RT (mins) 55.6 50.5 50.5	Charge 4 3 2	Tag	 Abundance 1.25E+06 3E+05 2.84E+05 	e Conf 1 1 1
gi 1267004 gi 2549739 Protein: g Protein: g tide ion vie conflictin; Accession gi 255101	000 (+9) () () () () () () () () () (15 387 cel 963 cel n resolution is for pep Peptides • 63 Tur	15 Il surface Il surface n otide ioo s Unique 18 n off all p	0 ce protein ce protein n 533 : Conflict 88 exptides	1.95E+03	precu precu		ide i 533 761 256	ein) [(ein) [(ons ο Σ [• (0 :	Clostridii Clostridii f gi 255 ∄ Score 0 63.7 L 61.3 2 123 0 122	um diffic um diffic 101963 Hits 2 4 5 5	Correlation 0.202 0.015 -0.361 -0.644	c26] q42] Mass 2413.295 2413.295 2413.295 2413.294 2413.294	Mass error (p -0.243 -0.699 -0_603	RT (mins) 55.6 50.5 50.5 53.1	Charge 4 3 2 2	Tag	 Abundance 1.25E+06 3E+05 2.84E+05 1.67E+05 	e Cont 1 1 1
gil267004 gil2549739 Protein: g Protein: g Conflictin; Accession gil254976 gil255101	000 (+9) m gi[254976 gi[255101 aws Protein g protein 5387 (+5) 1963 (+2)	15 387 cel 963 cel n resolution is for pep Peptides • 63 Tur	15 Il surfac Il surfac n otide io s Unique 18 n off all p	0 ce protein ce protein n 533 : Conflict 8 8 exptides	n (S-layer n (S-layer n (S-layer : Protein Sco 7.07E+03	precu	✓ 2 ✓ Image: second s	421 prote prote ide i \$533 761 2256 4421 2	ein) [(ein) [(ons o Σ [• (0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 :	Clostridii Clostridii f gi 255	um diffic um diffic 101963 Hits 2 4 5 5 5	Correlation 0.202 0.015 -0.361 -0.644 0.427	c26] q42] Mass 2413.295 2413.295 2413.294 2413.294 1076.586	Mass error (p -0.243 0.0288 -0.699 -0.603 -0.603	RT (mins) 55.6 50.5 53.1 30.5	2 Charge 4 3 2 2 2 2	Tag	 Abundance 1.25E+06 3E+05 2.84E+05 1.67E+05 2.49E+07 	e Conf 1 1 1 1 1
gi 1267004 gi 2549739 Protein: g Protein: g Conflicting Accession gi 254976 gi 255101	000 (+9) iii gi[254976 gi[255101 ews Protein g protein 6387 (+5) 1963 (+2)	15 387 cel 963 cel in resolution is for pep Peptides 63 Tur	15 Il surfac Il surfac n otide io s Unique 18 n off all p	0 ce protei ce protei n 533 Conflict 88 veptides	1.95±+03 n (S-layer n (S-layer 2. Protein Scc 7.07E+03 1.E+03	precu precu	2 rsor p rsor p Pepti 	421 prote prote ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ;	ein) [ein) [ein) [οns ο Σ [• (ο ο ο ο ο ο	Clostridia Clostridia f gi 255	um diffic um diffic 101963 Hits 2 4 5 5 5 5 5	Correlation 0.202 0.015 -0.361 -0.644 0.427 0.302	c26] q42] Mass 2413.295 2413.295 2413.294 2413.294 1076.586 1194.628	Mass error (p -0.243 0.0288 -0.603 -0.603 -0.603 -0.603 -0.603 -0.603 -0.603	RT (mins) 55.6 50.5 53.1 30.5 31.5	2 Charge 4 3 2 2 2 2 2 2	Tag	 Abundance 1.25E+06 3E+05 2.49E+07 2.97E+07 	: Conf 1 1 1 1 1 1 1
gil267004 gil2549739 Protein: g Protein: g tide ion vie Conflictin; Accession gil254976 gil255101	000 (+9) iii gi[254976 gi[255101 ews Protein g protein 6387 (+5) 1963 (+2)	15 387 cel 963 cel in resolution is for pep Peptides • 63 Tur	15 Il surface Il surface n ottide io s Unique 18 n off all p	0 ce protei ce protei n 533 Conflict 88 septides E	1.95±+03 n (S-layer n (S-layer 2: Protein Scc 7.07E+03 1.E+03	precu precu	V 2 rsor r rsor r r Pepti # V ! V ! V ! V ! V ! V ! V ! V ! V ! V ! V ! V !	421 prote prote ± 533 761 256 421 2 3 4	ein) [(ein) [(ons ο Σ [] (ο (ο (ο (ο (ο (ο (ο (ο (ο (ο	Clostridia Clostridia f gi 255	um diffic um diffic 101963 Hits 2 4 5 5 5 5 5 5	Correlation 0.202 0.015 -0.361 -0.361 0.427 0.302 0.455	Mass 2413.295 2413.295 2413.295 2413.294 1076.586 1194.628 1246.568	Mass error (p -0.243 0.0228 -0.699 -0.603 F _0194 0.25	RT (mins) 55.6 50.5 53.1 30.5 31.5 28.5	2 Charge 4 3 2 2 2 2 2 2 2 2	Tag	 ✓ Abundance 1.25E+05 2.84E+05 1.67E+05 2.49E+07 2.72E+07 	e Conf 1 1 1 1 1 1 1 1 1 1
