

Progenesis QI User Guide

Analysis workflow guidelines

Waters

THE SCIENCE OF WHAT'S POSSIBLE.™

Contents

Introduction	3
How to use this document	3
Data used in this user guide	3
Workflow approach to data analysis	4
Restoring or starting a new Metabolomics Tutorial	5
Stage 1: Data import and QC review of data set	7
Stage 2A: Automatic Processing of your data	8
Stage 2B: After Automatic Processing	12
Stage 3: Licensing	14
Stage 4: Review Alignment	15
Stage 5: Experiment Design Setup for Analysed Runs	22
Stage 6A: Peak Picking	25
Stage 6B: Reviewing Normalisation	34
Stage 7: Review Deconvolution	
Stage 8: Identify Compounds	44
Identification of compounds using the search method Progenesis MetaScope	45
Incorporating Theoretical Fragmentation	47
Creating and using databases	50
Additional Identification Methods	55
Stage 9: Review Compounds	66
Compounds separated by Drift Time	72
Exporting Compound Data to Pathways Tool(s)	74
Stage 10: Compound Statistics	77
Appendix 1a: Stage 1 Data Import and QC review of data set	85
Appendix 1b: Stage 1 Data Import from UNIFI (Waters Only)	90
Appendix 2: Stage 2 Processing warnings and failures	93
Appendix 3: Licensing runs (Stage 3)	96
Appendix 4: Manual assistance of Alignment	97
Appendix 5: Within-subject Design	102
Appendix 6: Using LipidBlast for Compound Identification	104
Appendix 7: Exporting compound identities to IPA	
Appendix 8: Power Analysis	
Appendix 9: Using Clip Gallery to Save and Export Pictures and Data	110

Introduction

This user guide takes you through a complete analysis of 22 LC-MS runs of metabolite containing samples for 3 biological groups (6 replicate runs per group) and a set of 4 QC runs using the unique Progenesis QI workflow. It starts with LC-MS data file loading then Alignment, followed by Analysis that creates a list of interesting compounds (molecules) which are identified with compound search and then explored within Compound Stats using multivariate statistical methods.

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 Progenesis QI workflow, therefore if you are using your own data files please refer to Appendix 1 (page 85) then start at page 8.

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 90 minutes and is divided into two sections. This means you can perform the first half focused on run alignment and analysis then complete the second half of analysis exploring comparative differences and compound identity at a convenient time. The table of contents allows you to focus on particular areas of the workflow.

Data used in this user guide

We would like to thank Dr Giorgis Isaac at Waters Corporation, Milford USA for providing the example data which has been adapted for this user guide as well as invaluable discussion on the handling of the data.

Workflow approach to data analysis

Progenesis QI 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 Compound Stats.



Stage



Description	Page
Import Data: Selection and review of data files for analysis	7
Automatic Processing: setting up steps for automatic processing	8
After Automatic Processing: how to work with auto analysed data	12
Licensing : allows licensing of individual data files when there is no dongle attached (Appendix 3)	14
Review Alignment: automatic and manual run alignment	15
Experiment Design Setup : defining one or more group set ups for analysed aligned runs	22
Peak picking : setting parameters for and performing peak picking of compound ions	25
Reviewing Normalisation: examine data normalisation methods	34
Review Deconvolution : review and edit the various adduct, forms of a compound	36
Identify Compounds: search identity of compounds using Progenesis MetaScope and or other search engines	44
Review Compounds : managing possible compound identities exploring identity and expression between conditions	66
Compound Statistics : performing multivariate statistical analysis on tagged and selected groups of compounds	77

Restoring or starting a new Metabolomics Tutorial

If working with your own data files then refer to Appendix 1 page 85.

To use the Tutorial data:

Open Progenesis QI and download the DataSet for the LC-MS user guide (.zip) from the **'User guide and tutorial data**' 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 downloaded (.zip) file and extract it to a folder, i.e. 'C:\Users\john_smith\Documents\Progenesis QI Tutorial

0 Progenesis QI	
File Experiments	nonlinear A Waters Company
Recent experiments	Search
0) Open Experiment Compared with the second	Search Progenesis Tutorial O
Organize ▼ New folder ③ BitTorrent Sync	Ezinfo version information Date modified Ezinfo is installed.
Desktop Decuments Music Videos Andy Borthwick Computer Quench Quench Sconputer Quench Quench Computer Quench	DMSeProgenesi 25/11/20151422 Quickly go to an ion map location Want to quickly validate your sample running by zooming to a known ion?
File name: Progenesis QI Tutorial HDMSe.Pro 👻	Open Cancel
Other experiments	Latest blog posts

Now you can restore the uncompressed Progenesis QI tutorial archive file. To do this, first locate the Tutorial Archive file using the **Open** button.

This opens the 'Import from archive' dialog.

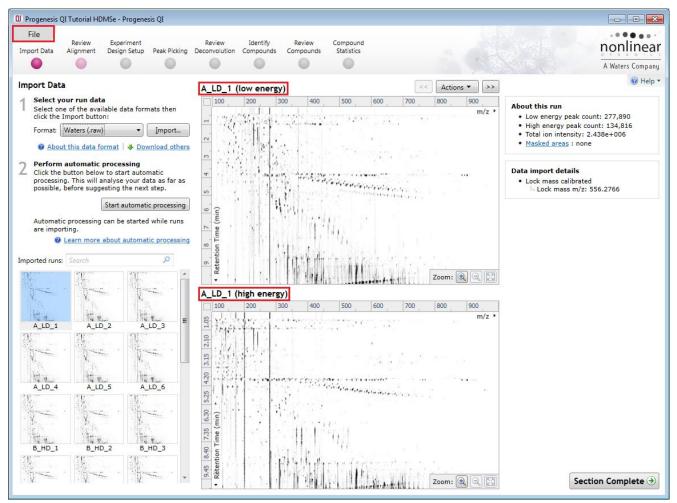
Select the **Create a new experiment** option and select the folder in which you placed the archive, using **Browse**.

QI Import Experiment from Archive				
	from archive kperiment from this archive, any changes to the experiment will be below, not back to the archive.			
 Replace an existin Experiment to replace Create a new experiment 	ace:			
Experiment name:	Progenesis QI Tutorial HDMSe			
Save to folder:	NProgenesis QI Tutorial_HDMSe\Progenesis Tutorial Browse			
	Import			

Then press Import.

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

The archive will restore with the software opening at the first stage in the workflow, Import Data. The panel on the left will display a 'thumbnail' for each run in the archive, indicating that they have been imported successfully.



