

Progenesis LC-MS User Guide

Analysis workflow guidelines

for version 4.1



Contents

Introduction	3
How to use this document	3
How can I analyse my own runs using LC-MS?	3
LC-MS Data used in this user guide	3
Workflow approach to LC-MS run analysis	4
Restoring the LC-MS Tutorial	5
Stage 1: Import Data and QC review of LC-MS data set	6
Stage 2: Automatic Alignment of your data	7
Stage 3: Licensing	9
Stage 4: Review Alignment	9
Stage 5A: Filtering1	2
Stage 5B: Reviewing Normalisation1	6
Stage 6: Experiment Design Setup for Analysed Runs1	9
Stage 7: Review Peak Picking and editing of results2	22
Stage 8: Peptide Statistics on selected features	61
Stage 9: Identify peptides	5
Stage 10: Refine Identifications	9
Stage 11: Resolve Conflicts4	0
Stage 12: Review Proteins4	5
Stage 13: Protein Statistics4	8
Stage 14: Reporting4	9
Creating an Inclusion list5	52
Congratulations!5	3
Appendix 1: Stage 1 Data Import and QC review of LC-MS data set5	64
Appendix 2: Stage 1 Data QC review and addition of exclusion areas5	9
Appendix 3: Licensing runs (Stage 3)6	0
Appendix 4: Manual assistance of Alignment6	51
Appendix 5: Within-subject Design6	6
Appendix 6: Power Analysis (Peptide Stats)6	8
Appendix 7 Search engine parameters (Stage 9) Mascot6	9

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 LC-MS workflow. It starts with LC-MS data file loading then Alignment, followed by Analysis that creates a list of interesting features (peptides) which are explored within Progenesis 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. However, the document covers all the stages in the LC-MS workflow, therefore if you are using your own data files please refer to Appendix 1 (page 54) then start at page 6.

How to use this document

You can print this user guide to help you work hands-on with the software. The complete user guide takes about 50 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 LC-MS?

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 LC-MS. Instructions on how to do this are included in a section at the end of the user guide document. Alternatively if you would like to arrange a demonstration in your own laboratory contact <u>support@nonlinear.com</u> and we will help you.

LC-MS Data used in this user guide

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

Workflow approach to LC-MS run analysis

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

	Review		Experiment	Review Peak	Peptide	Identify	Refine	Resolve	Review	Protein	
Import Data	Alignment	Filtering	Design Setup	Picking	Statistics	Peptides	Identifications	Conflicts	Proteins	Statistics	Report
											-

Stage	Description	Page
Import Data	LC-MS Import Data: Selection and review of data files for analysis.	5
	Automatic Alignment: Automatic Reference selection and alignment	7
Licensing	Licensing : allows licensing of individual data files when there is no dongle attached (Appendix 3)	9
Review Alignment	Review Alignment: automatic and manual run alignment	9
Filtering	Filtering : defining filters for peaks based on Retention Time, m/z , Charge and Number of Isotopes.	13
	Review Normalisation: explains LC-MS normalisation	16
Experiment Design Setup	Experiment Design Setup : defining one or more group set ups for analysed aligned runs	19
Review Peak Picking	Review Peak Picking : review and validate results, edit peak detection, tag groups of peaks and select peaks for further analysis	22
Peptide Statistics	Peptide Statistics : performing multivariate statistical analysis on tagged and selected groups of peptides	31
Identify Peptides	Identify Peptides: managing export of MS/MS spectra to, and import of peptide ids from Peptide Search engines	35
Refine Identifications	Refine Identifications: manage peptide ids and filters	39
Resolve Conflicts	Resolve Conflicts : validation and resolution of peptide id conflicts for data entered from Database Search engines	40
Review Proteins	Review proteins: review protein and peptide identity	45
Protein Statistics	Protein Statistics: multivariate statistical analysis on proteins	48
Report	Report: generate a report for proteins and/or peptides	49

Restoring the LC-MS Tutorial

Open Progenesis LC-MS and download the Compressed (.zip) Tutorial Archive file from the the 'View online 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 LC-MS tutorial archive file. To do this, first locate the LC-MS Tutorial Archive file using the **Open** button and press Open.

Progenesis LC-MS		
File		
Experiments		nonlinear
Perform analysis	Combine analysed fractions	Getting started with Progenesis Here are some resources to help you
Recent experime	nts	Search P Here are some resources to help you get started with Progenesis LC-MS: • Data analysis with Progenesis LC-MS: • data comparison of the progenesis LC-MS: • data comparison of th
	CC-MS Tutorial.Progenesis.LCmA Progenesis.LC-MS Archive File name LC-MS Tutori	rchive Date modified: 19/02/2013 11:30 Size: 280 MB dLProgenesisLcmsAr ▼ Experiments and Archives (".Pr. ▼ Open ▼ Cancel
Other experimen	ts Open	A Last Lan Tear Tear Providence and the Progenesis CoMet v2.0 released for LC-M5-based metabolomics 41.4797.36561 Happy Holidays from the Nonlinear

This opens the 'Import from archive' dialog.

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



nport fron	port from archive					
Import	LC-MS Tutorial from archive					
💿 Rep	lace an existing experiment					
Experim	nent to replace:					
Orea Name:	te a new experiment LC-MS Tutorial					
Folder:	$\label{eq:c:Users} C:\label{eq:c:Users} C:$					
		Import Cancel				

Then press Import.

Restoring tutorial	

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 LC-MS tutorial will now open at the LC-MS Import Data stage (see below).



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

Note: the **Experiment Properties** are available from the File menu. These were selected when the experiment was created (see Appendix 1, page 54).

Tip: the **'Exclude areas from selected run'** facility allows you to examine and exclude areas (usually early and/or late in the LC dimension (Retention Time)) that appear excessively noisy due to capture of data during column regeneration (see Appendix 2, page 59). This is not required for this data set.



Experiment Properties

Runs in this experiment: 6

Machine resolution: High resolution Peak processing: Profile data X

Close

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

Now start the Alignment process

Stage 2: Automatic Alignment of your data

The process of alignment of your data can be started before the import of your data has been completed by clicking on **Start alignment process.**

2	Start the alignment process While your runs are importing, click the button below to:					
	:	Select <u>alignment re</u> Determine the best Align all runs to tha	ference candidates of the candidates It reference run			
			Start alignment process			

During this process the software will Align all your runs to a Reference run which can either be selected automatically by the software or manually selected by you.

In this tutorial example you have 6 runs which have already been imported so to start the process click on **Start alignment process**

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

Progenesis LC-MS provides three methods for choosing the alignment reference run, as seen below:

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

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.

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.

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 54 for more details on using the other options).

You will now be asked if you want to Align your runs automatically or manually.

▲ Start Alignment 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 No, I'll align my runs manually
< Back Finish Cancel

Select automatically and click finish.

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

Alignment Processing (7%)	•••				
Alignment processing Current step: Choosing an alignment reference	.e				
 ✓ Imported runs: 6/6 ⑦ Reference run: pending 				1	
Aligned runs: pending	Alignment Processing (80%)		×		
	Imported runs: 6/6				
	Aligned runs: 3/5				
		Alignment Complet	te		
		Alignment proce	ssing com	nplete.	
		 Imported runs: Reference run: 	6/6 C1		
		 Aligned runs: 	5/5		
			Revie	w chromatography	Review alignment $$

Once the Reference run has been chosen the automatic alignment is then performed. As the whole process proceeds you get information on what stage has been performed and also the % of the process that has been completed.

When the Alignment completes you can either review the chromatography or go to the Review Alignment using the options on the Alignment Dialog.

Click Review Alignment.

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.



For details on how to use Licensing go to Appendix 3 (page 60)

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



Layout of Alignment

To familiarize you with Progenesis LC-MS Alignment, this section describes the various graphical features 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 A2 to make it current. You will now be looking at the alignment of A2 to C1 in the Unaligned view.

Run	Include?	Vectors	Score
A1	V X	221	84.3%
A2	🗸 📈	240	82.2%
A3		244	83.9%
C1	\checkmark \times	Ref	
C2		362	97.8%
C3		422	99.0%

Now adjust the size and position of the current focus. First select the size by clicking on the Focus
grid size. Darken or lighten the runs using the contrast buttons. Then click on the Ion Intensity Map
to 'locate' the current focus. The other 3 views will update to reflect the new focus.

	Ion Intensity Map
Contrast	
Ion maps: 🚿 Alignment target 💺 Run being aligned	
Focus grid size: 🔲 Make box square	
◎ 1 ◎ 2 ◎ 4 ◎ 8 ◎ 16 ◎ 32 ◎ □	Alignment quality: 🔳 Good 😳 OK 🛑 Needs review

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 chosen 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 features appear to move up and down. Once correctly aligned, they will appear to pulse. During the process of adding vectors, this view will change to show a zoomed view of the area being aligned to help accurate placement of manual vectors.

Whole Run (Window C): shows the **focus** for the other windows. When you click on the view the orange 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 using the controls in the bottom left of the screen or 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 (**note**: this can be switched off using the options in the View menu) which focuses your review of the alignment process.

Total Ion Chromatograms (Window D): shows the current **total ion** chromatogram (green) overlaid on the Reference chromatogram (magenta). As the features 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 **Whole Run** view (Window C).

This view assists in the verification of the feature alignment.

Note: the icon to the right of the 'Window' titles expands

Total Ion Chromatograms 📰 🛛 the view .

Reviewing quality of alignment vectors

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

If the alignment has worked well then in Windows A and C the grid lines (option under **View** menu) should show minimal distortion, Window B (Transition) will show features 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 set the grid size to 16, (and untick the Make box square option) using the '**Focus grid size'** control at the bottom left of the window. 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 red area combined with a low score clearly indicates a 'miss alignment' and may require some manual intervention (see Appendix 4, page 61).



Stage 5A: Filtering

Now that you have reviewed your aligned Runs, you are ready to analyse them. Move to the **Filtering** stage, by either clicking on **Section Complete** (bottom right) or on Filtering on the workflow.



Peak Picking Parameters

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

🔼 Peak Picking Parameters	Peak Picking Parameters
Runs for peak picking Peak picking limits Maximum charge Retention time limits Choose runs for peak picking	Runs for peak picking Peak picking limits Maximum charge Retention time limits Sensitivity
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 feature outlines, it will still have outlines added to it and will be available in the experiment design setup. Learn more about why you might not want to select all runs. Year Ca Ca Ca Ca Ca Ca Ca Ca Ca Ca Ca Ca Ca C	You can adjust the sensitivity of the peak picking algorithm using these different methods. Each 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.
	Chromatographic peak width The chromatographic peak width gives the length of time over which an ion has eluted. If you set a minimum peak width, any ion that has eluted over a shorter period will be rejected.
Start peak picking Cancel	Start peak picking Cancel

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 feature detection pattern. This may be important when one or more of the runs appear noisy due to non-optimal chromatography or sample handling.

Note: features 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 (as shown above).

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. 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.

Peak Picking Parameters	Peak Picking Parameters
Runs for peak picking Peak picking limits Maximum charge Retention time limits	Runs for peak picking Peak picking limits Maximum charge Retention time limits
Maximum allowable charge	Retention time limits
You can set the maximum charge of ions to be detected. Ions with a charge greater than this value will be rejected. Maximum ion charge: 20	You can set the minimum and maximum retention time for peak picking. Toos 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 features detected, in this example there are 14624 features.



If required you can remove features based on position, charge state, number of isotopes or combinations of these feature properties.

As an example we will filter the features based on charge 'charge state'.

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

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

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



Therefore all features with a charge state of 1 or 8 and above will appear blue on the main view.



To remove these features press Delete 1032 Non Matching Features.

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

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

Normalisation Note that all remaining features will be used in the normalisation calculation. If deleting a substantial number of features, you should review the normalisation afterwards.	
Review normalisation >>	440

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

Recalculating normalisation								

The **Review Normalisation** page will open displaying plots for the normalisation of all the features 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 'not 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 LC-MS 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 feature (normalisation target abundance/run abundance).



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

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



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

🔼 LC-MS Tut	orial - Progenes	is LC-MS										
File Import Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Statistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics	Report	nonlinear
Review no Normalisatio different san By assuming unaffected b the factor by normalise ba	ormalisation n is required to nple runs. 1 that a significa y experimental v which the sam ack to its refere	allow comp ant number conditions, aple as a wh nce.*	parisons across of features are we can use ole varies to	Normalisatio Normalise to Normalise to Normalise to Don't use any	n Graphs Ni all proteins all proteins a set of hous / normalisatio	ekeeping pro	Method teins					

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 features.

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

Now return to filtering by clicking on the button		8-
on the bottom left of the screen	< Continue filtering features	

1

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

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

.

Stage 6: Experiment Design Setup for Analysed Runs

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

There are two basic types of experimental designs:

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

🔼 LC-MS Tutorial	- Progenesis LC-MS													
File	Review	Experiment	Review Peak	Peptide	Identify	Refine	Resolve	Review	Protein				. •	
Import Data A	lignment Filtering	Design Setup	Picking	Statistics	Peptides	Identifications	s Conflicts	Proteins	Statistics	Report			nor	nlinear
	• •	•											DIN	A M I C S
New														Ø Help ▼
Which expe	riment design t	ype do you w	ant to use f	or this exp	periment	?								
oo Bet	ween-subject	Design			Creat	te	0-0 0-0 Withir	n-subject	Design					
Do samples only one co subject des	from a given subj ondition? Then use sign.	ect appear in the between-	A		Dele	te	Have you take subject under use the within	n samples fr different co -subject des	om a given onditions? Then sign.	I		Before	During	After
To set up th runs accord of the samp	nis design, you simp ing to the condition les. The ANOVA ca	ly group the n (factor level) Ilculation			A2 <u>Remov</u> A3 <u>Remov</u>		Note: you must subject for eve subject design.	t have a sam ry condition	ple from every to use a within	- р	atient X	X1	X2	X3
assumes that and therefo	at the conditions ar ire gives a statistica	e independent il test of	c		Dele	te	For example, you would choose this type of design for a time series experiment where					¥1		¥2
equal.	e means of the con	itions are all			C1 Remov	<u>/e</u>	every subject h	nas been sam	pled at each ti	me ^P	Patient Y	38.	12	
					C2 <u>Remov</u> C3 Remov	/e	To set up this o	lesign, you t	ell the software	, –		74		
	Add.cond						not only which condition (factor level) each run belongs to but also which subject it came from. The software will then perform a repeated measures ANOVA.							23 1945
							A standard ANC	OVA is not ap	propriate becau	ise				
							the data violate independence. ANOVA individu eliminated or r condition differ more powerful	es the ANOV. With a repe Jal differenc educed as a rences (whic test).	A assumption of ated measures es can be source of betw h helps to creat	een te a				
							The within-sub as an extension to include com repeated meas	ject design o of the paire parison betw ures.	an be thought o ed-samples t-te: veen more than	of st two				

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

Additional information on how to apply the Within-subject Design is in Appendix 5 page 66

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.

Progenesis LC-MS									
Create a new experiment design									
Name: AC	~								
Start with an empty layout									
O Copy layout from:	Ŧ								
Import design from a file:	at file formats can I use?								
Group by:	(no valid groups) 🔹								
	Create design Cancel								

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

	aria L.C. MS									
Eilo										
Import Data Alignment	Experimen Filtering Design Setu	Review Pe p Picking	ak Peptide Statistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics	Report	nonlinear
AC I × 🗅 New										🕑 Help 🔻
Setup conditions Setup the conditions that (e.g., control, drug A, etc your samples to the corre	you want to compare be), and then assign each o ct condition. A1 <u>Remo</u> A2 <u>Remo</u> A3 Remo	Runs ow f te ve ve	Add Selected F	Runs to Cont	dition Sea	rch C2		<u>م</u>	C3	
Add condition									S	ection 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 spread sheet 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**.