gil2549739 gil2549739 Protein: g Protein: g stide ion vie Conflicting Accession gil254976 gil255101	000 (+9) yi[254976 yi[255101 wws Protein 6387 (+5) 1963 (+2)	15 387 cel 963 cel n resolution is for pep Peptider 63 Tur	15 Il surface Il surface n ottide io s Unique 18 n off all p	0 ce protein ce protein n 533 conflict 88 meptides	n (S-layer n (S-layer n (S-layer n (S-layer : Protein Scc 7.07E+03	precu precu	✓ 2 rsor r rsor r rsor r ✓ 1 ✓ 2 ✓ 1 ✓ 2 ✓ 2 ✓ 2 ✓ 2 ✓ 2 ✓ 2	421 prote prote ide i \$ 533 761 2256 421 2 3 4 5	ein) [(ein) [(ons o Σ [[• (• (• (• (• (Clostridie Clostridie f gi 255	um diffic um diffic 101963 Hits 2 4 5 5 5 5 5 5 5 5 5	Correlation 0.202 0.015 -0.361 -0.644 0.425 0.418	Mass 2413.295 2413.295 2413.295 2413.294 1076.586 1194.628 1246.568 1206.661	Mass error (p -0.243 0.0288 -0.699 -0.603 F -0.194 -0.25 -0.0082	RT (mins) 55.6 50.5 53.1 30.5 31.5 28.5 27.3	2 Charge 4 3 2 2 2 2 2 2 2 2 2 2 2	Tag	 Abundance 1.072+03 Abundance 1.25E+06 3E+05 2.84E+05 1.67E+05 2.49E+07 2.97E+07 2.72E+07 2.28E+07 	: Conf 1 1 1 1 1 1 1 1 1 1
gil2549739 gil2549739 Protein: g Protein: g Conflicting Accession gil254976	000 (+9) yi[254976 yi[255101 wws Protein 6387 (+5) 1963 (+2)	15 387 cel 963 cel n resolution is for pep Peptider 63 Tur	15 Il surfac n potide io s Unique 18 n off all p	0 ce protein ce protein n 533 conflict 88 meptides	195E+03 n (S-layer n (S-layer * Protein Scc 7.07E+03 1E+03	precu	V 2 rsor r rsor r Pepti V V 1 V 2 V 1 V 2 V 2 V 2 V 2 V 2 V 2	421 prote prote 533 761 2256 421 2 3 4 5 6	ein) [ι ein) [ι οns ο Σ ξ • (ο ο (ο ο (ο ο (ο ο ο (ο ο ο ο ο ο	f gi 255 ➡ Score 0 63.7 1 61.3 2 123 0 122 0 59.6 1 51.7 5 89.8 0 76.1 2 47.9	um diffic um diffic 101963 Hits 2 4 5 5 5 5 5 5 5 5 5 5 5 5	" " " " " " " " " " " " " " " " " " "	C26] q42] Mass 2413.295 2413.295 2413.294 2413.294 2413.294 1076.586 1194.625 1246.568 1206.661 1250.615	Mass error (p -0.243 0.028 -0.699 -0.603 -0.194 -0.194 -0.25 -0.0082 0.224	RT (mins) 55.6 50.5 53.1 30.5 31.5 28.5 27.3 39.6	2 Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Tag	 Abundance 1.25E+06 3E+05 2.84E+05 1.67E+05 2.49E+07 2.97E+07 2.72E+07 2.28E+07 1.5E+07 	: Conf 1 1 1 1 1 1 1 1 1 1 1 1
) gil267004) gil2549739 Protein: <u>g</u> Protein: <u>g</u> otide ion vie Conflicting Accession <u>9</u> gil254976 <u>9</u> gil255101	000 (+9) gi[254976 gi[255101 aws Protein g protein 6387 (+5) 1963 (+2)	15 387 cel 963 cel in resolution is for pep Peptide: 63 Tur	15 Il surfac Il surfac n otide io s Unique 18 n off all p	0 ce protein ce protein n 533 Conflict 88 meptides	I J95E+03 Torreto Sec Torret	precu	V 2 rsor r rsor r Pepti V V 1 V 2 V V V V V V V V V V V V V V V V V V V V	<pre>ide il prote prote ide il # 5533 761 256 4221 2 3 4 5 6 7</pre>	ein) [' ons ο Σ Φ Ο Ξ <t< td=""><td>f gi 255 ♪ Score 0 63.7 □ 61.3 2 123 □ 122 0 59.6 □ 55.6 89.8 9 □ 76.1 2 47.9 □ 46.4 140.4 140.4</td><td>In diffic In dif</td><td>Correlation 0.202 0.015 0.202 0.015 0.202 0.015 0.202 0.203</td><td>C26] q42] Mass 2413.295 2413.295 2413.294 2413.294 2413.294 1076.586 1206.661 1250.615 922.526</td><td>Mass error (p -0.243 0.0288 -0.699 -0.603 -0.36 F-0194 0.25 -0.0082 0.224 -1.21</td><td>RT (mins) 55.6 50.5 50.5 53.1 30.5 31.5 28.5 27.3 39.6 35</td><td>2 Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2</td><td>Tag</td><td> Abundance 1.25E+06 3E+05 2.49E+07 2.97E+07 2.72E+07 1.5E+07 1.88E+07 </td><td>e Conf 1 1 1 1 1 1 1 1 1 1 1 1 1</td></t<>	f gi 255 ♪ Score 0 63.7 □ 61.3 2 123 □ 122 0 59.6 □ 55.6 89.8 9 □ 76.1 2 47.9 □ 46.4 140.4 140.4	In diffic In dif	Correlation 0.202 0.015 0.202 0.015 0.202 0.015 0.202 0.203	C26] q42] Mass 2413.295 2413.295 2413.294 2413.294 2413.294 1076.586 1206.661 1250.615 922.526	Mass error (p -0.243 0.0288 -0.699 -0.603 -0.36 F-0194 0.25 -0.0082 0.224 -1.21	RT (mins) 55.6 50.5 50.5 53.1 30.5 31.5 28.5 27.3 39.6 35	2 Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Tag	 Abundance 1.25E+06 3E+05 2.49E+07 2.97E+07 2.72E+07 1.5E+07 1.88E+07 	e Conf 1 1 1 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 88 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).