The data file format for this tutorial was **.Raw**. It was acquired from a **SYNAPT-G2S** with the ionisation polarity set to **positive**.

You can look at the current **Experiment Properties** using the file menu. This displays details of the adduct definitions that were used during data import.

Note: total adduct abundances will be reported once Peak Picking has been performed.

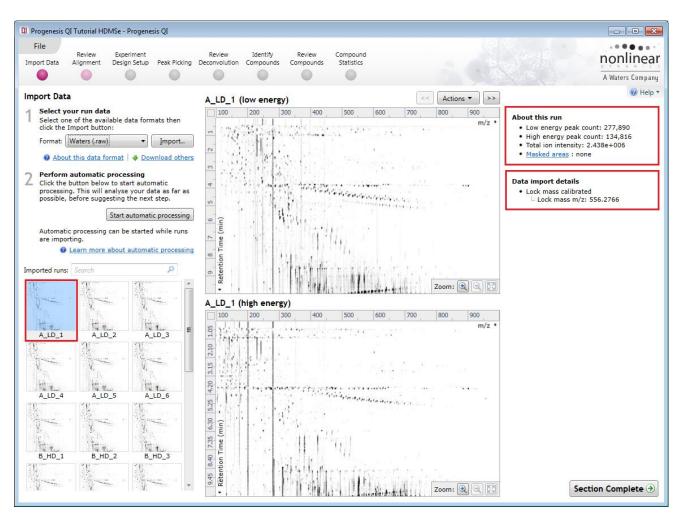
As this example is using HDMSe data it contains both low and high energy exact mass data.

The HD (High Definition) refers to the presence of Ion Mobility data (Drift Time)

Ionisation polarity: Positive Runs in this experiment: 22 Feature detection: High resolution Peak processing: Profile data Adduct definitions used in the experiment:				
Name	Adduct Mass	Charge	Abundance	
M+2H	2.014552864	2		
M+H+Na	23.99649711	2		
M+H-2H₂O	-35.01385294	1		
M+H-H ₂ O	-17.0032882!	1		Ξ
M+H	1.007276432	1		
M+NH₄	18.03382553	1		
M+Na	22.98922068	1		
M+CH₃OH+H	33.03349118	1		-

Stage 1: Data import and QC review of data set

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



Each run appears as a 2D ion intensity map in the list. The current Run is displayed in the main window showing both low and high energy data. At this stage you will be warned if there are any data import errors for any of the files. The files will be highlighted in red and the error will appear to the right of the screen.

Note: if you have imported one or more runs that are either: not required for the experiment or are displaying data import errors (such as incorrect polarity) these runs can be removed by right clicking on the run in the list and selecting **Remove run**.

Details about the current run are displayed, top right, showing the Low and High energy peak counts and Lock mass calibration status.

Tip: the **'Mask areas for peak picking'** facility (under **Actions**) allows you to exclude areas (usually early and/or late in the LC dimension (Retention Time) that appear excessively noisy due to capture of data during column regeneration. This is **NOT** used for this data set.

Now click Start automatic processing.

Stage 2A: Automatic Processing of your data

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

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



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

- Select an alignment reference
- Automatic alignment of all runs to a reference run
- Define an Experiment design
- Automatic peak picking for compound ion detection

In this tutorial example you have 22 HDMSe LC-MS runs.

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

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

Progenesis QI provides three methods for choosing the alignment reference run, as described below:

1. Assess all runs in the experiment for suitability

This method compares every run in your experiment to every other run for similarity, then selects the run with the greatest similarity to all other runs as the alignment reference.

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

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

This method asks you to make a selection of reference candidates; the automatic algorithm then chooses the best reference from this subset of runs.

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

3. Use this run

This method allows you to manually choose the reference run.

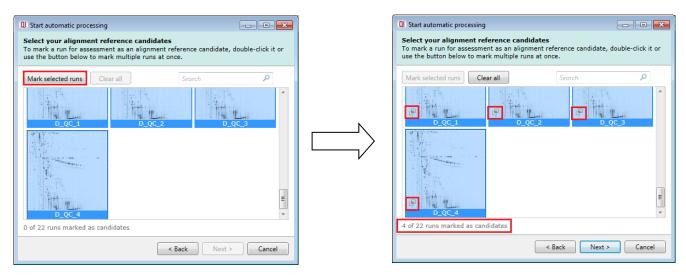
Manual selection gives you full control, but there are a couple of risks to note:

• If you choose a pending run which subsequently fails to load, alignment will not be performed.

• If you choose a run before it fully loads, and it turns out to have chromatography issues, alignment will be negatively affected (for this reason we recommend that you let your reference run fully load and assess its chromatography before loading further runs).

For this tutorial we will select the second option and click Next.

Now select all four of the QC runs as the Alignment Reference Candidates and click **Mark selected runs**. These will be marked with a Candidate icon. Now click **Next**.



You will now be asked if you want to align your runs automatically. Check box is ticked and select Next.

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? Ves, automatically align my runs
< Back Next > Cancel

The next page of the processing wizard will ask you to define an **Experiment Design.** To enter a design at this stage details of Experiment conditions must be available either as a MassLynx .SPL file or as a .CSV file.

Note: if these are not available leave the Set up option unticked and click next. You can enter a design(s) following automatic processing.

For this tutorial the .spl file **Progenesis QI_HDMSe Tutorial.spl** is available in the folder you restored the Tutorial Data Set.zip in.

0 Start automatic processing
Experiment design Experiment designs allow you to group and compare your samples according to their experimental conditions.
By defining an experiment design in advance, statistical measures such as ANOVA can be calculated automatically. Set up an experiment design Enter a name for the experiment design: All Conditions Load the criteria for grouping runs from this file: Browse Group runs by: noveline Browse
What file formats are supported? Kext > Cancel

Select the appropriate criteria by which to group by; in this example select 'Condition' and then click **Next** to move forward to the next page of the processing wizard.

This asks you if you want to **Perform peak picking** and allow you to **Set parameters**.

0 Start automatic processing
Peak picking Peak picking is the process by which we locate the compound ions and their isotopic peaks in your samples.
After the automatic alignment is finished, do you want to start automatic peak picking? Perform peak picking Set parameters
< Back Finish Cancel

For the purposes of this User guide we will use the default settings for peak picking.

Note: for more details on setting Peak Picking parameters refer to the section on Peak Picking (page 25).

Now click **Finish** to start the automatic processing.