🔼 Progenesis LC-MS		×						
Create a new experiment design								
Name: AC-2		~						
© Start with an empty layou	t							
© Copy layout from: AC		Ŧ						
Import design from a file:	Tutorial Groups What file formats can I use?]						
Group by:	Conditions •							
	Conditions							
	Date of Collection Location	icel						

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

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

🔼 LC-MS Tutorial - Progenesis LC-MS								- • •
File								
Review Experiment Rev Import Data Alignment Filtering Design Setup P	view Peak Peptide Picking Statistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics	Report	nonlinear
	•		•		•			D Y N A M I C S
AC AC-2 I × New								🕜 Help 🔻
Setup conditions	Runs Add Sele	ected Runs to	Condition 🔻	Search		Q		
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	10							
A1 <u>Remove</u>								
A2 Remove								
A3 Remove								
C <u>Delete</u>								
C1 <u>Remove</u>								
Add condition								
Add condition								
							-	ection Complete 🏵

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 features using the visual tools provided and edit features if required.

The review stage has 4 display modes: 1D, 2D, 3D and Feature 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 features on the aligned LC-MS runs.

Exploring analysed data using the Data displays

Window A: shows the list of features 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 features are included in the selection for the next section of the analysis.

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

progenesis

Run

-Cinh

The 1D Display

Window B: displays the Mass spectrum for the current feature on the selected Run (in window D).

Window C: displays the Chromatogram for the current feature on the selected Run (in window D).

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 feature editing tools are located in this window (see page 26 for functional explanation).

Clicking on the Expression Profile tab in Window D shows the comparative behaviour

of the feature across the various biological groups based on group average normalised volume. The error bars show +/- 3 standard errors.



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.







Retention time (m

25.5

26

Chromatogram (m/z = 1099.5684)

24.5

300000 250000 200000

150000 100000 50000 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 Montage mode, Window B displays a montage of the current feature 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 feature detection in detail to validate the correct detection of even fully overlapping features as shown above.



The 3D Display

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



The number of 3D views displayed in the montage is controlled using the <u>Select runs</u> link on the 3D Montage tab. The views can be set to **Rotate** automatically or you can rotate them manually by clicking and dragging them with the mouse.

Editing of features in the View Results stage

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



1. Locate the feature at approx 1004.77 m/z and 30.453 min using the Find tool.

Find specified location					×
Mass		1004.7700	*	m/z	
Retention time on:	Alignment reference	30.453	*	minutes	
or	select a sample 🔹	30.453	×	minutes	
				Go	

2. Select the Edit tool and click on the feature 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 feature to update the view.

5. To add a peak to an existing feature, ensure that **Edit** is selected then click inside the feature to reveal the handles.





- 6. Click on the 'plus' handle on the peak to add it.
- 7. Then click outside the feature to update the view.
- 8. Note: If you are not satisfied with the editing use the Undo button and retry.
- 9. Finally note: that a tag is automatically added to the edited feature in the table and the features id number is changed to the next available one at the end of the list.

Revie	ew Featur	es Experir	ment desig	gn: 🛛	C	
Y	No filter app	blied		Cre	eate	
#	Anova (p)	Fold	Tag	• N	lotes	*
9345	0.556	1.14		D	Adc	
202	0.556	1.13		D	Adc	
13468	0.557	1.53		D	Adc	
9753	0.557	1.35		a	Adc	
181	0.557	1.08		a	<u>Adı</u>	
7524	0.558	1.19		a	Adc	
4290	0.559	1.39		D	Adc	
6110	0.559	1.18		a	Adc	÷
•					•	

Revie	ew Featur	es Expe	eriment des	ign: AC
Y	No filter app	olied		Create
#	Anova (p)	Fold	Tag	▼ Notes *
9345	0.556	1.14		Adc
202	0.556	1.13		Adc
13468	0.557	1.53		Adc
9753	0.557	1.35		🔍 Adc
14626	0.557	1.08		📃 Adc
7524	0.558	1.19		🔬 Adc
4290	0.559	1.39		Adc
6110	0.559	1.18		🔍 Adc 🚽
•				•

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 features for Peptide Statistics

There are a number of ways to 'refine' your 'Ranked List' of analysed features before examining them with the Statistical tools in **Peptide Statistics**. These make use of simple 'Selection' and 'Tagging' tools that can be applied to the various groupings created in Stage 6 (page 19). An example is described below.

First expand the 'Features' table to show all the details by clicking on the 'Expander bar' to the right of the Review Features table.

Revie	w Featur	es Experi	ment desi	gn: AC	
7	No <mark>f</mark> ilter app	olied		Create	
#	Anova (p)	Fold	Tag	▼ Notes ▲	
284	6.3E-11	Infinity		Adc	
828	6.35E-11	Infinity		🔍 Adc	
6314	6.47E-11	Infinity		🔬 Adc	
10970	6.75E-11	Infinity		🔍 Adc	
11767	7.27E-11	Infinity		Adc	
1455	9.01E-11	3.5E+08		🔍 Adc	
8761	1E-10	Infinity		🔬 Adc	
8045	1.01E-10	Infinity		🔬 Adc	1
1391	1.01E-10	Infinity		🔍 Adc	
1142	1.21E-10	Infinity		🔬 Adc	
14023	1.27E-10	Infinity		🔍 Adc	
11985	1.36E-10	Infinity		🔍 Adc	
7057	1.48E-10	Infinity		Adc	
8028	1.56E-10	Infinity		🔍 Adc	
ا				•	
Dele	ete 1 selected	feature			

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

LC-N	MS Tutorial -	Progenesis	LC-MS													
File	e R : Data Ali	gnment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Statistics	Identify Peptides	Refine Identificatio	ons	Resolve Conflicts	Review Proteins	Protein Statistics	Report	ņç	onlin	ear
Revie	ew Featu	res Experi	ment desig	jn: AC		•										
7	No filter ap	plied		Create												
#	Anova (p) Fold	Tag	 Notes 	ł	lighest	Lowest	m/z	z	Mass	RT (mins)	RT window	Abundance	Intensity	Max CV	A
2	0.159	1.09		Add note	<u></u> A		С	805.441	3	2413.301	54.875	7.65	1.25E+08	1.08E+08	8.15	
18	0.463	1.05		🔜 Add note	e C		Α	1207.6552	2	2413.296	54.899	4.9	9.09E+07	5.04E+07	8.52	
24	3.18E-06	5.72E+04		📃 Add note	e C		Α	1100.5867	3	3298.738	44.928	2.43	8.39E+07	7.87E+07	53.4	
78	3.41E-06	1.15E+03		🔜 Add note	a A		С	1176.227	3	3525.659	48.238	9.44	6.91E+07	2.22E+07	23.1	
14	4.37E-07	856		📃 Add note	A		С	656.8612	2	1311.708	44.137	4.21	6.19E+07	1.17E+08	15	
20	2.22E-07	786		🔜 Add note	e C		Α	988.9849	2	1975.955	50.605	4.65	5.14E+07	9.19E+07	14.4	•
7	8.31E-08	130		📃 Add note	e C		Α	663.8693	2	1325.724	46.597	4.14	5E+07	1.69E+08	6.68	
23	2.08E-06	5.15E+03		📃 Add note	e C		Α	900.9713	2	1799.928	39.272	3.02	4.57E+07	8.16E+07	37.5	
56	5.85E-06	3.79E+03		📃 Add note	A		С	1061.007	2	2119.999	53.288	6.75	4.3E+07	2.52E+07	33.6	
41	5.57E-06	1.33E+03		📃 Add note	A		С	997.4477	2	1992.881	31.814	2.57	4.24E+07	3.81E+07	34.3	
94	0.0589	2.84		📃 Add note	e C		Α	976.8129	3	2927.417	54.357	13.3	3.72E+07	1.58E+07	55.3	
47	7.38E-06	3.58E+03		🔜 Add note	C		Α	1032.4669	3	3094.379	32.959	3.63	3.53E+07	3.95E+07	32.8	*
•			III												+	
Del	ete 1 selecte	d feature			Run	Expression Pr	ofile									
					Run:			-			x		0.01 -			
					Aggreg	ate										
					Sel	ect Ed	it \Xi 5).1/-	2		-		<u> </u>			
					🕞 Spl	it 🗁 Me	rae e	-					는 41.695-			
							<u>1</u> 04	315-		-1114		1	-ig	1.1		
					L-4 Ad		ete .e	007					eltio			
1D Dis	play 20 M	ontage 3D I	Montage		⊮7 Un	do 🥂 Rec	do g ob	00/-					濋 86.966-			
Sho	ow all outlin	es			Find		62	027								
🛛 🖾 Mu	Itiple colum	ins per condi	ition		Kon	lon charge -	1	····	nor		000	900		1000	1500	2000
Contra	ist:				Key:	on charge =	2	004 6	SUD	000 8 m	/z 000	003	500	1000 m/2	7000	2000
	_					lon charge =	3									
						ion charge =	4									
						on charge ≥	5									
Monta	ge size:															
	•	•	0	0										Section	n Comple	te 🤿

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

LC-M	5 Tutorial -	Progenesi	s LC-MS													
File import D	Rata Alig	eview gnment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Statistics	Identify Peptides	Refine Identificatio	ons	Resolve Conflicts	Review Proteins	Protein Statistics	Report	ņç	onlir	nec
Reviev	v Featu	res Expe	eriment desi	gn: AC		•										
Y 1	No filter ap	plied		Create												
#	Anova (p)	Fold	Tag	✓ Notes		Highest	Lowest	m/z	z	Mass	RT (mins)	RT window	Abundance	Intensity	Max C	^
29	0.00372	6.81		Add not	<u>e</u> (с	Α	630.866	5	3149.294	33.122	1.23	1.01E+05	4.42E+05	40.8	
33	0.0187	3.03		Add not	<u>e</u> ,	A	С	681.8701	2	1361.726	52.984	0.666	1.01E+05	4.15E+05	42.8	
35	0.0203	1.74		Add not	<u>e</u> (с	Α	1196.9362	3	3587.787	39.3	0.921	1E+05	2.57E+05	20.8	
64	2.85E-08	Infinity		Add not	<u>e</u>	с	Α	1585.5178	3	4753.532	46.597	0.359	1E+05	1.7E+05	16.9	
86	0.0798	2.62		Add not	e	<u> </u>		706.6327	4	2822.502	63.027	1.26	1E+05	5.39E+05	62.6	
71	0.0177	540		Add not	e	No tags to a	issign	1157.6035	3	3469.789	46.561	0.808	1E+05	2.24E+05	126	
41	0.0748	1.49		Add not	e	New tag		1081.0327	2	2160.051	36.226	0.754	1E+05	1.96E+05	25.4	
91	0.000569	3.03		Add not	e	Quick Tags	•	703.845	2	1405.675	45.667	1.94	9.99E+04	2.53E+05	17.7	
31	0.00365	20.8		Add not	e 🔗	Edit tags		1019.4452	3	3055.314	53.762	0.925	9.99E+04	1.94E+05	69.3	
2	0.000149	6.8		Add not	e	Add to clin	gallery	643.3413	2	1284.668	18.939	1.32	9.99E+04	2.37E+06	16.9	
81	0.00089	4.49		Add not	e	A	C	817.9225	2	1633.83	39.467	0.743	9.98E+04	2.87E+05	22.4	
79	0.0448	150		Add not	e	c	Δ	814 8726	4	3255.461	37.56	0.71	9.98E+04	4.7E+05	118	-
																•
Delet	e 2342 sele	cted featu	res		Run Run:	Expression P	rofile						0.01			_
					Aggre Se Se Carlos A	egate elect E olit PM dd De	dit 50 erge enge 54.3	315-					41.695-	9		
) Displ	ay 2D Mo	ontage 3[O Montage		19 U	ndo 陀 Re	edo 2 58.	007-			6		巖 86.966-			
Shov	v all outline	25			-			5	•							
Mult	iple colum	ns per con	dition		Kev:	Ion charge =	62.	804	805	806 8	07 808	809	500	1000	1500	200
ontrast	:				11	Ion charge =	2			m	v/z			<i>m/2</i>	r	
1						Ion charge = Ion charge = Ion charge ≥	3 4 5									
ontage	e size:															
0		•	0	•										Section	n Compl	ete

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

X
OK Cancel

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

Revie	ew Featur	es Experi	iment desi	gn: AC
Y	No filter app	blied		Create
#	Anova (p)	Fold	Tag	 Notes
2929	0.00372	6.81		🔍 Add not
1833	0.0187	3.03		🔍 Add not
4035	0.0203	1.74		Add not
5464	2.85E-08	Infinity		🔍 Add not
1686	0.0798	2.62		📃 Add not
4771	0.0177	540		🔍 Add not
2941	0.0748	1.49		🔍 Add not
3391	0.000569	3.03		Add not
5681	0.00365	20.8		Add not
8312	0.000149	6.8		Add not
2381	0.00089	4.49		Add not
2379	0.0448	150		Add not
•				
Dele	ete 2342 selec	ted feature	s	

Now right click on any feature 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**.

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

Revie	ew Featu	ires Experir	ment desi	ign: AC		•
	No filter aj	pplied		Create		
#	Anova (p) Fold	Tag	 Notes 		Highest
5464	2.85E-08	Infinity		Ad	d note	С
1686	0.0798	2.62		a Ad	d note	С
4771	0.01	Edited	_	Ad	d note	С
2941	0.074	Most abund	ant	Ad	d note	А
3391	0.00			Ad	d note	Α
5681	0.00	New tag		∏ Ad	d note	C
8312	0.00	Quick Tags		•	Anova p-value.	
2381	0.000 😭	Edit tags			Max fold chang	je
2379	0.044	Add to clip	gallery		No MS/MS dat	a
1720	0.323	1.32			No protein ID	
2212	0.000678	12.9		Ad	d note	А

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

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

The table now displays features with multiple tags.

Review Features Experiment design: AC No filter applied Create... # Anova (p) Fold Notes Tag Ŧ 0.159 1.09 Add note... 2 1.05 18 0.463 Add note... 24 3.18E-06 5.72E+04 Add note... 1.15E+03 Add note... 78 3.41E-06 14 4.37E-07 856 Add note... 20 2.22E-07 786 Add note... 7 Add note... 8.31E-08 130 23 2.08E-06 5.15E+03 Add note... 56 5.85E-06 3.79E+03 Add note... 41 1.33E+03 Add note... 5.57E-06 94 0.0589 2.84 Add note..

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

For example: to focus the table on displaying those features that have an **Anova p-value** \leq **0.05** click on **Create** on the filter panel above the table.