Progenesis Qup 11	l utorial fo	or bork i	rogenesi	a di tot bio														
File Ret nport Data Align	eview gnment	Filterin	Ex ng De	sign Setup	Review Peak Picking	Peptide Statist	Ion ics	Ide Per	entify otides	QC Met	Refine rics Identificat	Review ons Protein	s Statistics	Report			n A	Waters Comp
Experiment design	n: AC	No filter a	applied		Create	Per	otide	ions	of g	i 2549763	87				W No	filter app	plied	Create
Accession		Peptides	Unique	Conflicts	Score		#	Σ	ф	Score	Hits Correlati	n Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	e Confli
qi 254976387 (+	(+5) 🔉	63	63	0	7.07E+03		1		2	75.2	5 -0.624	2413.301	2.48	55.6	3		1.24F+08	0
a gi 255101963 (+	(+2)	0	0	0	0		17		0	128	5 0,464	2413.296	0.265	55.6	2	—	9.38E+07	0
gi 209571234		25	12	18	2.5E+03		533		0	63.7	2 -0.931	2413.295	-0.243	55.6	4	—	1.25E+06	0
) gi 260682215 (+	(+1)	24	11	18	2.08E+03	V	761	0	1	61.3	4 -0.790	2413.295	0.0288	50.5	3		3E+05	0
a::126700407		16	16	0	1.64E+03	V	1256	0	2	123	5 0.844	2413.294	-0.699	50.5	2	<u> </u>	2.84E+05	0
gij120700407					4.055.00			-			F 0.404	2412 204	-0.602	52.1	2	<u> </u>	1.675.05	0
gi[125700407] gi[254973900 (+ Protein: gi[25 Protein: gi[25	(+9) 	15 887 cel 963 cel	15 I surfac I surfac	o ce protein ce protein	n (S-layer p n (S-layer p		2421 or pro or pro	tein) tein)	0 [Clo [Clo	ostridium (difficile QCD	66c26] 63q42]	-0.003		L	-	1.072+03	•
gij254973900 (+ Protein: gij25 Protein: gij	(+9) 549763 551019 Protein proteins	15 887 cel 963 cel resolution for pep	15 Il surfac Il surfac n otide ior	e protei e protei	195E+03 ▶ n (S-layer p	precurso precurso Pe	2421 or pro or pro	tein) tein) teins	0 [Clo [Clo	ostridium ostridium gi 255101	difficile QCD difficile QCD 963	66c26] 63q42]	-0.005		L	-	1.072+03	4
gij254973900 (+ Protein: gij25 Protein: gij25 Protein: gij25 conflicting pr Accession	(+9) 	15 887 cel 963 cel resolution for pep Peptides	15 Il surfac Il surfac n otide ior	e protein e protein n 533 Conflict	n (S-layer p n (S-layer p n (S-layer p	Precurso Precurso Perecurso	2421 or pro or pro ptide #	tein) tein) tons	0 [Clo [Clo 0 [Clo	122 ostridium (ostridium (gi 255101 Score	difficile QCD difficile QCD difficile QCD 963 Hits Correlati	66c26] 63q42]	Mass error (p	RT (mins)	Charge	Tag	Abundance	e Conf
gij254973900 (+ Protein: gij25 Protein: gij25 Prote	(+9) 2549763 2551019 Protein proteins 7 (+5) •	15 387 cel 963 cel resolution for pep Peptides 63	15 Il surfac Il surfac n otide ion s Unique 63	e protein e protein n 533 Conflict	n (S-layer p n (S-layer p n (S-layer p : Protein Scor 7.07E+03		2421 or pro or pro ptide # 533	tein) tein) ions Σ	0 [Clo [Clo 0 0	122 ostridium (ostridium (gi 255101 ¹ Score 63.7	difficile QCD difficile QCD 2063 Hits Correlati 2	2413.294 66c26] 63q42] n Mass 2413.295	-0.003 Mass error (p -0.243	RT (mins) 55.6	Charge 4	Tag	 Abundance 1.25E+06 	e Conf
gij254973900 (+ Protein: gij25 Protein: gij25 ptide ion views Conflicting pr Accession gij254976387 gij255101963	(+9) 111 25549763 2551019 Protein proteins 7 (+5) 9 8 (+2)	15 387 cel 963 cel resolution for pep Peptides 63 0	15 I surfac I surfac n otide ion 63 0	e protein e protein n 533 Conflict 0	n (S-layer p n (S-layer p n (S-layer p r (S-layer p 2 Protein Scor 7.07E+03	Precurso precurso Pe re	2421 or pro or pro ptide # 533 761	tein) tein) ions Σ	0 [Clo [Clo 0 0 1	122 ostridium o ostridium o gi 255101 ⁴ Score 63.7 61.3	difficile QCD difficile QCD difficile QCD P63 Hits Correlati 2 4	2413.294 66c26] 63q42] n Mass 2413.295 2413.295 2413.295	-0.035 Mass error (p -0.243 0.0288	87 (mins) 55.6 50.5	Charge 4 3	Tag	 Abundanc 1.25E+06 3E+05 	e Conf
gij254973900 (+ Protein: gij25 Protein: gij25 Protein: gij25 ptide ion views Conflicting pr Accession gij254976387 gij255101963	(+9) 111 2549763 2551019 Protein proteins 7 (+5) • 8 (+2)	15 887 cel 963 cel resolution for pep Peptides 63 0	15 Il surfac Il surfac n otide ion 63 0	o ee protein n 533 Conflict 0	I.SEE+03 I.SEE+03 I.SEE+03 Protein Score 7.07E+03 0	Perecurso	2421 or pro ptide # 533 761 1256	tein) tein) ions	0 [Clo [Clo of s 0 1 2	122 ostridium (ostridium (gi 2551011 Score 63.7 61.3 123	difficile QCD difficile QCD difficile QCD P63 Hits Correlati 2 4 5	2413.294 66c26] 63q42] n Mass 2413.295 2413.295 2413.295 2413.295 2413.295 2413.295 2413.295	-0.003 Mass error (p -0.243 0.0288 -0.699	RT (mins) 55.6 50.5 50.5	2 Charge 4 3 2	Tag	 Abundanc 1.25E+06 3E+05 2.84E+05 	e Conf 0 0 0