The Alignment process starts with the automatic selection of **D_QC_4** as the reference

Q Automatic Processing	(23%)				
Automatic process Current step: Choosing a	-				
 Importing runs: Selecting reference: Aligning runs: Creating design: 	22 of 22 processed Choosing an alignment reference Pending Pending		QI Automatic Processin Automatic process Current step: Aligning '	sing	
Peak picking:	Pending		 Importing runs: Selecting reference: Aligning runs: Creating design: 	22 of 22 processed D_QC_4 1 of 21 processed Pending	
		Cancel	Peak picking:	Pending	
					Cancel

Once Alignment completes the Design is applied and Peak Picking commences.

Q Automatic Processing	(47%)	—	
Automatic process Current step: Aligning 'B	-		
 Importing runs: Selecting reference: Aligning runs: 	22 of 22 processed D_QC_4 7 of 21 processed		QI Automatic Processing Automatic processing Current step: Analysing
Creating design: Peak picking:	Pending Pending	Cancel	✓ Importing runs: 22 of 22 processed ✓ Selecting reference: D_QC_4 ✓ Aligning runs: 21 of 21 processed ✓ Creating design: Created
		Cancel	Cancel

Finally the Automatic Processing completes reporting the number of compound lons found.

Q Processing Complete		×
Automatic process	• •	
 Importing runs: Selecting reference: 	22 of 22 processed D OC 4	
 Aligning runs: 	21 of 21 processed	
 Creating design: Peak picking: 	Created 6411 compound ions found	
	i	Close

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: if Progenesis QI, during import of data and/or performing the various stages of Automatic Processing, detects problems with the data then the Processing dialog will report on the arising issue(s).

For example: failing to import one or more files.

Details on	the	various	warning	s and	managir	ig failure
messages	are	availab	le in App	pendix	2 (page	93)

i ng complete (with warnin 1 seconds	gs).
22 of 22 processed 1 failed to import	
D_QC_3	
20 of 20 processed	
Created	
6069 compound ions found	
	1 seconds 22 of 22 processed A 1 failed to import D_QC_3 20 of 20 processed Created

When the Automatic Processing completes click **Close**, the software remains at the Import Data stage of the workflow where you can review the chromatography of the imported runs and remove/replace any runs that have failed to import.

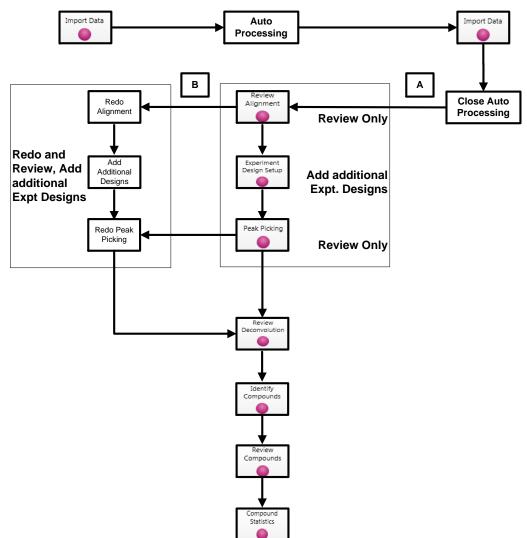
Waters
THE SCIENCE OF WHAT'S POSSIBLE."

Stage 2B: After Automatic Processing

When Processing completes, access to the later steps in the analysis workflow will depend on the steps performed in the automatic processing. Here we outline the various paths and options you can take to proceed with the analysis following Automatic Processing.



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



In this example where the data has been automatically processed once you **Close** the completed automatic processing dialog you can either:

- **Option A**: Review the Chromatography at the Import Data stage (page 13) and then continue with Reviewing Alignment quality (page 15).
- **Option B**: During the review process you decide to redo the alignment (page 15) and/or Peak Picking (page 25).

Note: the tutorial data described here does not require redoing either alignment or peak picking so you can step straight to Review Deconvolution. However, this guide is designed to support the analysis of your own data sets as well, so it covers all stages of the workflow.

Tip: it is good practice to close your experiment following automatic analysis and create an archive (see FAQs for details on creating an archive), and then reopen it and continue with analysis.

Review Chromatography

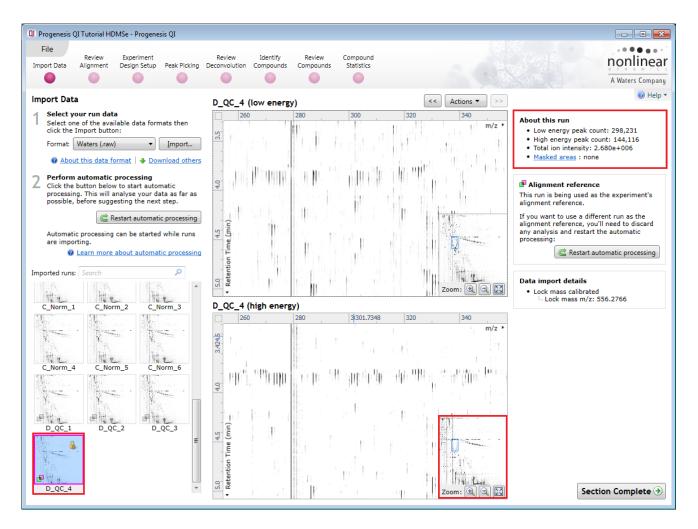
Each data file appears as a 2D representation of the run. In the case of MSe and HDMSe data both high and low energy maps are displayed.

At this stage you will be warned if any of the data files have failed to import. The reasons for import failure(s) include: run is centroided when importer is expecting a profile (continuum) run, the run has a different polarity to the other runs etc.

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

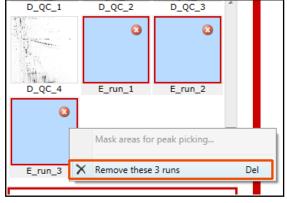
Note: for this data set do not remove any files

You can examine the quality of the imported runs using the 2D representation of the runs making use of the Zoom tools (bottom right of the ion maps) and Panning tools.



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.