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

Create a filter
Show or hide features based on a selection or to create the filter. For more guidance, pleas
Available tags: Most abundant (2342 features) Max fold change ≥ 2 (9881 features)

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

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

#	Anova (p)	Fold	Tag 🔹	 Notes 	Highest	Lowest		#	Anova (p) Fold	Tag 🔹	Notes	Highest	Lowest
1440	2.94E-05	74.5		Add note	С	А		1440	2.94E-05 74.5		Add note	С	А
951	0.00113	138		Add note	с	А		951	0.00113 138		Add note	С	А
623	0.000379	12.6		Add note	с	А		623	0.000379 12.6		Add note	С	А
3206	0.0224	11.3		🔬 Add note	с	А	N	3206	0.0224 🥃 Mos	t abundant	ld note	С	А
643	1.73E-05	112		Add note	с	А		643	1.73E-05 🥃 Ano	va p-value ≤ 0.0	5 Id note	С	А
692	0.000628	411		Add note	с	А		692	0.000628 📒 Max	fold change ≥ 2	d note	С	А
853	7.89E-06	5.18E+05		Add note	с	А		853	7.89E-06 New	/ tag	ld note	С	Α
11868	0.0289	49.9		Add note	А	С		11868	0.0289 Quid	ck Tags	Id note	Α	С
5801	0.000122	21.3		Add note	Α	С		5801	0.000122	tanc	ld note	Α	С
13055	9.77E-08	Infinity		Add note	А	С		13055	9.77E-08	to clin callen	ld note	Α	С
1842	0.00615	2.99		Add note	Α	С		1842	0.00615	anery	d note	Α	С
5030	0.000154	32.2		Add note	Α	С		5030	0.000154 32.2		Add note	Α	С

Create a tag for them called **Significantly up in C**, tagging 3959 features.

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

Tag filter applied features may be hidden	Filter the features Create a filter Show or hide features based on a selection or boxes to create the filter. For more guidance	of their tags. Move tags to the appropriate ty please see the <u>online reference</u> .
Make sure that only the tag for the Most abundant features is shown and press OK .	Available tags: Anova p-value ≤ 0.05 (7965 features) Max fold change ≥ 2 (9881 features) Significantly up in C (3959 features)	Show features that have all of these tags: Most abundant (2342 features) Show features that have at least one of these tags: Hide features that have any of these tags:
	Clear the filter	OK Cancel

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

Stage 8: Peptide Statistics on selected features

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** features.

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

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



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

Principal Component Analysis (PCA)

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

Are there any outliers in my data? And does my data cluster according to my experimental conditions? It answers this question by:

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

PCA can be used to determine whether there are any outliers in the data and also look at how well the samples group. The groupings that can be observed on the 2D PCA plot can be compared to your experimental conditions and conclusions can be drawn regarding possible outliers in your data. Selecting features in the table will highlight the features 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 feature'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 feature, using the expression variance, sample size and difference between the means.

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

Note: Power analysis is discussed in Appendix 6 (page 68)

Correlation Analysis

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

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

On pressing OK the PCA will recalculate using these 8618 features, you can (to save time) stop this calculation by pressing **Cancel calculation** and then set up Correlation Analysis for the 7965 features.

Filter the features Create a filter Show or hide features based on a selection boxes to create the filter. For more guidance	of their tags. Move tags to the appropriate e, please see the <u>online reference</u> .
Available tags: Most abundant (2342 features) Max fold change ≥ 2 (9881 features) Significantly up in C (3959 features)	Show features that have all of these tags: Anova p-value ≤ 0.05 (7965 features) Show features that have at least one of these tags: Hide features 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 'feature correlation based on similarity of expression profiles'



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

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

The question is answered by:

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

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



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

To highlight all the features 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 features in the table and create a Tag for these features (Up regulated in C).

Create new tag		×
Up regulated in C		
	OK Cance	

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

Create new tag		×
Up regulated in A		
	OK 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 features. This shows how tabulated information about features 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 additional tagging will be required.

A Filter the features	×
Create a filter Show or hide features based on a selection of th create the filter. For more guidance, please see	neir tags. Move tags to the appropriate boxes to the <u>online reference</u> .
Available tags:	Show features that have all of these tags:
Most abundant (2342 features) Max fold change ≥ 2 (9881 features) Significantly up in C (3959 features) Up regulated in C (3959 features) Up regulated in A (4006 features)	Anova p-value ≤ 0.05 (7965 features) Show features that have at least one of these tags: Hide features that have any of these tags:
Clear the filter	OK Cancel

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

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

Stage 9: Identify peptides

Progenesis LC-MS does not perform peptide identifications itself. 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 LC-MS, using a number of different file types, and matched to your detected features.



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

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

The Peptide Search page shows the number MS/MS that have been matched to each feature in the Feature list (see above). MS/MS scans are matched to a feature if their precursor m/z and aligned retention time fall within the area of one of the isotopes of the feature. The MS/MS scans which are matched to the displayed features 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 features displayed in the Features list (in this case all the features that have an **Anova p-value≤ 0.05**. 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.

Filter the table to show only the features tagged **Significant p<0.05** as shown.

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

The total number of spectra included in this set is **20769** as shown on the Export button.

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

Note: many of the abundant features have a large number of spectra associated with them.

LC-	MS Tutori	ial - Progenesis L	.C-MS														×
Fi	e																
Impo	t Data	Alignment	Filtering	Experiment Design Setur	Revi	ew Peak P cking St	eptide atistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics F	Report		nor	line	ar
		•				•	•				0	•	•				c s
Iden	tify Pep	tides				MS/MS S	pectra				_						
Run m	s/ms ion s	earches by expo	rting peak li earch results	st files to a	an a	A Batch in	clusion optio	ns for creatin	g export list of m	s/ms spectra							
be imp	orted bacl	k in to link prote	in IDs to de	tected peaks			Rank	greater that	an 🔹 5			Feature intensity	less than	•			
Featur	es:																
7	Tag filte	r applied		Edit			Feature ID	less than	•			Precursor intensity	less than	•			
	features	may be hidden		Lan	·		Charge	less than	•		Pre	cursor intensity (%)	less than	•			
#	MS/MS	 Proteins 	Score	Tags	· ^			(
56	59	0			54	S	can number	less than	•			Run name	contains	•			
34	58	0			103		Exported	Cogual to				Pentide sequence	containe				_
/	54	0		-				equal to	•	•			Containa	•			
78	52	0					Isotope	less than	•			Protein accession	contains	•			
91	50	0			in the second se		ID	(Bestein des sisting					_
64	46	0					ID SCOLE	less than	•			Protein description	contains	•			
3820	45	0			103							Include i	n export	Exclude fr	om export	lear all filter	rs
59	45	0			103		_						. copon				
135	43	0			10	Export	Rank #	Run	Scan number	Exported	Feature intensity	Precursor intensity	(%)	Charge	Precursor m/z	Isotope	ld
20	41	0			10		4	132 A3	4540	Yes	9.0e+006	3.8e+006	41.7	2	980.9709	1	A
21	41	0			6		5	132 A2	4602	Yes	1.4e+007	5.5e+006	39.8	2	980.9717	1	
30	40	0			65		6	132 A1	4461	No	1.1e+007	3.7e+006	34.8	2	980.9718	1	
32	39	0		-	55	4		132 A3	4580	No	9.0e+006	1.5e+006	16.3	2	980.9711	1	+
425	39	U			-	Feature	number	120 m/	7 052 7080	retention	time 53 245	min charge +3					-
Perf	orming t	the search				Pup:C1	Scon numb	or6594	2 , 52.1 , 60,	retention	1 time 55.245	inin, charge +5					9
Select	the search	h program you'r	e using:			Runora	I	e1.0504		5E+05-							
Maso	ot		-			51.945	-										
Help						·E 52.829		111	111 .	4E+05-							
	Export 115	44 ms/ms spect	ra			j 53.525	_ Ψ		en sit	3E+05-							
MS	AS Preproo	cessing				臣 54.315	-		Ē	2E+05-							
	imit fragr.	ment ion count	40	×		8 54 969				15.05							
[/ Deisoto	ping and charge	e deconvolu	tion		55.500				1E+00-		Th					
						55.503		53 95	54	0+	rahar i shini	. <u>بالطالبان في اللبين.</u> 1000		2000		2000	
	Import	search results						m/z		5		1000	m/z	2000		5000	-
	Clear all	identifications													C		
	ereer un														Section	omplet	e 🥑

Performing the search Select the search program you're using:
Mascot
Help
Export 20769 ms/ms spectra
MSMS Preprocessing
Limit fragment ion count 40
Deisotoping and charge deconvolution
Import search results
Clear all identifications

To control the number of spectra for each feature, expand the **Batch inclusion options**.

For example: We will make use of the 'Rank' value to reduce the number of Spectra being used for each feature 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 feature.

	Export	Rank	: #	Run	Scan number	Exported	Feature intensity	Precursor intensity	(%)	Charge	Precursor m/z	Isotope Id so	ore	*
		23			4000	M	1.0000	1.4000	10	2	CEC 0010	1		_
	1	24	The rank of	of each	h MS/MS spec	trum found	d by comparing it	s '%' values against	all other	spectra n	natched to the	same feature.		
E		25	9	A3	5042	No	9.3e+007	8.8e+005	1.0	2	656.8614	1		
		26	9	A1	5379	No	1.2e+008	1.0e+006	0.9	2	656.8610	1		
		27	9	A3	5374	No	9.3e+007	7.7e+005	0.8	2	656.8615	1		-
4]				•	

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

Set the Rank filter to 'greater than' 5 and click **Exclude from export** this reduces the number to spectra to export to **11544**

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 from 1000 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 11544 spectra to export, as described above:

- 1. Select appropriate search engine i.e. Mascot
- 2. Click 'Export current query set' 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 7 (pages 69) for details of the 'Search Engine' parameters used in this example

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 (Protein Search Results.xml). Select the 'Mascot' method and import this file to see results like those below.

On importing the Search results the Features table updates to reflect the identified proteins and the relevant score for each searched feature.

LC-	MS Tutorial -	Progenesis LC-MS	5										[- • ×
Fil	e													
Impor	t Data Alig	eview Inment Filterir	Experime ng Design Se	nt Review tup Pick	v Peak Peptide ing Statistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics I	Report		nor	linoar
									•	•	0			
Ident	ify Peptid	es			MS/MS Spectra									
Run m	/ms ion searc	<u>thes</u> by exporting p program Search	peak list files to a results files may	a then	 Batch inclusion option 	ns for creati	ng export list of m	s/ms spectra						
be imp	orted back in	to link protein IDs	to detected pea	ks.	Rank	greater th	an 🔻 5			Feature intensity	less than	•		
Featur	es:													
7	Tag filter ap	oplied	E	dit	Feature ID	less than	•			Precursor intensity	less than	•		
-	reatures may	, be maden			Charge	less than	•		Pre	cursor intensity (%)	less than	•		
#	MS/MS 🔻	Proteins S	Score Tags	• ^										
56	59	4 gi 2170250 10	02		Scan number	less than	•			Run name	contains	•		
34	58	2 gi 2549763 87	7.7	94	Exported	a sust to		_		Pentide sequence	(anataina	_		
7	54	6 gi 1121811 Bi	2.8	100	Exported	equal to		•		r opnice coqueries	contains	•		
12/	54	0	11	53.	Isotope	less than	•			Protein accession	contains	•		
78	50	3 gi 2551019 11	25		🔼 Import search	results			—X —					
64	46	2 gij2349705 72	2.5		6078	anala laitea la:				Protein description	contains	•		
3820	45	0			0978 58	arch hits h	ave been importa	eo ano assig	ned to reatures					1 11 51
59	45	ů O								Include	n export	Exclude fr	om export	lear all filters
135	43	4 gi 2170250 Bi	2.3	4						Precursor intensity	(%)	Charge	Precursor m/z	Isotope Id :
20	41	5 gi 2067250 94	4.2	1	V				ОК	6.2e+00	7 83.5	2	595.3201	1 🔺
21	41	1 gi 3661641 44	4.5 🧧	10						9.4e+00	7 81.7	2	595.3196	1
30	40	3 gi 2170250 B(0.9 📄	10.	V 4	10 A1	3852	Yes	1.1e+008	7.3e+00	7 66.7	2	595.3195	1
32	39	4 gi 2170250 57	7.5 📄	101	4	10 40	4011	V	7 4 007	4.900	7 CAD	2	EDE 2104	
425	39	0		· .	Feature number	10 m/z	505 210 ro	tention tir	me 36 021 mi	n charge +2	_	_	_	
Perfo	orming the	search			Due: A2 Seen number	0, 111/2	. 595.519, 10	cencion ci	me 50.921 mi	n, charge +2				
Select	the search pr	ogram you're usin	g:		Run AS Scan humb	el.3970		4E+05-						
Masc	ot	•	•		35.673-									
Help					·Ē 36.569			3E+05-						
	Export 11543 r	ms/ms spectra			O		Sity .					1		
MSN	1S Preprocess	ing					lite	2E+05-						
E L	imit fragmen	t ion count 40	A V		툞 38.304 -			1E+05-						
B	/ Deisotopin	g and charge deco	onvolution		39.156 -									
					595	596 597	598 599	0-	ladad (1)	500		بالعرب الليب	1000	
	Import sea	rch results				m/z		3			m/z			~
	Clear all ide	ntifications											Section	Complete 🔿
			_										Jeccion	acculate ()

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

Performing the search Select the search program you're using:	
Mascot 🗸	
Peaks (.mgf & .pepXml)	
Mascot	
MSPepSearch (tsv)	
PLGS (*.xml)	
PLGS v2.3 (final peptide report)	
ProteinPilot (group2xml)	A
Proteome Discoverer (.xls;.xlsx)	v
Scaffold Spectrum Report	olution
SEQUEST (dta & out files)	
SEQUEST (dta & pepXml files)	
SEQUEST (.sqt and .ms2)	
SpectrumMill (*.tsv & .pkl)	
Phenyx	
Clear all identifications	

Stage 10: 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** 'Acceptance Criteria' on which to base the sequential filtering of the Peptide results, the following thresholds will be applied:

- 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 Doesn't contain 'Clostridium difficile'