gij2570407 gij2570407 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25	(+9) #549763 2551019 Protein proteins 7 (+5) • 8 (+2)	15 387 cel 363 cel resolution for pep Peptides 63 0	15 Il surfac Il surfac n otide ion 5 Unique 63 0	e protein e protein n 533 Conflict 0	n (S-layer p n (S-layer p n (S-layer p 2 Protein Scor 7.07E+03	Peere	2421 or pro ptide # 533 761 1256 2421	tein) tein) ions	0 (Clo	122 ostridium (ostridium (gi 255101) Score 63.7 61.3 123 122	Hits Correlati 2 5	n Mass 2413.294 2413.294 n Mass 2413.295 2413.295 2413.294 2413.294 2413.294 2413.294	-0.003 Mass error (p -0.243 -0.228 -0.639 -0.603	RT (mins) 55.6 50.5 50.5 53.1	2 Charge 4 3 2 2 2	Tag	 Abundanc 1.25E+06 3E+05 2.84E+05 1.67E+05 	e Conf 0 0 0
gij254973900 (« Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 Protein: gij25 ptide ion views Conflicting pr Accession gij254976387 gij255101963	(+9) #549763 2551019 Protein proteins 7 (+5) • 8 (+2)	15 387 cel 363 cel resolution for pep Peptides 63 0	15 I surfac I surfac n otide ion 5 Unique 63 0	e protein e protein n 533 Conflict 0	n (S-layer p n (S-layer p · Protein Scor 7.07E+03	Peter	2421 or pro ptide # 533 761 1256 2421 2	tein) tein) ions	0 (Clo) [Clo) [Clo 0 1 2 0 0	122 ostridium (ostridium (gi 255101) Score 63.7 61.3 123 122 59.6	3 0.494 III IIII Iifficile QCD IIII Iifficile QCD IIII 063 IIIIII 2 4 5 5 5	2413.294 666c26] 63q42] n Mass 2413.295 2413.295 2413.294 2413.294 1076.586	-0.003 Mass error (p -0.243 0.0288 -0.699 -0.603 -0.36	RT (mins) 55.6 50.5 53.1 30.5	2 Charge 4 3 2 2 2 2	Tag	 Abundanc 1.25E+06 3E+05 2.84E+05 1.67E+05 2.49E+07 	e Conf 0 0 0 0 0
gi[254700407 gi]254973900 (« Protein: gi]25 Protein: gi]25 Protein: gi]25 ptide ion views Conflicting pr Accession ♀ gi]254976387 ♀ gi]255101963	(+9) 111 15549763 1551019 Protein proteins 17 (+5) • 3 (+2)	15 887 cel 963 cel resolution for pep Peptides 63 0	15 I surfac n otide ion 63 0	te protein te protein n 533 Conflict 0	n (S-layer p n (S-layer p n (S-layer p v Protein Scot	Perecurso	2421 or pro ptide # 533 761 1256 2421 2 3	tein) tein) ions	0 (Clo) (Clo)	122 postridium (pstridium (gi 255101' Score 63.7 61.3 123 122 59.6 51.7	S 0.494 III IIII difficile QCD QCD P63 IIII 2 4 5 5 5 5 5	n Mass 2413.294 666c26] 63q42] 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 1076.586 1194.628	Mass error (p -0.243 0.0288 -0.699 -0.603 -0.36 -0.194	RT (mins) 55.6 50.5 53.1 30.5 31.5	2 Charge 4 3 2 2 2 2 2 2	Tag	 Abundanc 1.25E+06 3E+05 2.84E+05 1.67E+05 2.49E+07 2.97E+07 	e Conf 0 0 0 0 0 0 0 0 0 0 0 0
gi[254973900 (+ gi[254973900 (+ Protein: gi[25 Protein: gi[25 ptide ion views Conflicting pr Accession ♥ gi[254976387 ♥ gi[255101963	(+9) 11 2549763 2551019 Protein proteins 7 (+5) 9 8 (+2)	15 387 cel 963 cel resolution for pep Peptides 63 0	15 Il surfac Il surfac n otide ion 63 0	te protein te protein n 533 Conflict 0 0	ISSE403 Protein Score Protein Score O	Peere	2421 pr pro pr pro ptide # 533 761 1256 2421 2 3 4	tein) tein) ions	0 (Cloid)	122 postridium (pstridium (23) (255101) Score 63.7 61.3 123 122 59.6 51.7 89.8	S 0.494 III III difficile QCD QCO P663 IIII 2 5 5 5 5 5 5 5 5	n Mass 2413.294 666c26] 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294	-0.003 -0.243 -0.699 -0.603 -0.36 -0.194 -0.25	RT (mins) 55.6 50.5 53.1 30.5 31.5 28.5	2 Charge 4 3 2 2 2 2 2 2 2 2 2	Tag	 Abundanc 1.25E+06 3.E+05 2.84E+05 1.67E+05 2.49E+07 2.72E+07 	e Conf 0 0 0 0 0 0 0 0 0 0 0 0 0 0