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 96)



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

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

File nport Data	Licensing	Review Alignment	Experiment Design Setup	Peak Picking	Review Deconvoluti	Identify on Compounds	Review Compounds	Compound Statistics			nonli	nea
								•			A Waters C	ompan
	cense Runs											
This installa icensed rui	ation is curren ns only.	tly restricted	to analyse	Run name						Licence state	License this run	^
Lo liconeo :	your runs, you	u need an ev	valuation or	C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	\checkmark	
	ce code whic			C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	V	
a sales rep	resentative.			C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	\checkmark	
Once licens	ed, your runs	s can be ana	lysed on	C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	V	
any installa	tion of the sof	ftware. The li	cence is	C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	V	
	lly included w	hen archivin/	g an	C:\Users\andy	borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	V	
experiment				C:\Users\ar	📙 Unable to	save experimen	t		×	QI T Unlicensed	V	
	have been lic lick here to m			C:\Users\ar						QI T Unlicensed	V	
available on	this compute	er.		C:\Users\ar			analysis withou	ut a valid license	If you close now your	QI T Unlicensed	V	Ξ
f vou have	one, you can	open a licenc	e file to	C:\Users\ar	anal	ysis will be lost.				QI T Unlicensed	V	
nstall.				C:\Users\ar						QI T Unlicensed	v	
f vou have	just installed	a donale, clic	k here	C:\Users\ar						QI T Unlicensed	V	
.,	,			C:\Users\ar				0	K Cancel	QI T Unlicensed	V	
				C:\Users\andy	borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	v	
				C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	V	
				C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	V	
				C:\Users\andy	borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	v	
				C:\Users\andy	borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	V	-
				C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	\checkmark	
				C:\Users\andy	.borthwick\D	ocuments\Custom	er Data\Progene	sis QI v2.2_Demo	Suites_Tutorial\Progenesi	s QI T Unlicensed	\checkmark	-
							Run licence	code:	·	Use L	icence Cod	e
											tion Comple	

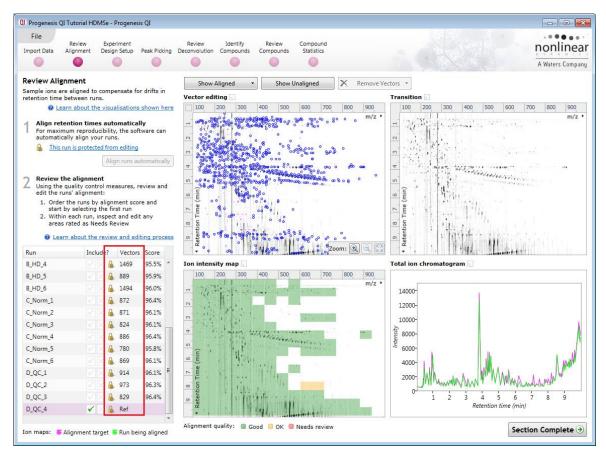
Note: although the analysis will be lost (Alignment Peak Picking etc) the experiment file and all the 'loaded pre-processed runs are retained'. This allows you to reopen the experiment, when you have a licenced code and/or dongle available, and redo the analysis steps without having to reload the data.

For more details on managing the licencing of your data refer to Appendix 3 (page 96)

Stage 4: Review Alignment

At this stage, **Review Alignment** opens displaying the alignment of the runs to the Reference run (D_QC_4).

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



Note: for this tutorial data set, refinement of the alignment is **NOT** required, so you can move to page 19 without altering the alignment.

However, when reviewing the quality of alignment after 'Automatic Analysis' of a data set you may decide that the alignment requires editing. Reviewing and editing are described in the following sections.

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

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

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

Details on editing alignment are described in Appendix 4 (page 97).

Layout of Alignment

To familiarize you with Progenesis QI Alignment, this section describes how to use the various graphical features used in the alignment of the LC-MS runs.



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

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

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

Total Ion 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 **Ion Intensity map** (Window C).

To best understand the process of alignment and the meaning of the scoring we will now undo the automatic alignment for C_Norm_4, then simulate a poorly aligned run by adding an incorrect manual vector.

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

- If you have not already unlocked the Analysis (note padlock icons in the runs table see right) then refer to page 15.
- In the Run table click on C_Norm_4 to make it current. You will now be looking at the alignment of C_Norm_4 to D_QC_4 in the Unaligned view.

Run	Include?	Vectors	Score	
B_HD_4	 X 4 	1469	95.5%	*
B_HD_5	🗹 🖂 🔒	889	95.9%	
B_HD_6	< X 🌡	1494	96.0%	
C_Norm_1	🗹 🗙 🔒	872	96.4%	
C_Norm_2	< X 🌡	871	96.1%	
C_Norm_3	🗹 🗙 🌲	824	96.1%	
C_Norm_4	🗸 🗙 🌡	886	96.4%	
C_Norm_5	X X 🌡	780	95.8%	
C_Norm_6	< X 🌡	869	96.1%	
D_QC_1	🗹 🗙 🌲	914	96.1%	Ξ
D_QC_2	< X 🌡	973	96.3%	
D_QC_3	🗹 🗙 🌲	829	96.4%	
D_QC_4	🖌 🖂 🌡	Ref		
				Ŧ

First click on **Remove Vectors** and then select remove All vectors in the whole run.

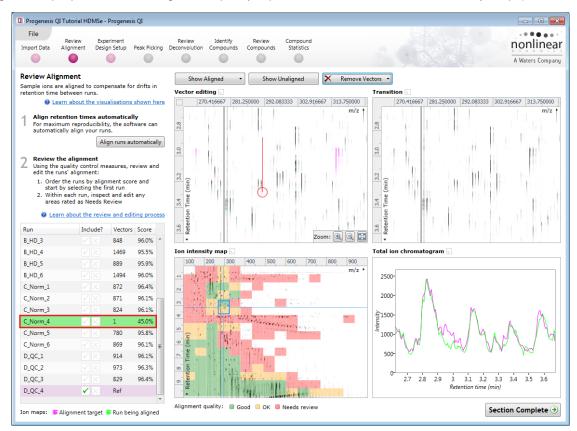


Note: the Retention Time alignment of this data is good between the runs as indicated by a high percentage score (in the absence of any vectors) and the Alignment quality showing as 'all' green.

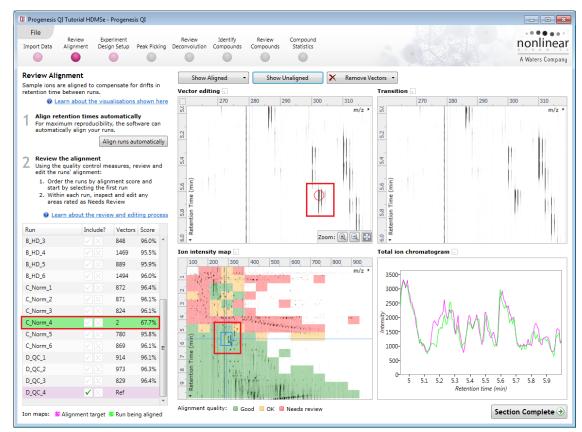
To simulate poor alignment, place a single manual vector on the Vector editing view (Window A). To do this, in the Vector Editing panel, click and drag out a single vector then release the mouse button. By doing this a single manual vector will appear with a length corresponding to the 'drag'.