LC-N	/IS Tutori	al - Progene	sis LC-M	IS														- • •
File																		
Import	Data	Review	Filter	rina	Exp	perime sign Set	nt Revie	ew Peak	Peptic	ie ics	Identify Pentides	Ref	ine rations	Resolve	e Review ts Proteins	Protein Statistics	Report	
	Data				Des	Sigir Sc	tup Pi		500130	6		- Contraint						nonimear
						-		·			•			-				
eatu	ires			_		Рер	tide Se	arch R	esults									
#	Total Hi	its m/z	RT(min	s Cha	A.	Bat	tch deletio	n optior	15				_					
7	30	663.87	46.60	2	=			Score	less thar	n .	• 40				Sequence	Length less	s than 🔻	
10	16	595.32	36.92	2									_			~		
11	18	5/3.80	24.19	2				Hits	less thar	۰ I	•					Charge les:	s than 🔹	
12	22	5/3.32	41.74	2				Macc	1		. —		_		S.a.	quence .		
16	25	572.22	44.14	2				111022	less that	1	•					docure (cou	itains *	
17	30	498.26	25.38	2			Mass erro	r (ppm)	less thar	, ·	•				Ac	cession cor	ntains 🔹	
20	25	988.98	50.61	2														
21	5	601.83	41.52	2				M/Z	less thar	n .	•				Des	cription doe	esn't contain 💌	
23	15	900.97	39.27	2							_					_		
24	20	1100.59	44.93	3			Retentio	on Time	less than	n '	•				Modifi	cations cor	ntains 🔹	
28	5	601.83	40.58	2										_				
30	9	614.35	19.34	2										De	elete matching searc	h results	Delete non-matchin	g search results Clear all filters
31	15	600.98	39.30	3														
32	20	484.26	24.15	2			#	Score	Hits	m/z	RT(mins) Charge	Mass	Mass err	r Sequence	Accessio	n Modifications	A
34	10	941.79	58.29	3			580	34.98	1	573.29	30.42	3	1716.8(0.38	🕥 VNTNVSALIA	🔇 gi 12669	7810 [14] Oxidatic	flagellin C [Clostridium difficile 63
41	20	997.45	31.81	2			582	66.07	5	547.31	50.36	3	1638.89	-0.43	🕥 IGVAQGVDA	🌒 gi 12669	9128	rubrerythrin [Clostridium difficile (
46	20	980.48	42.30	2		V	582	66.07	5	547.31	50.36	3	1638.89	-0.43	🎯 IGVAQGVDA	🌒 gi 32620	4727	Rubrerythrin [Clostridium papyros
49	25	702.36	43.00	3			591	53.42	3	497.27	30.70	2	992.52	-0.80	🎯 ELITFGADK	🎯 gi 12669	7970	alpha-subunit of electron transfer
56	20	1061.01	53.29	2			593	9.09	1	907.98	43.48	2	1813.94	-5.40	SEGEIVQVI 😵	🔇 gi 22787	7281 [1] Oxidation	H(+)-transporting two-sector ATP
69	5	623.83	37.74	2		7	596	84.21	5	611.33	26.65	2	1220.65	-0.69	TAATGFGVA\	🌒 gi 12669	7752	NAD-specific glutamate dehydroc
71	50	611.99	25.57	3			607	85.69	5	732.62	52.09	4	2926.43	0.10	TYNNGYSNA	🎯 gi 25510	1963	cell surface protein (S-layer precu
76	15	557.64	42.39	3		V	607	85.69	5	732.62	52.09	4	2926.43	0.10	TYNNGYSNA	🎯 gi 25497	6387	cell surface protein (S-layer precu
78	15	1176.23	48.24	3			628	32.10	3	442.91	46.62	3	1325.72	-0.58	IFFEGTLASTIK	S gi 87239	956	s-layer protein, partial [[Clostridiu
82	5	760.90	39.25	2			628	32.10	3	442.91	46.62	3	1325.72	-0.58	IFFEGTLASTIK	😻 gi 20672	5029	S-layer protein A, partial [[Clostrid
84	25	1053.03	43.00	2			628	32.10	3	442.91	46.62	3	1325.72	-0.58	IFFEGTLASTIK	😻 gi 20672	5031	S-layer protein A, partial [[Clostrid
86	30	1170.55	59.31	2			628	32.10	3	442.91	46.62	3	1325.72	-0.58	IFFEGTLASTIK	😻 gi 25497	6387	cell surface protein (S-layer precu
91	10	9/6.48	52.04	3			628	32.10	3	442.91	46.62	3	1325.72	-0.58	IFFEGTLASTIK	S gi 92380	869	s-layer protein, partial [[Clostridiu
96	10	832.18	54.48	4			628	32.10	3	442.91	46.62	3	1325.72	-0.58		9 gi 11218	1139	s-layer protein, partial [[Clostridiu
99 101	20 20	528.82	39.90	2			630	60.59	5	683.35	29.02	2	1364.68	-0.47		gi 20957	1234	cell wall protein V [[Clostridium] d
100	20	825.69	44.90	4			630	60.59	2	683.35	29.02	2	1304.6	-0.47		gi[25509	9012	cell surface protein (putative hem 🛫
111	10	002.42	40.24	4	-	-						_						4
-		-79.46 III	417.20	•		1846	search re	sults. 48	6 matchin	g batch d	lelete opt	tions.						Section Complete 🏵

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 (486 matching out of 1846)

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



Now **Clear all filters** and then apply the next filter (Hits: less than 2) followed by the remaining two filters (page 39)

Having applied all 4 filters the **Peptide Search Results** should be reduced to **1129**.

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

Stage 11: Resolve Conflicts

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

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

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

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



Depending on the ordering, make '**cell surface protein precursor etc '** the current protein by clicking on it in Window A (a circular orange symbol indicates current protein).

This protein has 33 peptides assigned (window B) which have a total of 27 conflicts. To view the conflicting assignments click on the **Protein Resolution** tab (window C).

LC-IVIS Tutorial - Proge	Ellesis EC-IMD															
File Review nport Data Alignmen	t Filterin	Ex g De	sign Setup	Review Peak Picking	Pepti Statis	ide stics	Identify Peptides	Refir Identifica	ne ations	Resolve Conflicts F	Review Proteins	Protein Statistics	Rep	ort	no	nlinec
Experiment design: A	C No filter a	pplied	(▼ Create	Pe	ptides	of gi 25	4976387					Y	No filter applied		Create
Accession	Peptides	Conflict:	Score	Tags • *	•	#	Score	Hits N	Mass	Mass error (p	RT (mins)	Charg	e Tags	 Abundance 	e Confl	ct: Peptide S ^
🎯 gi 254976387 (+6)	 33 (18) 	27	3.75E+03			607	85.7	5 2	926.431	0.0963	52.1	4		6.56E+05	1	S TYN
🌍 gi 255101963 (+4)	30 (15)	27	3.32E+03			91	72.5	5 2	926.432	0.427	52	3		1.09E+07	1	S TYN
🌚 gi 254976385 (+1)	11 (2)	17	1.06E+03			5672	102	4 1	800.888	-0.523	19	2		2.88E+04	1	S VET
🌍 gi 255656776 (+1)	9 (0)	17	786			/ 12	73.1	5 1	144.634	-0.13	41.7	2		2.28E+07	1	SLL:
🌍 gi 126700407	9 (1)	16	886		1	/ 16	73.1	5 1	144.634	0.371	40.7	2		2.69E+07	1	SLL:
장 gi 209571234 (+3)	22 (12)	11	2.25E+03		J	1097	67.8	5 1	144.633	-0.576	38	2		2.69E+05	1	🔇 GLL
			1.005.00				105	5 11	832 036	-0.565	24.3	2		0.01 E+05	1	🕥 ILN.
gi 384359782 (+8) Protein: gi 25497 Protein: gi 25510 Peptide Views Protein	21 (11) "" 76387 cell 01963 cell Resolution	11 surfac surfac	e protein	(S-layer pr (S-layer pr	ecurs	or prot	tein) [Clo tein) [Clo	ostridium	difficile	e QCD-66c26 e QCD-63q42	1	2		5.512+05	1	4
 gij384359782 (+8) Protein: gij25497 Protein: gij25510 Peptide Views Protein Conflicting protein 	21 (11) " 76387 cell 01963 cell Resolution ins for fea	II I surfac I surfac ture 60	r.88E+03 e protein e protein) (S-layer pr (S-layer pr	ecurs ecurs Pept	or prot or prot	tein) [Clo tein) [Clo tein) [Clo	ostridium ostridium 101963	difficile	e QCD-66c26 e QCD-63q42	1	2	-	5.512+05	1	F
gij384359782 (+8)	21 (11) 76387 cell 76387 cel	II I surfac I surfac ture 60 Conflict	1.88±+03 ee protein ee protein 7 7) (S-layer pr (S-layer pr	ecurs Pept	or prot or prot ides of	tein) [Clo tein) [Clo f gi 255'	ostridium ostridium 101963 Hits Mass	difficile difficile	ass error (p	I I RT (mins)	Charge	Tags	 Abundance 	Conflict:	Peptide Sec
gij384359782 (+8) Protein: gij25497 Protein: gij2551 eptide Views Protein Conflicting protein Accession gij254976387 (+6)	21 (11) " 76387 cell 01963 cell Resolution ins for fea Peptides 0 33 (18)	II I surfac I surfac ture 60 Conflict 27	1.88±+03 e protein 7 t: Protein Sci 3.75±+03	(S-layer pr (S-layer pr ore	Pept	or prot or prot ides of # \$	tein) [Clo tein) [Clo f gi 255 score 85.7	ostridium ostridium 101963 Hits Mass 5 2926	difficile difficile s M	ass error (p 1 0.0963	240 I I I I RT (mins) 52.1	Charge 4	Tags	 Abundance 6.56E+05 	Conflict:	Peptide Set
gij384359782 (+8) Contain: gij25497 Protein: gij25497 Protein: gij25510 Protein Conflicting protei Accession gij254976387 (+6) gij255101963 (+4)	21 (11) " 76387 cell 01963 cell Resolution ins for fea Peptides 33 (18) 20 (15)	II I surfac I surfac ture 60 Conflict 27	1.88±+03 e protein (re protein (7 7 t: Protein Sci 3.75E+03 2.225C 02	(S-layer pr (S-layer pr ore	Pept	ides of 607 12	tein) [Clo tein) [Clo f gi 255 core 85.7 73.1	ostridium ostridium 101963 Hits Mass 5 2926 5 1144	difficile difficile s M 6.431	ass error (p 1 0.0963 -0.13	240 I I I I I I I I I I I I I I I I I I I	Charge 4 2	Tags	 Abundance 6.56E+05 2.28E+07 	Conflict: 1 1	Peptide Sec
gij384359782 (+8) Protein: gij25497 Protein: gij25510 eptide Views Conflicting protei Conflicting protei Accession gij254976387 (+6) gij255101963 (+4)	21 (11) 76387 cell 11963 cell Resolution Ins for fea Peptides • 33 (18) 20 (15) Turr	II I surfac I surfac ture 60 Conflict 27 27 27 0 off all p	7 t: Protein Sc: 3.75E+03 2.325_03 eptides	(S-layer pr (S-layer pr)))))))))))))))))))))))))))))))))))	Pept	566 or prot cides of # \$ 607 12 16	tein) [Clo tein) [Clo f gi 255 core 85.7 73.1 73.1	ostridium ostridium 101963 Hits Mass 5 2926 5 1144 5 1144	difficile difficile s M .634 .634	ass error (p) 0.0963 -0.13 0.371	RT (mins) 52.1 41.7 40.7	Charge 4 2 2	Tags	 Abundance 6.56E+05 2.28E+07 2.69E+07 	Conflict: 1 1 1	Peptide Sec TYNN GILD GILD
gij384359782 (+8) Protein: gij25497 Protein: gij25510 eptide Views Protein Conflicting protei Accession gij254976387 (+6) gij255101963 (+4)	21 (11) 76387 cell 11963 cell Resolution Peptides • 33 (18) • 33 (18)	II I surfac I surfac ture 60 Conflict 27 27 27 27 27	1.88±+03 ee protein (ee protein (7 t: Protein Sc: 3.75±+03 2.255±+03 2.255±+03	(S-layer pr (S-layer pr)))))))))))))))))))))))))))))))))))	Pept	566 or prol cides of # S 607 12 16 1097	125 tein) [Clc tein) [Clc f gi 255 icore 85.7 73.1 73.1 67.8	Stridium ostridium 101963 Hits Mass 5 2926 5 1144 5 1144 5 1144	s M .634 .634	ass error (p 1 0.0963 -0.13 0.371 -0.576	240 RT (mins) 52.1 41.7 40.7 38	Charge 4 2 2	Tags	 ✓ Abundance 6.56E+05 2.28E+07 2.69E+07 2.69E+05 	Conflict: 1 1 1 1	Peptide Sec ^ TYNN E GILD GILD GILD
gij384359782 (+8) C Protein: gij25497 Protein: gij25510 Protein: gij25510 Conflicting prote Accession @ gij254976337 (+6) @ gij255101963 (+4)	21 (11) 76387 cell 01963 cell Resolution ins for fea Peptides • 33 (18) 20 (15) Turr	II I surfac I surfac ture 60 Conflict 27 37 off all p	7 t: Protein Sc: 3.75E+03 2.32E-03 eptides	(S-layer pr (S-layer pr ore	Pept	566 or prot or prot ides of # \$ 607 12 16 1097 91	tein) [Clo tein) [Clo tein) [Clo f gi 255' icore 85.7 73.1 67.8 72.5	Stridium ostridium 011963 Hits Mass 5 2926. 5 1144. 5 1144. 5 1144. 5 2926.	Million difficile difficile s M i.431 .634 .633 i.432	ass error (p 1 0.0963 -0.13 0.371 -0.576 0.427	XT (mins) 52.1 41.7 40.7 38 52	Charge 4 2 2	Tags	 Abundance 6.56E+05 2.28E+07 2.69E+07 2.69E+05 1.09E+07 	Conflict: 1 1 1 1 1	Peptide Set ^ TYNN GILD GILD GILD GILD TYNN
gij384359782 (+8) <pre> rotein: gij25493 Protein: gij25510 gij25510 conflicting protei Accession @ gij25976387 (+6) @ gij255101963 (+4) </pre>	21 (11) # *6387 cell 11963 cell 11963 cell ins for fea Peptides • 33 (18) 20.05 Turn	II I surfac I surfac ture 60 Conflict 27 37 n off all p	7 t: Protein Sc 3.75E+03 2 32C, 03 eptides	(S-layer pr (S-layer pr ore	Pept	566 or prot or prot ides of # \$ 607 12 16 1097 91 1205	tein) [Clc tein) [Clc f gi 255' icore 85.7 73.1 67.8 72.5 85	Destrictium pstrictium pstrictium 001963 Hits Mass 5 2926 5 1144 5 1144 5 2926 5 1144 5 2926 5 2926 5 2926	s M .634 .634 .633 .432 .634 .633	e QCD-66c26 e QCD-63q42 ass error (p 1 0.0963 -0.13 0.371 -0.576 0.427 -0.452	RT (mins) 52.1 41.7 40.7 38 52 52	Charge 4 2 2	Tags	 Abundance 6.56E+05 2.28E+07 2.69E+05 1.09E+07 6.4E+05 	Conflict: 1 1 1 1 1 1 1 1 1	Peptide Set TYNN GILD GILD GILD TYNN TYNN
gij34359782 (+8) ← Protein: gij2549? Protein: gij25510 eptide Views Protein Conflicting protei Accession @ gij2597637 (+6) @ gij255101963 (+4)	21 (11) ## #6387 cell 11963 cell Resolution [ins for fea Peptides • 33 (8) 20 (5) Turn	II I surfac I surfac ture 60 Conflict 27 27 27 27 27 27 27 27 27 27 27 27 27	7 t: Protein f 7. t: Protein Sc 3.75E+03 a.37E+03 a.37E+03 c	(S-layer pr (S-layer pr)))))))))))))))))))))))))))))))))))	Pept	566 or prot cides of # S 607 12 16 1097 91 1205 5672	tein) [Clc tein) [Clc f gi 255' core 85.7 73.1 67.8 72.5 85 102	bit bit 101963 101963 Hits Mass 5 2926 5 1144 5 1144 5 2926 5 1144 5 2926 5 2926 5 2926 4 1800	Million difficile difficile s M .634 .634 .633 .432 .432 .432 .888	ass error (p 1 0.0963 -0.13 0.371 -0.576 0.427 -0.452 -0.523	RT (mins) 52.1 41.7 40.7 38 52 52 19	Charge 4 2 2 2 2	Tags	 Abundance 6.56E+05 2.28E+07 2.69E+07 2.69E+07 1.09E+07 6.4E+05 2.88E+04 	Conflict: 1 1 1 1 1 1 1 1	Peptide Set TYNN GILD GILD GILD GILD TYNN TYNN VETG
gij34359782 (-8) C Protein: gij25497 Protein: gij25497 Protein: gij25510 eptide Views Protein: gij25510 Conflicting prote Accession @ gij254976387 (+6) @ gij255101963 (+4)	21 (11) # /6387 cell /6387 cell 11963 cell Resolution ins for fea Peptides 9 33 (8) 20 (45) Turn	II surfac surfac ture 60 Conflict 27 37 off all p	7 r Protein f 3.75E+03 apr. 03 apr. 03 apr. 03 apr. 03 apr. 03 apr. 03 apr. 03 apr. 03 apr. 03 apr. 04 apr. 04 c Protein f c P	(S-layer pr (S-layer pr ore	Pept	ides of <i>x</i> prot <i>x</i> s 607 12 16 1097 91 1205 5672 2646	tein) [Clc tein) [Clc f gi 2555 icore 85.7 73.1 73.1 67.8 72.5 85 102 54.1	bit bit 101963	Million difficile difficile s M .634 .634 .634 .634 .633 .429 .888 .483	ass error (p 1 0.0963 -0.13 0.371 -0.576 0.427 -0.452 -0.523 0.682	RT (mins) 52.1 41.7 40.7 38 52 52 19 63.3	Charge 4 2 2 2 2 4	Tags	 Abundance 6.56E+05 2.69E+07 2.69E+07 2.69E+07 2.69E+07 2.69E+07 2.88E+04 6.25E+04 	Conflict: 1 1 1 1 1 1 1 1 1 1 1	Peptide Set TYNN GILD GILD GILD GILD TYNN VETGI DSVDI
gij34359782 (+8) Protein: gij25497 Protein: gij25517 Protein: gij25517 Protein: gij2597 Protein: gij2597 Protein: gij2597 Protein: gij2597 Protein: gij2597 Protein: gij2597 gij2497 Protein: gij2597	21 (1) (6387 cell 11963 cell Resolution ins for fea Peptides • 33 (18) 20 (15) Turr	II I surfac I surfac ture 60 Conflict 27 off all p	r Brotein (re protein (re protein (r r r r r r r r r r r r r	(S-layer pr (S-layer pr ore	Pept	566 or prot ides of # S 607 12 16 1097 91 1205 5672 2646 34	tein) [Cld tein) [Cld f gi 2555 icore 85.7 73.1 73.1 67.8 72.5 85 102 54.1 87.7	stridium ostridium 01963 Hits Mass 5 2926 5 1144 5 1144 5 2926 5 1144 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2822	Million s M 6.634 .634 .634 .633 .6432 .6483 .633 .4432 .888 .483 .351 .351	ass error (p 1 0.0363 -0.13 0.371 -0.576 0.422 -0.452 1.71	RT (mins) 52.1 41.7 40.7 32 52 19 63.3 58.3	Charge 4 2 2 2 2 4 3	Tags	 Abundance 6.56E+05 2.28E+07 2.69E+05 1.09E+07 6.4E+05 2.88E+04 6.25E+04 2.15E+07 	Conflict: 1 1 1 1 1 1 1 1 1 1 1 1 1	Peptide Sec TYNN GILD GILD GILD GILD TYNN VETG DSVDI TIND
 gij384359782 (+8) Protein: gij25497 Protein: gij25510 conflicting protein Conflicting protein Accession gij254976337 (+6) gij255101963 (+4) 	21 (1) (6387 cell 11963 cell Resolution ins for fea Peptides • 33 (18) 20.055 Turr	II I surfac I surfac ture 60 Conflict 27 off all p	r astros re protein (re protein (r r r r r r r r r r r r r	(S-layer pr (S-layer pr ore	V V Pept V V V V 1 V 1 V 1 V 2 V 2 V 2 V 2	S66 or prol or prol cides of 12 16 1097 91 1205 5672 2646 34 228	tein) [Clc tein) [Clc f gi 255: icore 85.7 73.1 73.1 67.8 72.5 85 102 54.1 87.7 78	stridium ostridium ostridium 01963 Hits Mass 5 2926 5 1144 5 2926 5 1144 5 2926 5 1144 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2926 5 2822 5 2822	Million difficile difficile s M .634 .634 .633 .4431 .633 .4432 .4433 .4353 .432 .351 .345		TT (mins) 52.1 41.7 40.7 38 52 19 63.3 58.3 58.3 58.3	Charge 4 2 2 2 4 3 2	Tags	 Abundance 6.56E+05 2.28E+07 2.69E+07 2.69E+07 1.09E+07 6.4E+05 2.88E+04 2.15E+07 3.83E+06 	Conflict: 1 1 1 1 1 1 1 1 1 1 1 1 1	Peptide Set A TYNN E GILD GILD GILD TYNN YEIG DSVDJ TIND
 gi]384359782 (-8) Protein: gi[25497 Protein: gi[25510 Protein: gi[25510 Conflicting protein Conflicting protein Accession gi]254976387 (+6) gi]255101963 (+4) 	21 (1) "" 76387 cell 11963 cell Resolution ins for fea Peptides • 33 (18) 10/751 Turr	II I surfac I surfac ture 60 Conflict 27 27 27 0 off all p	7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	(S-layer pr (S-layer pr ore	Pept V V V V V V V V V V V V V V V V V V V	or prot or prot cides of # \$ 12 16 1097 91 1205 5672 22646 34 228 736	tein) [Clc tein) [Clc f gi 255 icore 85.7 73.1 67.8 72.5 85 102 54.1 87.7 78 76.5	ostridium ostridium ostridium ostridium ostridium itts Mass 5 2926 5 1144 5 1144 5 2926 5 2926 5 2926 5 2926 5 2926 5 2822 5 2822 5 2822 5 2822	Million difficile difficile s M .634 .634 .633 .432 .432 .433 .432 .433 .432 .433 .433 .434 .3351 .345 .346	0.000 a QCD-66c26 b QCD-63q42 ass error (p 1 0.0963 -0.13 0.371 -0.575 0.427 -0.452 -0.523 0.682 1.71 -0.154 0.236	RT (mins) 52.1 41.7 38 52 52 19 63.3 58.3 58.3 58.3 51	2 4 2 2 2 2 4 3 2 3	Tags	 Abundance 6.56E+05 2.28E+07 2.69E+07 2.69E+07 6.4E+05 2.88E+04 6.25E+04 2.15E+07 3.83E+06 4.9E+05 	Conflict: 1 1 1 1 1 1 1 1 1 1 1 1 1	Peptide Set TYNN' GILD GILD GILD TYNN' TYNN' VEIG DSVD TIND TIND TIND