gij2507300 (+ gij2507300 (+ Protein: gij25 Protein:	(+9) 11 2549763 2551019 Protein proteins 7 (+5) 9 8 (+2)	15 387 cel 963 cel resolution for pep Peptides 63 0	15 Il surfac Il surfac n otide ion 63 0	e protein e protein n 533 Conflict 0	195±+03 + n (S-layer p n (S-layer p n (S-layer p 7.07±+03 0	Peere	2421 pr pro pr pro ptide # 533 761 1256 2421 2 3 4 5	tein) tein) ions	0 (Clo	122 postridium (postridium (gi 255101) Score (63.7 61.3 123 122 59.6 51.7 89.8 76.1	S O(A94 "" "" "" "" Officile QCD QCD Officile QCD "" Officile QCD ""	n Mass 666c26] 63q42] 2413.295 2413.295 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294	-0.243 0.0228 -0.699 -0.603 -0.36 -0.194 0.25 -0.0082	RT (mins) 55.6 50.5 53.1 30.5 31.5 28.5 27.3	2 Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2	Tag	 Abundanc 1.25E+06 3E+05 2.84E+05 1.67E+05 2.49E+07 2.97E+07 2.72E+07 2.28E+07 	e Conf 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
g gj2500000 gj2507300 (c Protein: gj25 Protein: gj25 ptide ion views Conflicting pr Accession @ gj254796387 @ gj255101963	(+9) #5549763 #551019 Protein proteins 7 (+5) 9 8 (+2)	15 387 cel 963 cel resolution for pep Peptides 63 0	15 Il surfac Il surfac n otide ion 63 0	e protein e protein n 533 Conflict 0	195±+03 + n (S-layer p n (S-layer p n (S-layer p 7.07±+03 0	Pere	2421 pr pro pr pro ptide # 533 761 1256 2421 2 3 4 5 6	tein) tein) ions Σ	0 (Clo	122 postridium (postridium (gi 255101) Score 63.7 61.3 123 122 59.6 51.7 89.8 76.1 47.9	S U.494 III IIII Jifficile QCD Correlati J63 IIII 2 4 5 5 5 5 5 5 5 5 5 5 5 5	n Mass 2413.294 2413.295 2413.295 2413.295 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294 2413.294	0.243 0.0288 -0.699 -0.603 -0.36 -0.194 0.25 -0.0082 0.244	RT (mins) 55.6 50.5 53.1 30.5 31.5 28.5 27.3 39.6	2 Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Tag	 Abundanc 1.25E+06 3E+05 2.84E+05 1.67E+05 2.49E+07 2.97E+07 2.72E+07 2.88E+07 1.5E+07 	e Conf 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
gijzbiordov gijzbioradov Protein: gij2 Protein: gij2 epide ion views Conflicting pr Accession gij25976387	(+9) #2549763 2551019 Protein proteins 7 (+5) • 8 (+2)	15 387 cel 387 cel 387 cel 7 7 7 7 7 7 7 7	15 I surfac I surfac n tide ion 63 0	e protein e protein n 533 Conflict 0 0	195±403 + n (S-layer p n (S-layer p 2 Protein Scor 7.07±403 0	Pere	2421 pr pro pr pro ptide # 533 761 1256 2421 2 3 4 5 6 7	tein) ions Σ	0 (Clo	122 postridium o sostridium o gi 2551011 Score 63.7 63.7 123 122 59.6 51.7 89.8 76.1 47.9 46.4	S OLAP4 III IIII D63 IIII 2 4 5 5 5 5 5 5 5 5 5 5 5	66c26] 63q42] n Mass 2413.295 2413.295 2413.294 2413.294 2413.294 2413.294 2413.294 1076.586 1194.628 1206.655 1225.655 922.526	0.003 Mass error (p -0.243 0.0288 -0.603 -0.36 -0.194 0.25 -0.0082 0.244 -1.21	RT (mins) 55.6 50.5 53.1 30.5 31.5 28.5 27.3 39.6 35	2 Charge 4 3 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Tag	 Abundanc 1.25E+06 3E+05 2.84E+05 2.49E+07 2.97E+07 2.22E+07 2.28E+07 1.5E+07 1.88E+07 	e Conf 0 0 0 0 0 0 0 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 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 18 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	•		#	Σ	몓	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge Tag	✓ Abundance	Conflict: 🔦
🎯 gi 126697969 (+1)	10	10	0	925			3657	0	1	67.4	2	0.900	1791.962	-0.00909	52.6	2 🧧	1.06E+05	1 =
🎯 gi 126697970	9	9	0	748		V	3098	0	1	63	4	0.991	1198.656	0.582	29.1	2 📒	8.84E+04	1
🎯 gi 254976383 (+2) 🜼	9	7	2	711		V	1415	0	0	68.7	3	0.846	1268.666	0.282	45.7	2 📒	1.17E+05	0
🎯 gi 255101959	0	0	0	0		V	892	٠	2	110	4	0.821	1889.897	0.0366	25.8	2 📒	2.5E+05	0
🎯 gi 126699128	7	4	5	860	-	V	680	٠	0	71.3	5	0.966	1319.646	-0.666	30.7	2 🧧	1.59E+05	- 0
•					.P.	•												•
1 Protein: gi 254976	383 ce	ll surfac	e proteir	n [Clostr	idium d	liffic	ile Q(CD-	66c2	26]								
Protein: gi 101802	05 Cwp	66 [[Clo	stridium] difficile	el													
Peptide ion views Protei	n resolutio	n																
Conflicting protein	s for pe	ptide ior	n 3657			Pep	otide	ions	of	gi 1018	0205							
Accession	Peptide	s Unique	Conflict	Protein S	core		#	Σ	中	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge Tag	 Abundance 	Conflict
🕥 gi 254976383 (+2)	<u> </u>	7	2	711		V	3098	•	1	63	4	0.242	1198.656	0.582	29.1	2 🦷	8.84E+04	1
🔮 gi 255101959	0	0	0	0		V	3657	٠	1	67.4	2	0.376	1791.962	-0.00909	52.6	2 🧧	1.06E+05	1
🕥 gi 10180205 (+1)	3	1	2	266 🗸		V	6071	٠	2	136	4	0.367	2319.16	0.126	60.1	2 📒	5.91E+04	0
						V	11170	٠	1			0.401	2319.10	0.184	60.1	3 💼	9.75E+03	0
< III					*	•	_	_	_									•