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

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



Now place a second vector at a different location **BUT** this time hold down the **Alt** key time click and release **without** deliberately misaligning the vector.

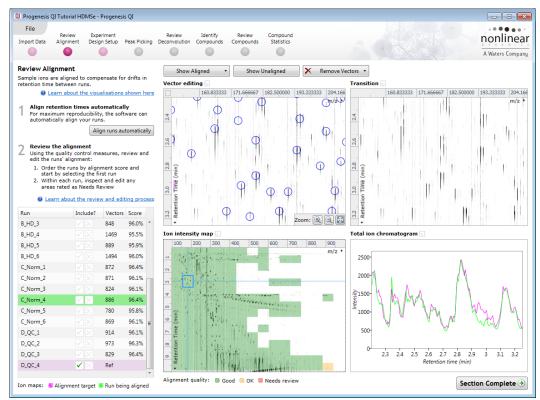


Notice how the correctly placed vector markedly improves the Score this is also reflected in the increased proportion of the Ion Intensity Map showing green indicating a good quality alignment.

Reviewing quality of alignment vectors

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

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



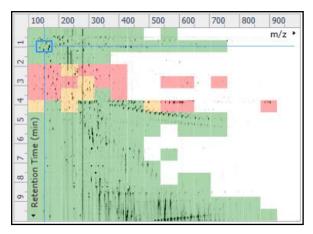
Using the **Simulated** miss-aligned example from the previous section to explain the review process for alignment, the alignment looks as below with a region of poor alignment (highlighted in red).

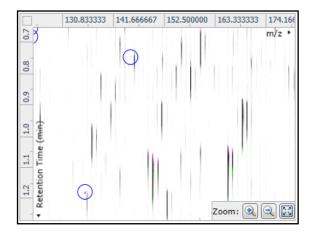


Reviewing Quality of Alignment

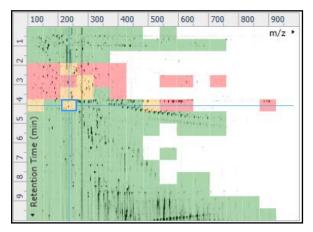
At this point the quality metric, overlaid on the Ion Intensity Map as coloured squares, acts as a guide drawing your attention to areas of the alignment. These range from Good (Green) through OK (Yellow) to Needs review (Red). Drag out a 'Focus' area that corresponds to one of the coloured squares.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

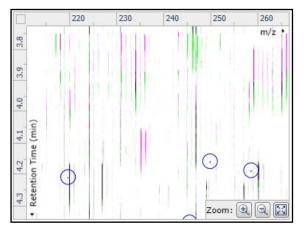
For a 'green' square the majority of the data appears overlapped (black) indicating good alignment. When viewed in the Transition view the data appears to pulse.



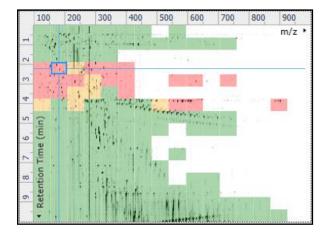


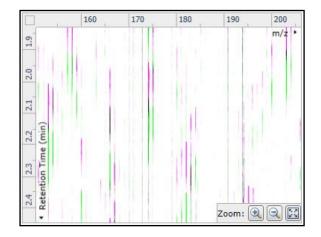
For a 'yellow' square some of the data appears overlapped (black) indicating OK alignment. When viewed in the Transition view some of the data appears to pulse.





For a 'red' square, little if any of the data appears overlapped (black) indicating questionable alignment.

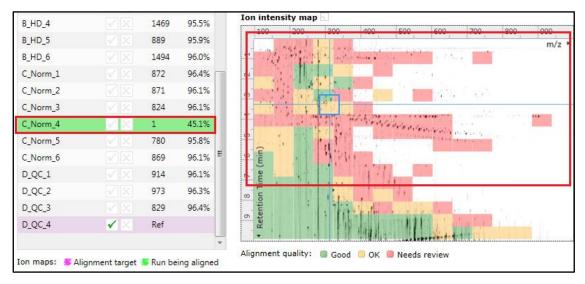




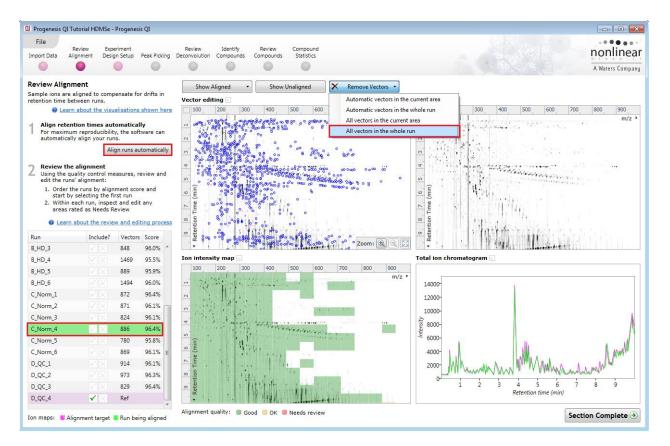
When viewed in the Transition view little data appears to pulse.

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

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



Note: a marked red area combined with a low score clearly indicates a 'miss alignment' and may require some manual intervention (see Appendix 4, page 97).



The alignment quality of this data set does not require any manual intervention so before going to the next section make sure you have Removed all manual vectors and re-performed the Automatic alignment. To do this for C_Norm_4 first select Remove 'All vectors in the whole run' and then click Align runs automatically.

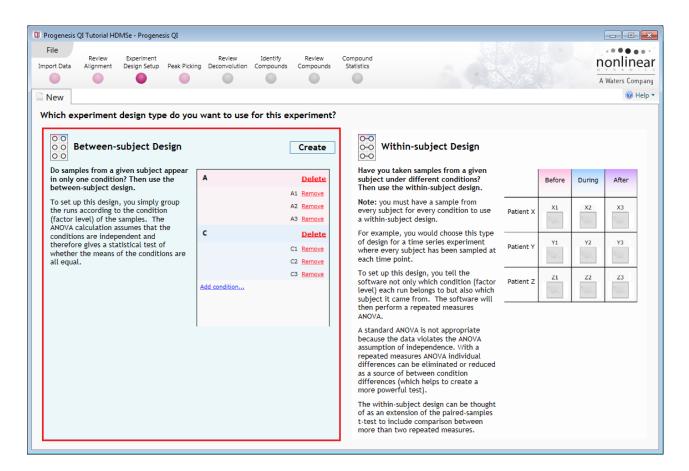
Having aligned the runs automatically click **Section Complete** to move to Experiment Design Setup.