In this case the conflicting peptide assignments are with '**The same protein**' (from a different strain) which also contains 27 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).

Note: as you un-assign the peptides the number of conflicts update 'on the fly' in all the windows.

🔼 LC-MS Tutorial - Progenesis LC-MS										- 0 💌
File Experiment Experiment Review Peak Import Data Alignment Filtering Design Setup Picking	Peptide Statistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics	Report		non	linear
Experiment design: AC Proteins Vo filter applied Create	Peptide	es of gi 254	976387				Wo filte	er applied	(Create
Accession Peptides Conflict: Score Tags	#	Score	Hits Mass	Mass error (p.	RT (mins)	Charge	Tags 🔻	Abundance	Conflic	t: Peptide ^
gi[254976387 (+6) o 33 0 3.75E+03	V 60	7 85.7	5 2926.431	0.0963	52.1	4		6.56E+05	0	TY:
gi 255101963 (+4) 0 0 0	9	L 72.5	5 2926.432	0.427	52	3		1.09E+07	0	S TY:
gi[2549/6385 (+1) 11 (2) 17 1.06E+03 [37] [3	✓ 56 ⁻	72 102	4 1800.888	-0.523	19	2		2.88E+04	0	S VE.
gi[2556567/6 (+1) 9 (0) 17 786	1	2 73.1	5 1144.634	-0.13	41.7	2		2.28E+07	0	GL:
gili26/0040/ 9 (1) 16 886 -	1	5 73.1	5 1144.634	0.371	40.7	2		2.69E+07	0	S GL -
Protein: gil254976387 cell surface protein (S-laver p	ecursor n	rotein) [Clo	stridium diffic	ile OCD-66c2	261					
Protein: gi[25101963 cell surface protein (S-laver p	ecursor p	rotein) [Clo	stridium diffic	ile QCD-63a4	421					
Peptide Views Protein Resolution		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,								
Conflicting proteins for feature 607	Peptides	of gi 25510	1963							
Accession Peptides Conflict: Protein Score	#	Score Hir	ts Mass I	Mass error (p	RT (mins)	Charge 1	Tags 🔻 Al	bundance (Conflict: I	Peptide S 🔦
	607	85.7 5	2926.431	0.0963	52.1	4	6.5	6E+05	0	TYNI E
🕥 gi 255101963 (+4) 0 0 0	12	73.1 5	1144.634	-0.13	41.7	2	2.2	8E+07	0 6	GILI
	16	73.1 5	1144.634	0.371	40.7	2	2.6	9E+07	0 🤇	GILI
	1097	67.8 5	1144.633	-0.576	38	2	2.6	9E+05	0 🧯	GILI
	91	72.5 5	2926.432	0.427	52	3 🧯	1.0	9E+07	0 🧯	TYNI TYNI
	1205	85 5	2926.429	-0.452	52	2	6.4	E+05	0 🧃) TYNI
	5672	102 4	1800.888	-0.523	19	2	2.8	8E+04	0 🧯	VET(
	2646	54.1 5	2765.483	0.682	63.3	4	6.2	5E+04	0 🧕	DSVI
۰ III ا	•	077 0	2022 251 III	1 71	C 0 3	-	a 21	CF. 07		
Protein options								Section	n Compl	lete 🌖

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 11 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.

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

Carl Edit protein building options	×
Protein grouping You can choose to to simplify the protein list by grouping together proteins which have been identified by similar sets of peptide sequences. A protein will be grouped with another if its identified peptide sequences are a subset of the identified sequences of the other.	
 Group similar proteins No protein grouping 	
Protein quantitation You can choose to have quantitation based on either only features which have no conflicting protein identifications or on all features identified on part of a protein.	
Quantitate from non-conflicting features Quantitate from all features	
OK Cance	2

With protein grouping switched on protein groups and the additional members are indicated by a bracketed number located after the Accession number. Taking **flagellin subunit** as an example, when the cursor is held over the accession number the group members (9 in total) appear in a tool tip.

Experiment design: A	c er applied		Create		Pept	tides	of gi 26	06820	17		Ø	No filte	er applied	a C	reate
Accession	Peptides	Conflict:	Score			#	Score	Hits	Mass	Mass error (p	RT (mins)	Charge	Tags	 Abundance 	Coi ^
🌍 gi 254976385 (+1)	11	0	1.06E+03			76	63.9	5	1669.888	-0.0985	42.4	3		4.82E+06	
🎯 gi 260682017 (+8) 📢	12	0	1.35F+03			138	103	5	1660 887	0.663	42.4	2		4.32E+06	(≡
🔮 gi 255656776 (+1)	gi 2606820)17 - flage	ellin subunit [Clo tive flagellar fila	ostridi	um dif	ficile C	D196] [Clostridiu	m diffic	ile 050-D50-	20111 0.407	22.9	2	Ē	3.11E+06	
🎯 gi 126700407	gi 1028148	37 - flagel	lin subunit FliC	[[Clos	tridium	n] diffi	cile]	munic	11E 030-F30-	0.167	38.9	2	Ē	5.6E+06	
🎯 gi 126697969 (+1)	gi 1266978	310 - flage	ellin C [Clostridio	um dit	fficile 6	30]	silal			0.157	38.9	3	Ē	3.12E+06	
🚳 gi 126697970	gi[1028146 gi[5668937	7 - flagelli	n [[Clostridium]	diffic	ile]	1) 01110	cilej			0.445	30.5	2	Ē	1.75E+06	
🎯 gi 126700790 (+1)	gi 7374572	26 - flagel	lin subunit FliC	[Clos	tridium	n] diffi	cile]			-1.19	34.8	3	Ē	7.66E+05	
🔮 gi 126698640	gi 2208629 gi 2208630	99 - FIIC [[)9 - FIIC [[Clostridium] dif	ficile]						-1.21	34.8	2	<u> </u>	7.54E+05	
< III	31		,		• [_							- F
No protein select Peptide Views Protein R Conflicting protein	ted Resolution NS			F	Peptic	les of	f conflict	ing pr	otein						
Accession	Peptide	s Conflic	t: Protein Score		#	S	core	Hits N	Aass N	lass error (p R	T (mins)	Charge T	ags	Abundance	Conflict
۲			þ		۲										F
Protein options													Se	ection Complete	• •

Having performed the conflict resolution with **Group similar proteins** and **Quantitate from non-conflicting features** now switch off the protein grouping.

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

Note: flagellin has **no unique** peptides (brackets after the peptides field) 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.