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	中	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Conflict: 🔶
🎯 gi 126697969 (+1)	10	10	0	925		V	3657	0	1	67.4	2	0.900	1791.962	-0.00909	52.6	2		1.06E+05	0 =
🎯 gi 126697970	9	9	0	748		V	3098	0	1	63	4	0.991	1198.656	0.582	29.1	2		8.84E+04	0
🎯 gi 254976383 (+2) 🛛	9	9	0	711		V	1415	0	0	68.7	3	0.846	1268.666	0.282	45.7	2		1.17E+05	0
🎯 gi 255101959	0	0	0	0		1	892	٠	2	110	4	0.821	1889.897	0.0366	25.8	2		2.5E+05	0
🎯 gi 126699128	7	4	5	860	-	1	680	٠	0	71.3	5	0.966	1319.646	-0.666	30.7	2		1.59E+05	0 -
•				Þ		•													P
T Protein: gi[254976																			
Protein: gi[254976 Protein: gi[101802 Peptide ion views Prote Conflicting protein	205 Cwp in resolutions for pe	n [[Clo	stridium] difficile]		Per	otide i	ions	of	gi 1018	0205								
Protein: gi[254976 Protein: gi[101802 Peptide ion views Prote Conflicting protein Accession	205 Cwp in resolutions for per Peptide	o 66 [[Clo n ptide ior s Unique	1 3098 Conflict:] difficile] Protein Sco	ore	Pep	otide i #	ions Σ	of ⊕	gi 1018 Score	0205 Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	✓ Abundance	Conflict
 Protein: gi/254976 Protein: gi/101802 Peptide ion views Prote Conflicting protein Accession gi/254976383 (+2) 	205 Cwp in resolutions for per Peptide	p 66 [[Clo n ptide ior s Unique 9	1 3098 Conflict:] difficile] Protein Sco 711	ore	Peŗ	tide i # 3098	ions Σ ○	of 1	gi 1018 Score 63	0205 Hits 4	Correlation	Mass 1198.656	Mass error (p 0.582	RT (mins) 29.1	Charge 2	Tag	 Abundance 8.84E+04 	Conflict 0
 Protein: gi[254976 Protein: gi[101802 Peptide ion views Prote Conflicting protein Accession gi[254976383 (+2) gi[10180205 (+1) 	205 Cwp in resolutions for pe Peptide 9 1	ptide ior s Unique 9	1 3098 Conflict: 0	Protein Sco 711 136	ore	Per	tide i # 3098 3657	ions Σ ο	of 1	gi 1018 Score 63 67.4	0205 Hits 4 2	Correlation	Mass 1198.656 1791.962	Mass error (p 0.582 -0.00909	RT (mins) 29.1 52.6	Charge 2 2	Tag	 Abundance 8.84E+04 1.06E+05 	Conflict 0 0
 Protein: gi[254976 Protein: gi[101802 Peptide ion views Prote Conflicting protein Accession gi[254976383 (-2) gi[10180205 (+1) 	205 Cwp in resolutions for pep Peptide 9 1	ptide ior s Unique 9 1	1 3098 Conflict: 0	Protein Sco 711 136	ore	Per	# 3098 3657 6071	ions Σ ο	of 1 1 2	gi 1018 Score 63 67.4 136	0205 Hits 4 2 4	Correlation 1.000	Mass 1198.656 1791.962 2319.16	Mass error (p 0.582 -0.00909 0.126	RT (mins) 29.1 52.6 60.1	Charge 2 2 2	Tag	 ✓ Abundance 8.84E+04 1.06E+05 5.91E+04 	Conflict 0 0 0
 Protein: gi[254976 Protein: gi[101802 Peptide ion views Prote Conflicting protein Accession gi[254976383 (+2) gi[10180205 (+1) 	205 Cwp in resolutions for pep Peptide 9 1	ptide ior s Unique 9 1	1 3098 Conflict: 0	Protein Sco 711 136	ore	Peŗ	# 3098 3657 6071 11170	ions Ω Ο	of 1 1 2 1	gi 1018 Score 63 67.4 136 	0205 Hits 4 2 4 	Correlation 1.000 1.000	Mass 1198.656 1791.962 2319.16 2319.16	Mass error (p 0.582 -0.00909 0.126 0.184	RT (mins) 29.1 52.6 60.1 60.1	Charge 2 2 2 3	Tag Tag	 Abundance 8.84E+04 1.06E+05 5.91E+04 9.75E+03 	Conflict 0 0 0