Stage 5: Experiment Design Setup for Analysed Runs

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

Note: If you have created a design as part of the automatic processing then it will appear as a Design tab. However if you have reperformed the analysis 'manually' then the only tab displayed will be **New**.

There are two basic types of experimental designs:



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

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

Additional information on how to apply the Within-subject Design is in Appendix 5 page 102

This experiment contains 4 conditions: A, B, C and D 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.

Q Create New Experiment Design	
Enter a name for the experiment	design:
All conditions	(\
How do you want to group the ru	ins?
Group the runs manually	
Copy an existing design:	Ŧ
Import criteria from a file:	Browse
Group runs by: <	no valid groups>
What file formats are supported	Create design Cancel

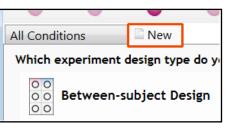
Give the new experimental design a name and then click Create design.

Al Progenesis QI Tutorial H File Import Data Alignment	Experiment	Review Identify Review Deconvolution Compounds Compounds	Compound Statistics	- deal?	nonlinear
All conditions I ×	New				A Waters Company
Setup conditions	New	Runs Add Selected Runs to Condition	▼ Search	٩	
Setup the conditions that	you want to compare below .), and then assign each of ect condition.	Add Selected Runs to Condition	B_HD_2	B_HD_3	A
C_Norm	Delete	C_Norm			
	C_Norm_1 Remove	A_LD			
	C_Norm_2 Remove				
	C_Norm_3 Remove	B_HD_4	B_HD_5	B_HD_6	
	C_Norm_4 Remove	1988 C	3360 -	2360	
	C_Norm_5 <u>Remove</u> C_Norm_6 <u>Remove</u>				
A_LD	Delete				E
	A_LD_1 Remove	D_QC_1	D_QC_2	D_QC_3	
	A_LD_2 Remove		and the second	1942 -	
	A_LD_3 Remove				
	A_LD_4 Remove				
	A_LD_5 Remove	D_QC_4			
Add condition	A_LD_6 <u>Remove</u>				-
				Se	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. B_HD) 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 **Progenesis QI_HDMSe Tutorial.SPL** file available in the Experiment Archive you restored at the beginning of this tutorial exercise.

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

Q Create New Experiment Desig	n 🗾	
Enter a name for the experime	ent design:	
Minus QC	(-	
How do you want to group the	runs?	
\bigcirc Group the runs manually		
Copy an existing design:	All conditions 👻)
Import criteria from a file:	genesis QI_HDMSe Tutorial.spl Browse)
Group runs by:	Condition	
	VERSION	
What file formats are support	Condition	
what he formats are suppor	Index	J

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 i.e in this example the **QC**.

File File	QI Tutorial HD Review Alignment	MSe - Progenes Experiment Design Setup	sis QI Peak Picking	Revie Deconvol		Review Compounds	Compound Statistics	nonlii A Waters C	neal
All condition	ons	Minus QC	Ι×Ι	New) Help
	ditions that drug A, etc)	you want to co , and then assi ct condition.		Runs	Add Selected R	uns to Conditio	n 🔻 Search	٩	
		A_LD_6	Remove	1					
B_HD			Delete						
		B_HD_1	Remove						
		B_HD_2	Remove						
		B_HD_3	Remove						
		B_HD_4	Remove						
		B_HD_5	Remove						
		B_HD_6	<u>Remove</u>						
e QC			Delete						
		D_QC_1	Remove						
			<u>Remove</u>						
			<u>Remove</u>						
		D_QC_4	Remove						
Add condition	<u>n</u>		*					Section Comple	ete 🤿

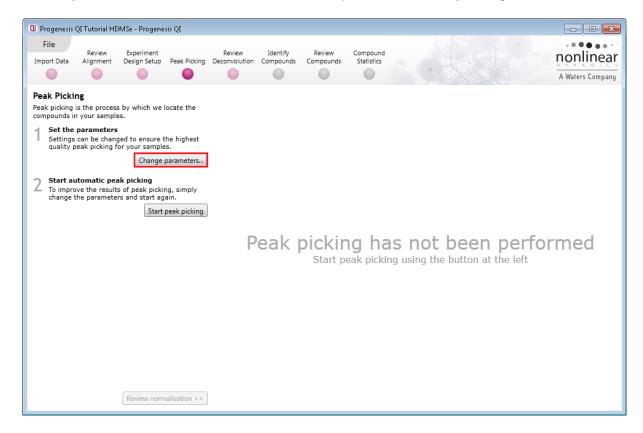
Note: On deleting a condition the runs will reappear in the Runs window.

Note: both designs are available as separate tabs.

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

Stage 6A: Peak Picking

Having set up one or more Experiment Designs for your data the Peak Picking stage will open as shown below. Here you can define thresholds for the detection of peaks across all of your aligned runs.



Note: if you have processed your data automatically and have NOT discarded your analysis as described on page 15 then the Peak picking will appear as shown on page 27.

If you have been reviewing the Alignment as described in the previous section then the Peak picking will look as shown above, indicating that you will need to **Start peak picking**

Peak Picking Parameters

The Peak Picking Parameters dialog opens by clicking on **Change parameters...** 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 compound ion outlines.

Retention time limits	Hi	gh energy limits	Adducts
Runs to include	Peak picking limits		
Choose runs for peak picking			
You can tick or un-tick each run to		Run	
control which will be used by the peak picking algorithm. Although any		A LD 1	
In which is left un-ticked will not ffect the compound ion outlines, it ill still have outlines added to it and ill be available in the experiment		A_LD_2	
		A LD_3	
will be available in the experiment design setup.		A LD 4	
Learn more about why you might not	V	A LD 5	
want to select all runs.		A LD 6	
		B HD 1	
		B HD 2	
		B_HD_3	
		B_HD_4	
		B_HD_5	
		B_HD_6	
		C_Norm_1	
		C_Norm_2	

Retention time limits	High energy lim	its	Adducts
Runs to include		Peak picking	limits
Sensitivity			
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.	Automatic Absolute ion ii % Base Peak The automatic noise estimati the noise level the sensitivity ions will be de fewer	sensitivity n on algorithm s in the data value, the m	to determine
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.	Apply a minim		th] minutes

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: feature outlines will be added to 'un-ticked' runs; however, 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).

Finally you can set **Retention time limits** for the detection. The default limits are displayed if no limits are defined.

Peak Picking Parameters		_				
Runs to include	Peak pick	Peak picking limits				
Retention time limits	High energy limits	Adducts				
Retention time limits You can set the minimum and maximum retention time for peak picking. Ions that elute before or after these values will be ignored.	Ignore ions before	0.0254666() minutes				
		OK Cancel				

Click **OK** to close the Peak Picking Parameters dialog.