rotein 🕎 N	o filter appli	ied	Cre	ate		eptide	es of gi 2	606820	17		4	No filt	ter applie	ed	Create
Accession	Peptides	Conflict:	Score	Tags	•	#	Score	Hits	Mass	Mass error (p	RT (mins)	Charge	e Tags	 Abundance 	Coi
) gi 5668937	12 (0)	75	1.35E+03			V 33	68 84.7	4	1423.649	-0.431	22.7	2		7.25E+04	1
) gi 260682017 🛛 🧿	12 (0)	75	1.35E+03			V 44	9 93.3	5	1676.838	-1.21	34.8	2		7.54E+05	(E
) gi 126697810	9 (0)	65	1.07E+03			184	15 99	4	1407.656	0.593	38.7	2		1.28E+05	
) gi 10281485	8 (0)	60	867			V 47	7 47.4	5	1692.835	-0.238	20.7	3		5.61E+05	1
🕽 gi 357548088	7 (0)	56	821			V 88	8 109	5	1692.833	-1.02	20.7	2		4.67E+05	
) gi 10281487	8 (0)	55	887			V 30	0 43.8	5	1676.838	-1.19	34.8	3		7.66E+05	1
ail73745726	5 (0)	46	662			15	7 125	5	2317.115	0.167	38.9	2		5.6E+06	
2 2 1	- (-7														
gi 209570719 Protein: gi 260 Protein: gi 566 ptide Views Prote	10 (0) 10 (0) 10 0682017 68937 fla sein Resolution	37 flagellin agellin [816 n subunit [C [Clostridium	⊧ lostri n] diffi	dium icile]	✓ 17 < diffici	5 60.9 le CD196	5	2317.115	0.157	38.9	3		3.12E+06	•
gi[209570719 Protein: gi[260 Protein: gi[566 ptide Views Prot Conflicting prot Accession	10 (0) 10 (0) 10 10 10 10 10 10 10 10 10 10	37 flagellin agellin [on feature	816 n subunit [C [Clostridiun : 3358	lostri n] diffi	dium icile] Pej	✓ 17 diffici diffici diffici diffici diffici	5 60.9 le CD196 of gi 560	5 5] 58937 Hits N	2317.115 III	0.157	38.9	3 Charge	Tags	3.12E+06	Conf
gi[209570719 Protein: gi[260 Protein: gi[566 ptide Views Prote Conflicting prot Accession gi[35668937	10 (0) 10 (0) 10 10 10 10 10 10 10 10 10 10	37 flagellin agellin [on feature Conflict	816 n subunit [C [Clostridiun : 3358 : Protein Score 1.35F+03	clostri n] diffi	dium icile] Pej	✓ 17	5 60.9 le CD196 of gi 566 Score 84.7	5 5 5 5 8 9 3 7 Hits M 4 1	2317.115 III Mass N 423.649	0.157 1ass error (p	38.9 RT (mins) 22.7	3 Charge 2	Tags	3.12E+06 ▼ Abundance 7.25E+04	Conf ^
gij209570719 Protein: gij260 Protein: gij566 ptide Views Prot Conflicting prot Accession gij5668937 gij260682017	10 (0) 0682017 68937 fla tein Resolution pteins for Peptides 12 (0) 9 12 (0)	37 flagellin agellin [on feature Conflict 75 75	816 n subunit [C [Clostridium 3358 Protein Score 1.35E+03 1.35E+03	clostri n] diffi	dium icile] Per	 ✓ 17 ✓ 17 ✓ diffici ✓ diffici	5 60.9 le CD196 of gi 566 Score 84.7 93.3	5 58937 Hits M 4 1 5 1	2317.115 III Aass N 423.649 676.838	0.157 Mass error (p -0.431 -1.21	38.9 RT (mins) 22.7 34.8	3 Charge 2 2	Tags	3.12E+06 ▼ Abundance 7.25E+04 7.54E+05	€ Conf
gij209570719 Protein: gij260 Protein: gij566 ptide Views Prot: Conflicting prot Accession gij5668937 gij260682017 gij260682017 gij260682017	10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 12 (0) 9 (0) 10 (0)	37 flagellin ggellin [on feature Conflict 75 75 65	816 n subunit [C [Clostridium : 3358 Protein Score 1.35E+03 1.35E+03 1.07E+03	clostri n] diffi	dium icile] Per	 ✓ 17 <l< td=""><td>5 60.9 le CD196 of gi 566 Score 84.7 93.3 99</td><td>5 58937 Hits M 4 1 5 1 4 1</td><td>2317.115 III Aass N 423.649 676.838 407.656</td><td>0.157 Aass error (p -0.431 -1.21 0.593</td><td>38.9 RT (mins) 22.7 34.8 38.7</td><td>3 Charge 2 2 2</td><td>Tags</td><td>3.12E+06 ▼ Abundance 7.25E+04 7.54E+05 1.28E+05</td><td>Conf ▲ 6 6 6</td></l<>	5 60.9 le CD196 of gi 566 Score 84.7 93.3 99	5 58937 Hits M 4 1 5 1 4 1	2317.115 III Aass N 423.649 676.838 407.656	0.157 Aass error (p -0.431 -1.21 0.593	38.9 RT (mins) 22.7 34.8 38.7	3 Charge 2 2 2	Tags	3.12E+06 ▼ Abundance 7.25E+04 7.54E+05 1.28E+05	Conf ▲ 6 6 6
gij209570719 Protein: gij260 Protein: gij566 pride Views Prote Conflicting protection Accession gij5668937 gij260682017 gij126697810 gij126697810 gij12687810	10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 12 (0) 12 (0) 9 (0) 8 (0)	37 flagellin agellin [on feature 75 75 65 55	816 n subunit [C [Clostridium : 3358 ! Protein Score 1.35E+03 1.35E+03 1.07E+03 887	e	dium icile] Per	 ✓ 17 ✓ 17 ✓ diffici ✓ 0 ✓	5 60.9 le CD196 of gi 564 Score 84.7 93.3 99 47.4	5 58937 Hits M 4 1 5 1 4 1 5 1 5 1	2317.115 	0.157 Mass error (p -0.431 -1.21 0.593 -0.238	38.9 RT (mins) 22.7 34.8 38.7 20.7	3 Charge 2 2 2 3	Tags	 3.12E+06 3.12E+06 7.25E+04 7.54E+05 1.28E+05 5.61E+05 	► Conf ▲ 6 6 6
gij209570719 Protein: gij260 Protein: gij560 ptide Views Prote Conflicting prot Accession gij260682017 gij126697810 gij126697810 gij10281487 gij10281485	10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 11 (0) 12 (0) 9 (0) 8 (0) 8 (0) 8 (0)	37 flagellin ggellin [on feature Conflict 75 65 55 60	816 n subunit [C [Clostridium : 3358 : Protein Score 1.35E+03 1.35E+03 1.07E+03 887 867	e	dium icile] Pep	 ✓ 17 ✓ difficit <l< td=""><td>5 60.9 le CD196 of gi 566 Score 84.7 93.3 99 47.4 109</td><td>5 58937 Hits N 4 1 5 1 4 1 5 1 5 1 5 1</td><td>2317.115 III Aass N 423.649 676.838 407.656 692.835 692.833</td><td>0.157 Aass error (p -0.431 -1.21 0.593 -0.238 -1.02</td><td>38.9 RT (mins) 22.7 34.8 38.7 20.7 20.7 20.7</td><td>3 Charge 2 2 2 3 2</td><td>Tags</td><td>3.12E+06 ▼ Abundance 7.25E+04 7.54E+05 1.28E+05 5.61E+05 4.67E+05</td><td>Conf • • • • • • • • • • • • • • • • • • •</td></l<>	5 60.9 le CD196 of gi 566 Score 84.7 93.3 99 47.4 109	5 58937 Hits N 4 1 5 1 4 1 5 1 5 1 5 1	2317.115 III Aass N 423.649 676.838 407.656 692.835 692.833	0.157 Aass error (p -0.431 -1.21 0.593 -0.238 -1.02	38.9 RT (mins) 22.7 34.8 38.7 20.7 20.7 20.7	3 Charge 2 2 2 3 2	Tags	3.12E+06 ▼ Abundance 7.25E+04 7.54E+05 1.28E+05 5.61E+05 4.67E+05	Conf • • • • • • • • • • • • • • • • • • •
gij209570719 Protein: gij266 Protein: gij566 pride Views Prot. Conflicting prot. Accession gij260682017 gij126697810 gij126697810 gij10281487 gij10281485 gij10281485 gij357548088	10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 12 (0) 9 (0) 8 (0) 8 (0) 7 (0)	37 flagellin agellin [on feature Conflict 75 65 55 60 56	816 n subunit [C [Clostridium : 3358 : Protein Scorr 1.35E+03 1.35E+03 1.07E+03 887 867 821	clostri n] diffi	dium icicile] Pep V	 ✓ 17 ✓ difficit ✓ difficit ✓ 17 ✓ 17 ✓ 188 ✓ 300 	5 60.9 Ie CD196 of gi 566 Score 84.7 93.3 99 47.4 109 43.8	5 58937 Hits N 4 1 5 1 4 1 5 1 5 1 5 1 5 1	2317.115 III Aass N 423.649 676.838 407.656 692.835 692.835 692.833 676.838 8	0.157 Aass error (p -0.431 -1.21 0.593 -0.238 -1.02 -1.19	38.9 RT (mins) 22.7 34.8 38.7 20.7 20.7 20.7 34.8	3 Charge 2 2 3 2 3 2 3 3 5 5 5 5 5 5 5 5 5 5 5 5	Tags	3.12E+06 ▼ Abundance 7.25E+04 7.54E+05 1.28E+05 5.61E+05 4.67E+05 7.66E+05	Conf ▲ 6 6 6 6 6 6
gij209570719 Protein: gij266 Protein: gij566 pride Views Prot. Conflicting prot Accession gij26068937 gij260682017 gij126697810 gij10281487 gij10281487 gij10281485 gij10281485 gij10281485 gij10281485 gij10281485	10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 10 (0) 12 (0) 9 (0) 8 (0) 8 (0) 7 (0) 5 (0)	37 flagellin gellin [on feature Conflict 75 75 65 55 60 56 46	816 n subunit [C [Clostridium 33558 Protein Scorr 1.35E+03 1.35E+03 1.35E+03 1.07E+03 887 867 821 662	e www.www.www.www.www.www.www.www.www.ww	dium icile] Pep V V V	 ✓ 17 ✓ 17 ✓ 17 ✓ 3358 ✓ 449 1845 ✓ 477 888 ③ 300 157 	5 60.9 IE CD196 of gi 560 Score 84.7 93.3 99 47.4 109 43.8 125	5 58937 Hits M 4 1 5 1 5 1 5 1 5 1 5 1 5 1 5 2	2317.115 III Aass N 423.649 676.838 607.656 692.835 692.835 692.833 676.838 317.115	0.157 Aass error (p -0.431 -1.21 0.593 -0.238 -1.02 -1.19 0.167	38.9 RT (mins) 22.7 34.8 38.7 20.7 20.7 20.7 34.8 38.9	3 Charge 2 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3	Tags	3.12E+06 ▼ Abundance 7.25E+04 7.54E+05 1.28E+05 5.61E+05 4.67E+05 7.66E+05 5.6E+06	Conf ▲ 6 6 6 6 6 6 6 6 5
gij209570719 Protein: gij266 Protein: gij566 ptide Views Prot. Conflicting prot Accession gij5668937 gij260682017 gij126697810 gij10281487 gij10281487 gij10281485 gij10281485 gij137548088 gij73745726	10 (0) 10 (0) 10 0682017 68937 fize leiein Resolution otelins for Peptides 12 (0) 9 (12 (0) 9 (0) 8 (0) 8 (0) 7 (0) 5 (0)	37 flagellin gellin [on feature Conflict 75 65 55 60 56 46	816 n subunit [C [Clostridium 33558 Protein Score 1.35E+03 1.35E+03 1.07E+03 887 867 821 662	e E E E E E E E E E E E E E E E E E E E	dium icile] Peŗ V V V	✓ 17 ✓ 17 ✓ diffici ✓ diffici × 4 × 4 × 449 × 1845 × 477 × 888 × 300 × 157 × 175	5 60.9 IE CD 196 of gi 560 Score 84.7 93.3 99 47.4 109 43.8 125 60.9	5 58937 Hits N 4 1 5 1 4 1 5 1 5 1 5 1 5 1 5 2 5 2	2317.115 III Alass N 423.649 676.838 407.656 692.835 692.833 676.838 317.115 317.115	0.157 Aass error (p -0.431 -1.21 0.593 -0.238 -1.02 -1.19 0.167 0.157	38.9 RT (mins) 22.7 34.8 38.7 20.7 20.7 20.7 34.8 38.9 38.9 38.9	3 Charge 2 2 3 2 3 2 3 2 3 2 3 2 3 2 3 2 3 3 2 3	Tags	3.12E+06 ▼ Abundance 7.25E+04 7.54E+05 1.28E+05 5.61E+05 4.67E+05 7.66E+05 5.6E+06 3.12E+06	Conf 6 6 6 6 6 6 5 5

Now set the Protein Options back to Group similar proteins

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 by right clicking on the table and selecting **Quick Tags...**

🔼 LC-MS Tutori	ial -	Proge	nesis I	LC-MS														
File					. .		_							- ·				
Import Data	Alig	nment		Filtering	Design 1	nent Setup	P	icking	Statistics Pe	entity ptides	Identifica	ie Kes ations Cor	nflicts	Proteins	Statistics	Report	nonli	near
•								•	•		0			•				M I C S
W No filter	apı	olied (Cr	eate	Search			م ر						Prote	in options	AC	•	Ø Help ▼
ssion	Pe	ptide o	ount	Confidence	e score	Anov	a (p)	Tag 🛛 💌	Max fold change	Highest	t Mean	Lowest Mear	n Descript	ion				
254976387 (+6)	33			3.75E+03		5.33E	-06	4	8.06	С		А	cell surfa	ice protein (S-layer precurs	or protein) [(Clostridium diff	icile QCD-6 *
126701103	1			51.7		7.17E	-06	4	6	А		с	ribose-5	-phosphate	isomerase 2 [C	lostridium di	fficile 630]	
54781345 (+1)	5			398		1E-05	5	4	8.26	А		с	(R)-2-hyd	droxyisocap	roate dehydrog	genase [Clost	ridium difficile	ATCC 9689 [≡]
126699971	5			334		1.2E-	05	4	11.8	А		с	thioredo	xin 2 (Trx2)	[Clostridium di	fficile 630]		
126697690	5		An	ova p-value	≤ 0.05		-05	4	6.01	А		с	ferredox	in/flavodoxi	n oxidoreducta	se subunit g	amma [Clostrid	ium difficile
126700634	3		Ma	ax fold chang	ge≥2		-05	4	6.41	А		с	PTS syste	em mannos	e-specfic trans	orter subuni	t IIB [Clostridiu	m difficile €
126699140	1		Ne	w tag			-05	4	12.4	А		с	ferredox	in-NADP(+)	reductase sub	unit alpha [Cl	ostridium diffic	ile 630]
126699940	1		Qu	iick Tags		•		Anova p-v	alue	A		с	transkete	olase, N-ter	minal (Sedohep	tulose-7-ph	osphate:D-glyce	eraldehyde
260682017 (+8)	12	1	Edi	it taos				Max fold o	hange	А		с	flagellin	subunit [Clo	stridium diffici	le CD196]		
126697752	5			498		3.43E		Sequence.		А		с	NAD-spe	cific glutan	nate dehydroge	nase [Clostri	dium difficile 6	30]
126698435	1			48.9		4.38E		Modificati	on	С		А	ABC tran	sporter olig	opeptide-fami	y extracellula	r solute-bindin	g protein [
126697684	3			158		5.19E		Peptide ta	gs contain	A		с	phospha	te butyryltr	ansferase [Clos	tridium diffic	ile 630]	
126697583	1			79.3		5.3E-	05	4	4	А		с	DNA bin	ding protei	n [Clostridium d	difficile 630]		
126701233	1			44.9		6.3E-	05	4	3.63	А		с	ferredox	in [Clostridi	um difficile 630]		
384359782 (+8)	21			1.88E+03		7.52E	-05	4	3.43	с		A	hemaggi	utinin/adhe	sin [Clostridiur	n difficile BI1]	-
٠																		• •
Selected J	Dro neas	otein	tł <u>ts</u>	nioredox	din 2	(Trx	2)	Clostri	dium difficil	le 630]							
							А								С			
h Normalised Abun 15.5 15.0 14.5 14.0							¥¢¢.											
4.0 H															N			
4																		Þ
Total number	of d	isplay	ed pr	oteins: 1	35												Section Com	iplete 🌛

As an example let us explore Thirodoxin 2.

The table indicates that this protein is most highly expressed in Condition A by 11.8 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.

🔼 LC-MS Tutorial - Progenesis LC-MS															
	LC-MS	Tutoria	I - Progenes	is LC-MS											
	File		Daviaw	Eve	neriment Peui	aw Dask Da	antide k	dentify	Dafina	Perolya	Daviaw	Protein			
Im	port Da	ta A	lignment	Filtering Des	ign Setup Pi	icking Sta	atistics Pe	eptides Io	dentification	is Conflicts	Proteins	Statistics	Report	nonl	inear
					0	•		•			•			D Y N A	M I C S
6	Deals	Dro	tain. th	-i-rodovin 7	(Tey2) [C]	tridium	difficile	4201							
Ľ	Васк	Pro	item: u		(11, x2) [Ci	ostriaium	dimeile	030]							
Acc	ession:	😻 gi	126699971												
Σ	#	Score	Anova (p)	Max Fold Change	Highest Mean	Lowest Mean	Abundance	m/7	Charge	Retention Time	Mass error (ppm)	Peptide S	equence	Modifications	
•	263	69.09	0.000171	16.4	A	с	2.285E+06	603.3518	2	47.914	0.04	S VLGLP	ΤΜΑΙΥΚ		
•		45.25	9.54E-06	10.1				711.0242		41.823		S VDEV	TKDDATVPNIENMIK		
•			8.33E-06	6.27				1066.0325		41.785	-0.43	S VDEV	TKDDATVPNIENMIK		
•	501	48.68	8.46E-05	12			4.41E+05	611.349		39.811		🕲 VLGLP	PTMAIYK	[7] Oxidation (M)	
•	1429	60.28	0.000947	16.7			2.289E+05	730.3582		46.185	-0.33	S DDAT	VPNIENMIK		
•	1730	44.06	0.000126	12.4	А	С	2.166E+05	716.3561	3	32.622	0.03	S VDEV	IKDDATVPNIEN <mark>M</mark> IK	[17] Oxidation (M)	
						Sta	andardise	d Expres	ssion Pr	rofiles	•				
			A						с						
	20														
2	1.5					-		-							
andar	1.0						\leq								
pet 1	0.5			8											
Mormal															
palb															
Standa	-0.5														
	Ĩ														

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

Modified proteins can be located by specifically searching for proteins containing modified peptides. Use the **Back** button to return to the Proteins List and right click on it and select **Modification** from the list of **Quick Tags**.