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

Accession	Peptides	Unique	Conflict	s Score	*		#	Σ	Φ	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Conflict:	^
🌍 gi 126700070	1	1	0	58.4			717	٠	0	67.6	4	0.962	2002.074	0.0596	51.1	3		2.73E+05	0	=
🎯 gi 126701092	1	1	0	119			159	0	1	138	5		1489.803	0.0648	51.2	2		1.96E+06	0	
🎯 gi 126698464	1	1	0	78.3			1020	0	1	59.6	3		1489.802	-0.816	51.2	3		1.13E+05	0	
🎯 gi 126697826	1	1	0	106		7	1956	٠	1	129	5	0.900	2002.073	-0.146	51.1	2		1.28E+05	0	
🎯 gi 126699793	1	1	0	63.9	-		165	0	0	64.4	5		902.533	-0.538	42.2	2		5.56E+05	0	-
<				•		•													÷	
1 Protein: gi 12669	9128 ru	brerythr	in [Clos	tridium diffi	cile 6	30]														
Protein: gi 12669	9078 ru	brerythr	in (Rr) [Clostridium	diffic	ile (630]													
Peptide ion views Prot	ein resoluti	on																		
Conflicting protei	ins for pe	eptide io	n 165			Pep	tide	ions	of	gi 1266	99078									
Accession	Peptides U	nique C	onflict: P	rotein Score			#	Σ	Φ	Score	Hits	Correlation	Mass	Mass error (p	RT (mins)	Charge	Tag	 Abundance 	Conf	
🎯 gi 126699128 🌼	5	5	0 59	9			165	•	0	64.4	5	0.928	902.533	-0.538	42.2	2		5.56E+05	0	E
🕥 gi 126699078	4	4	0 45	6	V	V	1020	٠	1	59.6	3	0.983	1489.802	-0.816	51.2	3		1.13E+05	0	
						7	159	٠	1	138	5	0.987	1489.803	0.0648	51.2	2		1.96E+06	0	
						V	250	٠	1	96.9	5	0.870	1696.905	2.95	51.6	2		1.15E+06	0	-
٠ III					- F	•						III							•	

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

<	۴	•
Refine Identifications Protein options	Recalculate al	bundances

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.



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 <u>IPA website</u>.

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.



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

IP 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 🔹
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 positive fold change.
< Back Export proteins to IPA Close

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 ×	
<u>File E</u> dit <u>V</u> iew F <u>a</u>	vorites <u>T</u> ools <u>H</u> elp	
INGEI	NUITY	
Welcome! F	lease login	Contact Customer Support
Email	smith@work.com	Customer Support
Password	•••••	Houre: 650.551.5111 Houre: 6am - 5pm (PST)
	Remember my password	support@ingenuity.com
	LOG IN	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.

IPA Edi	it View Window	Help								Drouide Feedback I Support Japurz Nikial	
_ <u>_</u> u	<u>. view <u></u>indow</u>	Ticib								Howae recable k juppon Sanacz wykier	Close I
			Ger	ies and Chemicals	Diseases and Fund	tions Pathways ar	nd Tox Lists				
EW ¥			Ente	r gene names/symbol	ls/IDs or chemical/o	rug names here			SEA	Advanced Search	0000
									200		QIAG
tacot l	Inload - New Dataset	2016-05-	20 10-20 46							Dataset Lipload Workflow Instructions	
	spioud rich bataset	1010 05		•						Paraset opious mortinov and actions	
. Select	t File Format:		Flexible For	mat	▼ Mor	e Info				Data Upload Workflow	
. Conta	ains Column Header:		🔿 Yes 🌘	No						Lice Detroit Lieland to impart your detroit file into IDA	
. Select	t Identifier Type:		UniProt/Sw	iss-Prot Accession	▼ Spec	ify the identifier type f	found in the dataset			Once uploaded, many different analysis options exist includ	ing
Array	nlatform used for evo	eriments:	Not specifi	ed/applicable	▼ Sele	t relevant array platfo	rm as a reference se	t for da	ita analysis	Biomarker Filter, Molecular Tox and Core Analyses. Review	
Array	plation dised for exp	enments:	Not specifi	co, applicable	• Sele	creievant array platro	in as a reference se	cror ua	1.4 41141315.	the different type of analyses and see which one best fits	
. Use tl	he dropdown menus t	o specify th	he column na	ames that contain ide	ntifiers and observa	tions. For observations	s, select the appropr	iate exp	pression value typ	your needs.	
										 To upload a dataset file, <u>click here</u>. 	
Raw D	ata (512) \ Dataset Sur	nmary (1)	/								
	1				lr-				[177 Open	
	ID 🔻	Observ	ation 1 🔹	Ignore 🔻	Ignore	▼ Ignore	▼ Ignore	-	Ignore	Look In: 🗀 Multple Rank. 👻 🕼 🏠 🎬 🐯 🕃	
		Exp Fol	d Chan 🔻							Batch (mult-timepoint)	
1	D24721	1 000021	55280087							L_with_Fold and_Normalized.txt	
2	C4YOR7	1.0361887	74042727								
3	C4YMC3	-1.163043	36093987								
4	C4YON7	-1.142480	88552437								
5	P43098	-1.001672	260838215								
6	C4YR46	1.1114419	95146305								
7	013430	1.0005161	1363124							File Name:	
8	P46587	1.1669751	10036399							Files of Type: All Files	•
9	P82610	1.1414367	75382931							Onen Conce	a
10	P46598	1.0600492	23965961							Card	
11	Q96VB9	-1.106136	536924535								
12	P41797	1.5074757	73107542							 Select the dataset file from your computer and click the 	
13	P28877	-1.206248	862059398							Open button.	
14	C4YK39	1.0549466	51189734							2. Calact Elavible format for the file format from the	
15	Q59KZ1	1.0399981	13845519							drondown menu	
16	C4YL05	-1.031227	70752554							uropuown menu.	
17	P46273	1.3116598	32609058							4. Select an Identifier Type from the dropdown menu. IPA	
18	C4YIL8	1.2690949	93012736							supports many identifiers and symbols and will attempt to	•
19	O94039	1.5532525	57659893							guess at the type of identifier in your dataset file. To	
20	013287	1.3853552	24141575							override the selection, uncheck the option and simply sele	ct
			_			3333333				the most appropriate one. If more than one type of identi	fier
										exists in your dataset, select all appropriate ones.	
										GenBank 💌	