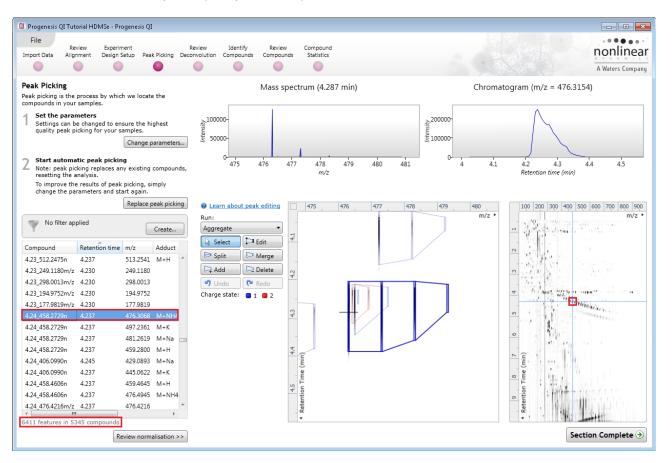
Press Start peak picking to start the detection process.

Analysing			

During peak picking, a progress bar will appear telling you that it is Analysing.

Following automatic peak picking the detected peaks are displayed in a colour according to their charge state. (**note**: number of charge states represented reflects Adduct Definitions used in the experiment)

The actual number of peaks (Compound Ions) detected is recorded on the bottom left of the screen.



To inspect the detection more closely drag out an area of interest (m/z by RT) on the full ion map.



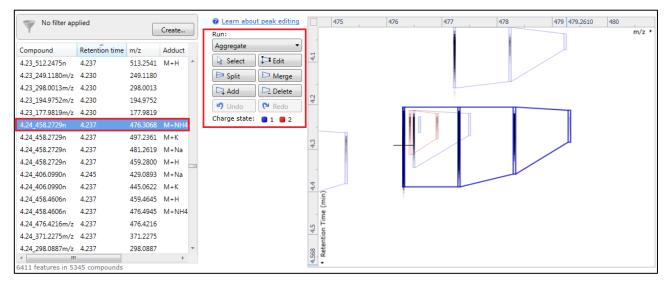
In this example the range viewed in the zoomed widow shows the 4 compound ion adducted forms (M+NH₄, M+K, M+Na and M+H) that contribute to the deconvolved neutral mass for the Compound 4.24_458.2729n in the table. Hover the cursor over each ion to show m/z and Retention Time.

Editing Peak Picking

The editing tools that are present at Peak Picking will allow you to Edit, Split, Merge, Add and Delete (**Note**: Undo and Redo are also available).

These tools can be applied to a selected Compound Ion or in the case of the Add tool it can be applied to an undetected Isotopic pattern.

For the purposes of understanding how the tools work we will focus on one isotopic pattern, the $M+NH_4$ for the Compound 4.24_458.2729n. Order on **Compound** in the table and scroll to m/z and RT..

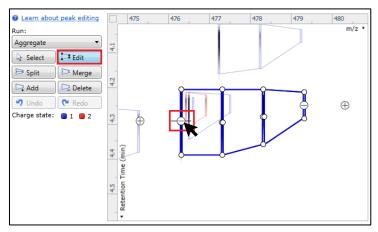


Select the compound ion to edit either by clicking on it in the table or inside the isotopic pattern on the ion map, in each case the selected ion will become centered in the view.

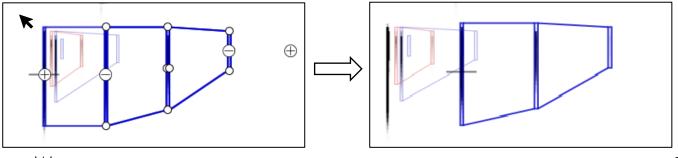
Edit tool

To remove a peak from an existing compound ion select the **Edit** tool and click on the compound ion to reveal the 'edit handles'. You can zoom in more by dragging an area around the peptide ion of interest.

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

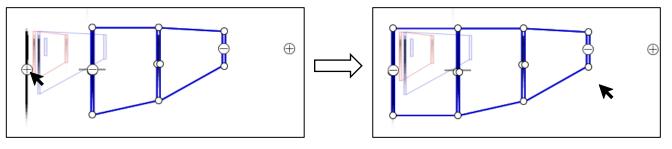


2. Click outside the boundary of the compound ion to update the view.



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

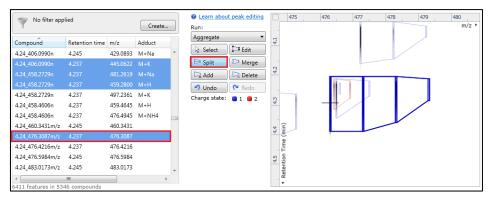
1. Click on the 'plus' handle on the peak to the left (under the plus handle) to add it



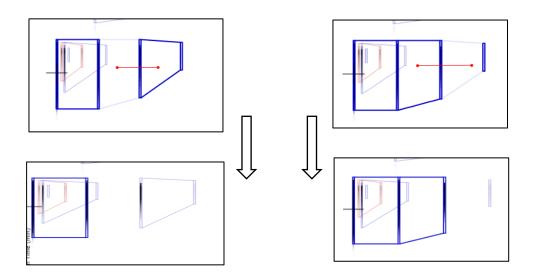
- 2. Then click outside the peptide ion to update the view.
- 3. Note: If you are not satisfied with the editing use the **Undo** button and retry.
- 4. Note: The Undo and Redo buttons will no longer be active if you leave and return to Peak Picking.

Split and Merge Tools

Using the edited compound ion, $M+NH_4$ adduct, for the Compound 4.24_458.2729n which, due to editing is now displayed as a single compound ion at 4.24_476.3087m/z.

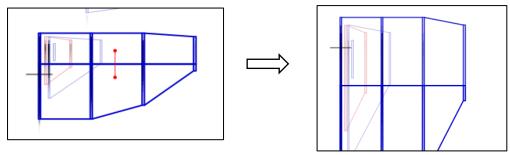


- 1. Click on the Split tool
- To split the vertically into 2 Charge state 1 compound ions, click and release in the centre of the ion pattern. A red dot appears, drag the cursor to right, a line extends across the isotope envelopes. Extending this line will alter how the pattern splits.