🞯 gi 54781345 (+1)	5	398		1E-05	-	8.26
😵 gi 126699971	e	224		1.2E-05	4	11.8
🔇 gi 12669769 💆	Anova p-value	≤ 0.05		1.32E-05		6.01
🔇 gi 12670063	Max fold chang	ezz		1.63E-05	4	6.41
🔇 gi 12669914	New tag			1.92E-05	4	12.4
🎯 gi 12669994	Quick Tags	•		Anova p-va	lue	
🔇 gi 26068201 😭	Edit tags			Max fold ch	ange	
🔇 gi 126697752	5	498		Sequence		
🔇 gi 126698435	1	48.9		Modificatio	n	
🔮 gi 126697684	3 158			Peptide tag	s contain	
🎯 gi 126697583	1	79.3		5.3E-05	4	4

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**.

🔼 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

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

A 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 tans: Show proteins that have all of these tans:											
Available tags:	Show proteins that have all of these tags:										
 Anova p-value ≤ 0.05 (124 proteins) Max fold change ≥ 2 (99 proteins) 	Modification with Oxidation (M) (9 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.

LC-MS Tutorial - Progenesis LC-MS												
File Review Experiment Review Peak Peptide Identify Refine Resolve Review Protein Import Data Alignment Filterigo Design Setup Picking Statistics Pentides Identifications Conflicts Proteins Statistics Report Process												
Import Data Alignn	ew nent Filterin	Experiment Design Setup	Review Pe Picking	eak Pep) Stat	otide Identify istics Peptides	Refine Identification	Resolve ns Conflicts	Review Protein Proteins Statistics Report	nonlinear			
Too Channel	- 4								_ 🕜 Help 🔻			
proteins may be	e hidden Ec	Search		ې				Protein options	· · ·			
Accession	Peptide count	Confidence score	Anova (p)	Tag 🛛 💌	Max fold change	Highest Mean	Lowest Mean	Description				
🕥 gi 254976387 (+6)	33	3.75E+03	5.33E-06	٩	8.06	С	А	cell surface protein (S-layer precursor proteir	n) [Clostridium difficile QCD-66c.			
💱 gi 126699971	5	334	1.2E-05	٩	11.8	А	С	thioredoxin 2 (Trx2) [Clostridium difficile 630]	1			
🎯 gi 126697690	5	491	1.32E-05	٩	6.01	Α	С	ferredoxin/flavodoxin oxidoreductase subuni	t gamma [Clostridium difficile 6:			
🔇 gi 260682017 (+8)	12	1.35E+03	3.33E-05	۷	3.87	Α	С	flagellin subunit [Clostridium difficile CD196]				
🔇 gi 384359782 (+8)	21	1.88E+03	7.52E-05	۲	3.43	С	Α	hemagglutinin/adhesin [Clostridium difficile	BI1]			
🎯 gi 126700790 (+1)	8	688	9.71E-05	۲	2.41	С	А	enolase (2-phosphoglycerate dehydratase) (2	2-phospho-D-glycerate hydro-ly			
🔇 gi 126700129	3	267	0.000199	۹	2.4	Α	С	translation inhibitor endoribonuclease [Clost	ridium difficile 630]			
🔇 gi 126697654	3	215	0.000219	۲	2.53	A	С	30S ribosomal protein S8 [Clostridium difficil	e 630]			
🔇 gi 54781347	4	341	0.00222	۹	2.47	Α	с	2-hydroxyisocaproate-CoA transferase [Clost	ridium difficile ATCC 9689]			
•									+			
Selected prote	ein: thiore	doxin 2 (Trx	2) [Clo	stridiun	n difficile 63	80]						
			А					С				
A C												
Total number of disp	played proteins	: 9							Section Complete 🥑			

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

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

Stage 13: 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 14: 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.**

Filter the proteins	
Create a filter Show or hide proteins based on a selection of the create the filter. For more guidance, please see t	ir tags. Move tags to the appropriate boxes to he <u>online reference</u> .
Available tags:	Show proteins that have all of these tags:
 Anova p-value ≤ 0.05 (124 proteins) Max fold change ≥ 2 (99 proteins) 	Modification with Oxidation (M) (9 proteins) Show proteins that have at least one of these tags: Hide proteins that have any of these tags:
Clear the filter	OK Cancel

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

- 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 9.
- 2. Expand the various Report Design options (by default they are all selected)
- 3. Un-tick as shown below
- 4. Click Create Report

🔼 LC-MS Tutorial - F	Progenesis LC	C-MS										
File												
Revi	ew Eiler	E)	periment	Review F	Peak Peptide	Identify	Refine	Resolve	Review	Protein	Papart	nonlinger
Alight	nent ritt	enng De		PICKIN		Peptides			Proteins			
			-		-	-			-	-		
Protein report Pep	tide report											
Proteins					Report Design	n						
Experiment design:	AC		•]								
🖘 Tag filter ar	onlied				Title LC-MS T	utorial						
proteins may	y be hidden		Edit		Select the section	s you wish to	o include in your	report:				
Accession	Anova (p)	Fold	Tag		💎 🔳 Overvie	w run						
gi 254976387 (+6)	5.33E-06	8.1	4	c	🔍 🔲 Data pr	ocessing me	thods					
gi 126699971	1.2E-05	11.8	4	tl	🔍 🔲 Experim	ent design						
gi 126697690	1.32E-05	6.0	4	fi	A Protein	report						
gi 260682017 (+8)	3.33E-05	3.9	4	fl	Include tables s	, howing prot	ein abundances	and peptides	identified for	each protein		
gi 384359782 (+8)	7.52E-05	3.4	الله الله	h								
gi 126700790 (+1)	9.71E-05	2.4	الله الله ال	e	🔽 Proteir	ı table						
gi 126700129	0.000199	2.4	4	tı	📝 Peptid	e tables						
gi 126697654	0.000219	2.5	الله الله	3	Protein	Details						
gi 54781347	0.00222	2.5	4	2	Roports the full	dotails of ou	on protain whic	h matchac va	ur current filte			
					Reports the full	details of ev	ery protein whic	in matches yo	or corrent line			
					🔽 Tags							
					Expres	sion profile						
					👻 🔲 Feature	table						
					👻 🔳 Feature	details						
					Create Rep	ort						
<												
				F								

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

LC-MS Tutorial											
Experiment: LC-MS Tutorial											
Report created: 18/02/2013 0	09:39:30										
Proteins											
Protein building options											
Protein grouping Group similar	r proteins										
Protein quantitation Using only fea	atures with no protein co	nflicts									
Accession Peptides Score Ar	nova (p) Fold Tags Descr	ription						Average Normalise	d		
*								Abundances			
gi 254976387 33 3752.13 5	i.33e-006 8.06 🔔 cell s	urface protein (S	layer p	recursor prot	tein) [Clo:	stridiun	n difficile	5.35e+007 4	1.32e+008		
gi 384359782 21 1879.25											
gi 260682017 12 1352.16	<u>gi 126699971</u>										
gi 126700790 8 688.34	thioredoxin 2 (Trx2) [Clostridium d	ifficile	e 630]							
gi 126697690 5 490.99	5 peptides										
gi 54781347 4 341.12	Sequence	Feature Score	e Hits	Mass	Charge	Tags	Conflicts	Modifications	In quantitati	on Average Norma	ised Abundances
gi 126699971 5 333.68		1 (20 (0.2		4 450 7040						A	C
gi 126700129 3 266.82	DDATVPNIENMIK	1429 60.2	3 5	1458.7019	2		0		yes	2.29e+005	1.37e+004
g1 126697654 3 214.79	VDEVTKDDATVPNIENMIK	366 45.2	5 4	2130.0508	3		0		yes	1.40e+006	1.38e+005
	VDEVTKDDATVPNIEN	1730 44.0	5 2	2146.0465	3		0	[17] Oxidation (M)	yes	2.17e+005	1.74e+004
	VLGLPTMAIYK	263 69.0	9 5	1204.6890	2		0		yes	2.29e+006	1.40e+005
	VLGLPTMAIYK	Accessio	n gi 1	2669997	'1						
	Tags										
	Most abundant	Description	n thior	edoxin 2 (T	rx2) [Cl	ostridi	um diffici	le 630]			
	Anova p-value ≤ 0.05	Peptide	s 5								
	Max fold change ≥ 2	Anov	a 1.20€	⊿8 ≥-005							
	Significantly up in C	. Fol	11.77	1							
	Up regulated in C	. 🦷	Anov	a p-value ≤	0.05						
	Op regulated in A		Max 1	old change	e≥2						
-			Modi	ication wit	th Oxida	tion (/	۸)				
					А				с		
		e			×						
		up 16.0									
		nq 15.5									
		15.0 -									
		E 14.5									
		2 <u>4</u> 14.0									
		SJA 13.5							<u>M</u>		
		-								+	

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.

Exporting Protein Data

Protein data can be exported in a csv file format. You can either export the **Protein and/or peptide measurements** using the options in the File Menu which are available at the Review Proteins stage.

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

First set the tag filter as shown below. Then select Export Protein Measurements from the File menu.

Filter the proteins Create a filter Show or hide proteins based on a selection of create the filter. For more guidance, please se	their tags. Move tags to the appropriate boxes to the the <u>online reference</u> .
Available tags: Anova p-value ≤ 0.05 (124 proteins) Max fold change ≥ 2 (99 proteins)	Show proteins that have all of these tags: Modification with Oxidation (M) (9 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 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.





Excel will open displaying the exported protein measurements

	A1 .	• (• f x							Y
	А	В	С	D	E	F	G	l l	Ē
1						Normalize	ed abundar	nce	
2						Α			
3	Accession	Peptide count	Anova (p)	Max fold change	Description	A1	A2	A3	
4	gi 254976387;gi 112	33	5.33E-06	8.063180835	cell surface protein (S-layer precursor protein) [Clostridium	55837799	57754081	4702	=
5	gi 126699971	5	1.20E-05	11.7737882	thioredoxin 2 (Trx2) [Clostridium difficile 630]	4957161	5277038	519	
6	gi 126697690	5	1.32E-05	6.011328857	ferredoxin/flavodoxin oxidoreductase subunit gamma [Clos	1778402	2084321	168	
7	gi 260682017;gi 102	12	3.33E-05	3.872261726	flagellin subunit [Clostridium difficile CD196]	27038907	24794152	3013	
8	gi 384359782;gi 209	21	7.52E-05	3.430214543	hemagglutinin/adhesin [Clostridium difficile BI1]	2614529	2673778	313	-
9	gi 126700790;gi 296	8	9.71E-05	2.414242661	enolase (2-phosphoglycerate dehydratase) (2-phospho-D-gl	404610.9	387979.1	441	
10	gi 126700129	3	0.000198776	2.398837692	translation inhibitor endoribonuclease [Clostridium difficile	5557391	5446163	618	
11	gi 126697654	3	0.0002192	2.532221342	30S ribosomal protein S8 [Clostridium difficile 630]	1014579	1149285	12(
12	gi 54781347	4	0.002217665	2.469813207	2-hydroxyisocaproate-CoA transferase [Clostridium difficile	1581089	1205546	138	
13									•
14 -	My Oxidised	Proteins 🦯 🖏 🦯						•	

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.



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 the Progenesis LC-MS and the extra MS/MS scans added to the experiment.

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

	Review		Experiment	Review Peak	Peptide	Identify	Refine	Resolve	Review	Protein	
Import Data	Alignment	Filtering	Design Setup	Picking	Statistics	Peptides	Identifications	Conflicts	Proteins	Statistics	Report

First return to Review Peak Picking using the Workflow icons.



Available tags:	Show features that have all of these tar
Most abundant (2342 features) Max fold change ≥ 2 (9881 features) Significantly up in C (6892 features)	Anova p-value \$ 0.05 (7965 features) No MS/MS data (8390 features) Show features that have at least one of these tags: Hide features that have any of these tag

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

Filter the table so that it is only showing features with an **Anova p-value** \leq **0.05** and **No MS/MS data**. To do this click **Create** and drag the two tags on to the Show features that have all these tags and click OK.

Select all of the features that are displayed, right click and create a New tag called Inclusion 1

Revi	ew Feature	es Experiment design	AC	•											
7	Tag filter app features may	plied be hidden	Edit												
#	Anova (p)	Fold Tag	 Notes 	Highest	Lowest	m/z	z	Mass	RT (mins)	RT window	Abundance	Intensity	Max CV	MS/MS	-
4022	1.13E-12	Infinity 📑	Add note	С	Α	1526.7795	2	3051.544	39.156	0.204	4.79E+04	2.76E+05	1.3	0	
141	3.125 12	T	Add note	С	Α	901.2218	2	1800.429	39.344	1.13	4.07E+06	1.43E+07	2.35	0	
12141	1E- 🧮 🕺	lost abundant	Add note	А	С	980.4783	2	1958.942	39.245	0.232	1.05E+04	5.4E+04	1.95	0	
3818	1.1	nova p-value ≤ 0.05	Add note	С	А	1441.1479	5	7200.703	39.344	0.484	2.51E+05	2.96E+05	173	0	
7793	1.4: M	lax fold change ≥ 2	Add note	А	С	847.7614	3	2540.262	55.142	0.202	8.84E+03	7.15E+04	2.09	0	
11083	1.6	gnificantly up in C	Add note	С	А	1011.0473	2	2020.08	41.307	0.22	5.45E+03	5.16E+04	2.09	0	
8201	1.7:	o MS/MS data	Add note	С	А	1002.0215	2	2002.028	32.829	0.21	8.65E+03	9.91E+04	2.19	0	
10123	2.7: N	ew tag	Add note	С	А	1157.2779	3	3468.812	53.605	0.368	2.03E+04	9.39E+04	2.68	0	
7553	2.7: Q	uick Tags	Add note	А	С	1207.0705	2	2412.127	34.551	0.566	3.91E+04	8.4E+04	2.86	0	
5343	5.98 🔗 Ed	dit tags	Add note	А	С	1372.4025	4	5485.581	32.829	0.379	2.25E+05	2.32E+05	3.96	0	
6314	6.4. Ad	dd to clip gallery	Add note	С	A	1293.8406	4	5171.333	30.462	0.479	4.8E+04	1.62E+05	3.56	0	
10970	6.75E-11	Infinity	Add note	Δ	C	691.9449	3	2072.813	33.464	0.566	8.85E+03	6.09E+04	3.09	0	Ŧ

Call the new tag Inclusion_1.

Create new tag	X
Inclusion_1	
	OK Cancel

Now use the new tag to filter the table to display only those features that show a Significant Change and **DO NOT** have any MS/MS spectra.

Available tags:	Show features that have all of these t
Most abundant (2959 features) Mark fold change 2.2 (10864 features) Significantly up in C (4155 features) Anova p-value 2 0.05 (61.8 features) No MS/MS data (9488 features)	Inclusion_1 (5305 features) Show features that have at least one of these tags: Hide features that have any of these tags:



Then select **Export Inclusion List...** from the file menu and select the appropriate format.

Finally save the file to an appropriate location

Kara Export Inclusion	on List			×
Save in:	Inclusion Lists	•	G 🦻 📂 🛄 -	
(Ba	Name	Date modified	Туре	Size
Recent Places		This folder is empt	y.	
Desktop				
Andy Borthwick				
Computer				
Computer	•	III		Þ
Network	File name: Inclusion_1		-	Save
	Save as type: Thermo Finn	igan inclusion list files (*.	txt) 👻	Cancel
				Help

Inclusion list retention time windows
Do you want to widen the retention time windows
by 0.00 🚔 minutes?
Yes No

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

If you require further information on the inclusion list file formats then click Help.