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 11: Exporting protein identities to MetaCore

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

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 MetaCore

QIP Progenesis QI.p Tutorial for DDA - Progenesis QI for prot	eomics				
File Review Experiment Import Data Alignment Filtering Oston	Review Peak Peptide Ion Identify Refine Picking Statistics Peptides QC Metrics Identification	ns Proteins Statistics	Report		A Waters Company
Review Proteins Using this screen, you can find the proteins of interest in your experiment.	No filter applied Create	Q			
Set the quantitation options	Accession Peptides Unique peptides Confidence sci	ore Anova (p) q Value 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.	OIP Export Pathways Information		3.03	C A	60 kDa chaperonin [Clost *
protein grouping and more.			9.65	C A	30S ribosomal protein S1
Protein options	Choose a pathways tool from the list below. You can find out m	ore or download new	2.79	C A	cell wall-binding protein
	plugins using the links below.		5.24	A C	thioredoxin reductase 3 [
2 Create a shortlist to review	Which pathways tool do you want to use?		5.1	C A	RNA-binding protein Hfq
on their measurements, to generate a shortlist	MetaCore	-	2.45	A C	30S ribosomal protein S8
for further review.	About this plugin	 Download other pluging 	3.34	A C	dinitrogenase iron-molyb
To a title title has a size and a size a fit	- About this plagm	- Download other plagins	11.5	A C	peptidyl-prolyl cis-trans i
the relevant column header.			3.18	A C	phosphate butyryltransfe
Deview the proteins			9.14	A C	phosphatase, 2C family [(
3 For each protein of interest, review its peptide			44	A C	cell surface protein [Clost
measurements and correlations:			3.45	C A	F0F1 ATP synthase subun
View peptide measurements			1.78	C A	electron transfer flavopro +
You can also double-click to review a protein.					•
4 Export data for further processing By exporting your data to external tools, there's no limit to your analysis.			in) [Clostridium	difficile QCD-66c2	6]
Export to pathways tool	< Back	Next > Cancel		C	
Export protein measurements Export peptide measurements Export peptide ion measurements	1955 1955 1957 1978 1979 1979 1979 1970 1970 1970 1970 1970			1\$1	
Experiment design Review your data from a different perspective:	4				÷
Current design: AC	Quantifiable proteins displayed: 145				Section Complete $ e e e e e e e e e e e e e $

Then click Next to Configure your export page .

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

OP 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:
Swiss-Prot Accession 🔹
Swiss-Prot Accession
GI Number ENSEMBL
EMBL
comparison with the selected control.
< Back Export proteins to MetaCore Close

If you have multiple conditions in your current experiment design you can select which condition should be used as the control

Finally click **Export proteins to MetaCore**.

If you have not already done so, you will be prompted to enter your MetaCore *User name* and *Password*; once you have done this, your data will be exported to your MetaCore account, and the MetaCore user interface will open in your default web-browser.

Windows Securit	y 💌
Log in to N Please enter yo Once you have	letaCore our username and password for your MetaCore installation. e logged in, your data will be exported to your user account.
	User name Password
	OK Cancel

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

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

Appendix 12A 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:

r					
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 🗸		
	NCBInr	Allow up to	1 V missed cleavages		
	NCBInr_UserGuide	Ouantitation	None	\checkmark	
Tavanamu	Eirmieutee (gram positive bacteri	ia)			
Taxonomy	Firmicules (gram-positive bacteri	ia)	•		
Fixed modifications	Carbamidomethyl (C)	> <	Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term) Amidated (Protein C-term)	^	
Variable modifications	Display all modifications	> <	Ammonia-loss (N-term C) 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		
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 12B Use Additional Export Fields in Mascot for PRIDE

For exporting mzIdentML 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	\checkmark	
Score	\checkmark	
Description*	\checkmark	
Mass (Da) [*]		
Number of queries matched		
Percent coverage**		
Length in residues**		
pI ^{**}		
Taxonomy**		
Taxonomy ID**		
Protein sequence**		
emPAI		
$\ ^{*}$ Occasionally requires information to be retrieved fr	om 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)		
Mass error (Da)		
Start		
End		
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