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 LC-MS, from Alignment through Analysis to generating lists of interesting features 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 SameSpots using your own runs and explore the Progenesis LC-MS workflow please go to Appendix 3: Licensing Runs (page 60).

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

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

To create a new experiment with your files select **New** give your experiment a name. Then select data type, the default is 'Profile data'.

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

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

Create New LC-MS Experiment	
Create a new label-free experiment named:	
LC-MS_Tutorial	
CData type	
Profile data	
Centroided data	
Resolution (full width at half maximum) 2000	
Machine type	High resolution mass spectrometer
High resolution mass spectrometer	e.g. Thermo LTQ Orbitrap, Bruker Maxis, Waters SYNAPT, Agilent QTOF, AB SCIEX TripleTOF
	Thermo LTQ Iontrap in Enhanced mode.
Experiment folder	
Save experiment in the same folder as the run data	Low resolution ion trap
Choose an experiment folder	e.g. Bruker HCT, Bruker HCT Ultra, Thermo LTQ XL
Brows	Thermo FT-ICR
Create experiment Canc	el

Select the 'Import Data file format', in this example they are mzXML files

Then locate your data files using Import...

Import Data Import your run data Select one of the available data formats then click the Import button: Import your run data Format Import your run data Start tig wersion: 1.0.4780.30257 Import Version: 1.0.4780.30257 below the Totorion (c. 4623.23396) Version: 1.0.4780.30276 dates Iook at format Int process Version: 1.0.4780.30276 mt process Version: 1.0.4780.318798 w, checking mt the version: 1.0.4780.318798 mt process version: 1.0.4780.318798 mt process version: 1.0.4780.318798 mt process version: 1.0.4780.318798 mt process versi	<pre>my comp comp comp comp comp comp comp comp</pre>	File mport Data	Review Alignment	Filtering	Experiment Design Setup	Review Peak Picking	Peptide Statistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics	Report	nonline
Whitey Bucker Datonics (.d) he button Version: 1.0.46790;20371 dates Deblow ti Uname Image: Construction of the	White Buker Datonics (.d) he button • Belowt Wersion: 1.0.4780.30276 dates • Deversion: 1.0.4780.20271 in • Alig Wersion: 1.0.4780.20271 • Alig Wersion: 1.0.4780.20271 • Alig Mersion: 1.0.4780.20271 • Alig Mersion: 1.0.4780.20271 • Alig Mersion: 1.0.4780.20271 • Alig Mersion: 1.0.4780.20271 • New folder Incom • Version: 1.0.4780.20271 Intermore Raw Files • Version: 1.0.4780.18798 w, checking • Wersion: 1.0.4703.18798 w, checking • Desktop • Clamani • Wersion: 1.0.4703.18798 • Muicit • More unus Scarch • No runs have been imported yet No runs have been imported yet	mport Dat Import Select o click the Format: @ Abo	ta your run dat ne of the avai Import butto mzXML files Waters .Raw Version: 2.0 Agilent Q-TC Version: 2.0	ta lable data fo n: folder .4780.458955 IF/TOF .4780.30257	rmats then Import nload other	5	Select files	« LC-MS)	LC-MS4.1 ▶			Search LC-MS 4	.1	e He
for any Version: 1.0.4703.18798 t might affect at Uterror FT-CR. Rw Files upported runs: Search P 6 items selected Date modified: 20.057208 08:25 Size: 2.08 GB	for any Version: 1.0.4703.1878 t might any runs affect at Them FF1CR. Raw Files any runs any runs nported runs: Search Image: Search and the search	While ye below to • Sel • Det • Alig Review Look at	Bruker Dalto Version: 1.0 mzML Files Version: 1.0 mzXML files Version: 1.0 NetCDF files Version: 1.0 Thermo .Ray	nics (.d) .4780.30276 .4699.20371 .4728.23855 .4623.23369 v Files	he button idates fates un ent process w, checking)	Organize - Favorites Deskto Downly Secent	New folde p oads .Places	econo an e	Name	-MS_TULORIALAR -MS_Test Inclu .mzxml .mzxml .mzxml	Ialysis sion	≣ ▼ []]	
	File name: "C3.mzml" "A1.mzml" "A2.mzm Dpen Cancel	for any affect a that hay nported run:	Thermo FT-I Version: 1.0	CR .Raw File	s t might any runs		E Desktop Librarie Docu Musi 6	es iments ic items selecti	ed Date modified Size	20/05/2008 0 2.08 GB	.mzxml .mzxml .mzxml .mzxml 8:25	1		

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: as the loading process starts you can also start the automatic alignment before the loading has completed. This is a 2 stage process that involves the selection of an Alignment Reference (either automatically or manually then the automatic alignment of all your runs to this Reference run.

Click Start alignment process to start the automatic alignment of your runs.

🔼 LC-MS_Tutorial - Prog	enesis LC-MS										
File	-	Experiment	Review Peak	Peptide	Identify	Refine	Resolve	Review	Protein	D 1	
Import Data Alignment	Filtering	Design Setup	Picking	Statistics	Peptides	Identifications	Conflicts	Proteins	Statistics	Report	nonlinear
Import Data	-			-	-	-			-	-	🕢 Help 🔻
A Import your run	data		A1				Actic	ons •			
Select one of the a click the Import bu	vailable data i itton:	formats then							About th • MS/I	iis run MS count: -	
Format: mzXML fi	les	Import							 Peak Tota 	count: - l ion intensi	tv: -
About this date	a format 🖊	Download othe	rs						• Mas	<pre>ked areas :</pre>	none
2 Start the alignm While your runs ar	e nt process e importing, c	lick the button									
 below to: Select alignment 	ent reference (candidates									
 Determine the Align all runs 	best of the cito that referen	andidates ace run									
	Start ali	ignment proces	;								
Review the chro	matography										
J Look at all of the r for any sample-run	uns in the list	below, checkin s that might	9								
affect analysis. Rig that have significa	ht-click to ren nt problems.	nove any runs	M/	aiting	n to	he im	nort	be			
Imported runs: Search		Q	VVC	The ion in	Itensity	map will be	e shown h	ere			
			1		,						
Importing Per	iding F	Pending									
A1	A2	A3									
Pending Per	iding F	Pending									
C1	C2	C3									
										_	
			r							5	Section Complete 🏵

Progenesis LC-MS provides three methods for choosing the alignment reference run, as seen below:

Start Alignment 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 e.
< Back	Next > Cancel

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

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.

🔼 Start Alignmen	t Processing			- • •
Select your alig To mark a run fo use the button b	gnment referen or assessment as below to mark mu	an alignment re an alignment re ultiple runs at on	eference candidate ice.	e, double-click it or
Mark selected r	uns Clear all		Search	Q
\$ 1.				*
æ	e	æ	æ	
A1	A2	A3	C1	
C2	e C3			·
6 of 6 runs mark	ced as candidates	5		
		<	Back Next	> Cancel

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.

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).

Once you have selected how to handle the choice of Reference run you will now be asked if you want to Align your runs automatically or manually.

Start Alignment 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?
< Back Finish Cancel

Select automatically and click finish.

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

Alignment Processing (7%)			
Alignment processing Current step: Choosing an alignment referen	ce		
✓ Imported runs: 6/6 [©] Reference run: pending		$\overline{\mathbf{V}}$	
Aligned runs: pending	Alignment Processing (80%)	X	
-	Alignment processing Current step: Aligning 'A3'		
	✓ Imported runs: 6/6 ✓ Reference run: C1 ② Aligned runs: 3/5		
	-	Alignment Complete	
		Alignment processing con Time taken: 2m 43s	nplete.
		✓ Imported runs: 6/6	
		 Aligned runs: 5/5 	
		Revie	w chromatography Review alignment 🌖

Once the Reference run has been chosen the automatic alignment is performed. As the whole process proceeds you get information on what stage has been performed and also the % of the process that has been completed.

Note: At this stage you have the option to Review the Chromatography or go straight to the review of the Automatic Alignment of your data.

Review Chromatography

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

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



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



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

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

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

Now move to the next stage in the workflow (page 9 in this user guide) by clicking Section Complete.

Appendix 2: 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 identifyable '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 features in the replicate group.

To do this select Masked areas from selected run on the bottom left of the screen.



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



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

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



Appendix 3: Licensing runs (Stage 3)

When setting up a **New experiment** if you are evaluating Progenesis LC-MS 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 Nonlinear Sales Person or purchase a licence code directly from Nonlinear.

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

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

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

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

LC-MS_Tutorial - Progenesis LC-MS									
File Review Import Data Licensing Alignment Filtering	Experiment Design Setup	Review Peak Picking	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Report	nor D y N	linear
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	D:\Customer E	ata\LC-MS\LC-N	4S 4.1∖A1.mzn	ld				Unlicensed	
licence code which can be obtained from a sales	D:\Customer E)ata\LC-MS\LC-N	IS 4.1∖A2.mzn	ld				Unlicensed	
representative.	D:\Customer E)ata\LC-MS\LC-M	IS 4.1\A3.mzn	ld				Unlicensed	
Once licensed, your runs can be analysed on	D:\Customer E)ata\LC-MS\LC-N	AS 4.1\C1.mzn	ld				Unlicensed	
any installation of the software. The licence is	D:\Customer E)ata\LC-MS\LC-N	IS 4.1\C2.mzn	ld				Unlicensed	
automatically included when archiving an experiment.	D:\Customer Data\LC-MS\LC-MS 4.1\C3.mznld Unlicensed								
If your runs have been licensed on another computer, <u>click here</u> to make the licences available on this computer.									
If you have one, you can <u>open a licence file</u> to install.									
If you have just installed a dongle, <u>click here</u> .			F	lun licence code:	XXX-XXXX-XX	XX-XXX		Use Licenc	e Code
								Section (Complete 🏵

A message confirming successful installation of your licences will appear.

Installation	n complete	×
i	Successfully installed licences for Progenesis LC-MS.	
	ОК	

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

Appendix 4: Manual assistance of Alignment

Approach to alignment

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

- 1. Click on Run C2 in the **Runs** panel, this will be highlighted in green and the reference run (C1) will be highlighted in magenta.
- 2. You will need approximately 5 10 **alignment vectors** evenly distributed from top to bottom of the whole run.
- 3. First either drag out a focus area or set the focus area to **8 or 16** using the Focus grid size on the bottom left of the screen.



Click on an area (see below) in the **Whole Run** view (C) to refocus all the windows. Adjust Contrast as required.



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



5. As you are holding down the left mouse button drag the green feature over the corresponding magenta feature of the reference run. The red circle will appear as shown below indicating that a positional lock has been found for the overlapping features.



Note: as you hold down the mouse button, window B zooms in to help with the alignment.

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

Progenesis LC-MS User Guide



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. Additing an additional vector will improve the alignment further. **Note** this time as you click to add the vector it 'jumps' automatically to the correct position using the information from the existing alignment vector.



Repeat this process moving the focus from top to bottom on the Whole Run view



Note: the number of vectors you add is recorded in the **Runs** table also with each vector addition the Score and alignment quality updates. This can help guide the number of manual vectors you need to add before applying the automatic alignment.

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

The number of manual vectors that you add at this stage is dependent on the misalignment between the current run and the Reference run. In many cases only using the Automatic vector wizard will achieve the alignment.

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



11. Then select Automatic Alignment to bring up the Automatic Alignment dialog and click **OK**. The automatic alignment process will begin, using the manual vectors you have added to aid in automatic vector placement.

Autom	atic A	lignment	×
Selec	t the n	uns for automatic alignment vector generation	Vectors
	A1	run has user vectors	5
V	A2	run has user vectors	5
V	A3	run has user vectors	5
	C1	this run does not need to be aligned as it is the alignment reference	Ref
v	C2	run has user vectors	5
V	C3	run has user vectors	5
		ОК	Cancel

Note: the tick boxes next to the 'Run name control' which control whether vectors will be generated for each run.

To review the vectors, automatic and manual return to page 11

Appendix 5: Within-subject Design

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

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

C-MS_Tutorial - Progenesis LC-MS File Review Experi wort Data Alignment Filtering Design	nent Ret Setup I	view Peak Picking	Peptide Statistics	Identify Peptides	Refine Identifications	Resolve Conflicts	Review Proteins	Protein Statistics	Report	no	nline
New hich experiment design type do y Between-subject Design Do samples from a given subject appear in only one condition? Then use the between-subject design. To set up this design, you simply group the runs according to the condition (factor level) of the samples. The ANOVA calculation assumes that the conditions are independent and therefore gives a statistical test of	ou want A C	t to use t	for this ex A1 A2 nesis LC-MS e a new	cperiment Delete Remove Remove experim	2 Have you subject u Then use Note: you every sub	ithin-sub taken samp nder differd the within- u must have ject for eve	ject Desig oles from a gi ent condition: subject desig a sample from ry condition to	n s? n. Patient X	Before X1 Y1	During X2	After X3
therefore gives a statistical test of whether the means of the conditions are all equal.	Add condi	Creat Name: Start	e a new Before I t with an e y layout fr	experim During and mpty layo	After Treatm ut Crea	te design	Cancel	Patient Y Patient Z Jut he ted A a	Z1	72 Z2	Z3
					repeated difference reduced a condition create a r The withi thought o paired-sa comparia repeated	measures Al es can be eli is a source o differences more powerf n-subject de f as an exte mples t-test on between measures.	NOVA individu iminated or of between (which helps ful test). esign can be nsion of the to include more than two	al to o			

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

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



You can create additional Experimental Designs using the New tab

All of these Experimental Designs are available at all the following stages in the LC-MS workflow.

Appendix 6: Power Analysis (Peptide Stats)

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

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

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 features 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 features with a power >0.8'

Using the Significant p<0.05 features (7965), as an example, view the power analysis.



This is displayed graphically showing that 78.6% of the 7965 features have a power of 80% or that 4 replicates would give you 96.4% of your data with power > 0.8.

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

Appendix 7 Search engine parameters (Stage 9) Mascot

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

MASCOT MS/MS Ions Search			
Your name	andy.borthwick	Email	andy.borthwick@nonlinear.com
Search title	Search_full_LC-MS_4.1		
Database(s)	ApoE MSDB NCBInr	Enzyme	Trypsin 🔻
		Allow up to	1 🔻 missed cleavages
	SwissProt	Quantitation	None 🔻
Taxonomy	· · · · Firmicutes (gram-positive bacteria) ▼		
Fixed modifications	Carbamidomethyl (C)	>	mTRAQ:13C(6)15N(2) (N-term) mTRAQ:13C(6)15N(2) (Y) NIPCAM (C) Oxidation (HW) Propionamide (C)
	Display all modifications 📃		Pyridylethyl (C) Pyro-carbamidomethyl (N-term C)
Variable modifications	Oxidation (M) Phospho (ST) Phospho (Y)	>	Sulfo (STV) TMT2plex (K) TMT2plex (N-term) TMT6plex (K)
Peptide tol. ±	9 ppm ▼ # ¹³ C 0 ▼	MS/MS tol. ±	0.6 Da 🔻
Peptide charge	2+ 🔹	Monoisotopic	Average
Data file	D:\Customer Data\LC-MS\LC-ME Browse		
Data format	Mascot generic 🛛 🔻	Precursor	m/z
Instrument	ESI-TRAP	Error tolerant	
Decoy		Report top	AUTO 🔻 hits
	Start Search		Reset Form

Database : NCBInr (circa 02/13) was used with the Taxonomy restriction set to Fermicutes

Fixed modifications: Carbamylation(C) and variable modifications: Oxidation (M), Phospho (ST) and Phospho (Y)

Peptide Tol: 9ppm

Instrument: ESI-Trap