- 3. To complete the split click again, the pattern will split depending on the length of the line
- 4. Note: Using the Undo tool will allow you to retry the splitting of the pattern

- 5. Note: The Undo and Redo buttons will no longer be active if you leave and return to Peak Picking
- 6. To split the horizontally into 2 Charge state 1 compound ions, **click and release** in the centre of the ion pattern. A red dot appears, drag the mouse up or down, a blue line extends across the isotope envelopes showing where the split will occur. Extending the red line (up and down) will alter where the pattern splits.

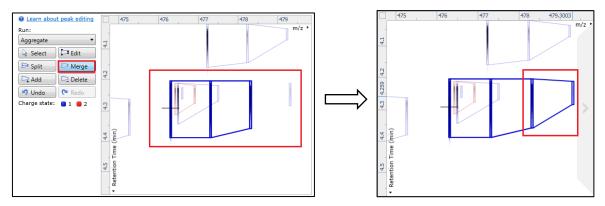


7. To complete the split click again, the pattern will split depending on the length of the red line.

Merge Tool

The merge tool can be used to combine 2 Isotopic patterns that were originally detected as separate compound ions as an example of this functionality try merging the split patterns generated above.

- 8. Select Merge tool and click on the first isotope overlay
- 9. Drag the cursor to the location of the second isotope overlay, if it is the correct distance for a Charge state 1 compound ion then the second overlay will appear joined to the first

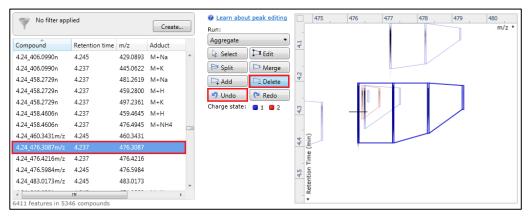


- 10. To complete the merge click the cursor again over the second isotope pattern
- 11. Before exploring the Add and Delete tools, click Undo to restore the original isotope pattern.

Add and Delete Tools

First use the Delete tool to remove the edited compound ion.

1. Select Delete and click on the compound ion in the ion map view





2. The blue, charge state 1, isotope pattern overlay will now be removed from the view.

ompound					Run:	-				m/z
	Retention time	m/z	Adduct		Aggregate 💌	4.1				
.24_406.0990n	4.245	429.0893	M+Na		Select Edit					
.24_406.0990n	4.237	445.0622	M+K		🖻 Split 🛛 🗁 Merge		. г.			_
.24_458.2729n	4.237	459.2800	M+H		Add 🕞 Delete	4.2				
.24_458.2729n	4.237	497.2361	M+K		🔊 Undo 陀 Redo			- I	1	
.24_458.2729n	4.237	481.2619	M+Na		Charge state: 1 2	_				
.24_458.4606n	4.237	459.4645	M+H			4.3				
.24_458.4606n	4.237	476.4945	M+NH4							
.24_460.3431m/z	4.245	460.3431				÷ 1	(uiu			
.24_476.4216m/z	4.237	476.4216				4.421	E.			
.24_476.5984m/z	4.245	476.5984					Ě			
.24_483.0173m/z	4.245	483.0173				5.5	lo			
.24_648.0331n	4.245	671.0252	M+Na	-			tent			
				P		-	Kete			

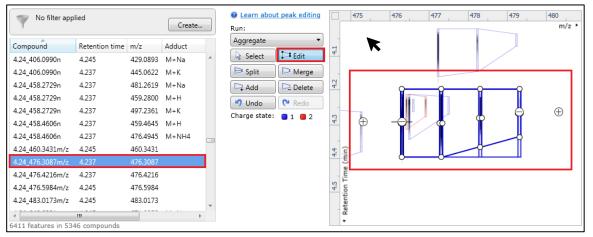
- 3. The isotope intensities will remain on the ion map, appearing as an 'undetected' isotopic pattern. The Compound entry in the table will also be deleted.
- 4. These can be restored using the **Undo** button, however, try using the **Add** button to create the peak picking overlay in the ion map window.

Click **Add** and then click on the top left of the first isotope and drag out an area covering the 4 isotope intensities.

1. As you drag the overlay out it will change colour reflecting the underlying pattern of intensity

W No filter app	blied	Crea	ate	Learn about peak editing Run:	475	476 477 478 479	QI Edit ion properties
Compound	Retention time	m/z	Add	Aggregate 🔹	4.1		Isotope width
4.24_458.2729n	4.237	459.2800	M+ ^	Select Edit			
4.24_298.0887m/z	4.237	298.0887		🖻 Split 🔛 Merge	2		Offset
4.24_406.0990n	4.237	445.0622	M+	📮 Add 📃 Delete	4		
4.24_458.4606n	4.237	459.4645	M+	💙 Undo 🛛 🍽 Redo			Ion charge
4.24_458.4606n	4.237	476.4945	M+	Charge state:	4.3		1
4.24_476.4216m/z	4.237	476.4216			4		
4.24_371.2275m/z	4.237	371.2275	_		-		
4.24_460.3431m/z	4.245	460.3431			4.4 (min)		
4.24_476.5984m/z	4.245	476.5984			e (1		
4.24_483.0173m/z	4.245	483.0173			- T		
4.24_771.2658m/z	4.245	771.2658			4.5 tion		
4.24_241.0748m/z	4.245	241.0748	-		4.5 Retention		Create Cancel
<			•		- Re		Create Cancel
6410 features in 53	45 compounds						

- 2. Use the property tools, (that appear as you release the cursor), to adjust the lsotope width and off set of the overlay on the underlying isotopic intensity pattern
- 3. Click Create to generate the new overlay and a new entry in the table
- 4. Click on the **Edit** tool and use it to refine the shape of the peak picking overlay by using the 'handles provided'



5. Finally, to set the final overlay click on an area outside the overlay on the ion map

Vo filter applied	Create	<u>Learn about peak editing</u> 475 476 477 478 478.481679 480 Run: m/z
4.24_406.0990n 4.237 445 4.24_406.0990n 4.237 445 4.24_458.2729n 4.237 481 4.24_458.2729n 4.237 459 4.24_458.2729n 4.237 497 4.24_458.2729n 4.237 459 4.24_458.4606n 4.237 476 4.24_458.4606n 4.237 476 4.24_458.4606n 4.237 476 4.24_458.4606n 4.237 476 4.24_476.3087m/z 4.237 476 4.24_476.4216m/z 4.237 476 4.24_476.5984m/z 4.245 476	Adduct 0893 M+Na 0622 M+K 2619 M+Na 2800 M+H 2361 M+K 4645 M+H 4945 M+H4 3431 3087 4216 5984 0173 •	Aggregate Select Split Merge Add Delete Undo Charge state: 1 2 C C C C C C C C C C C C C

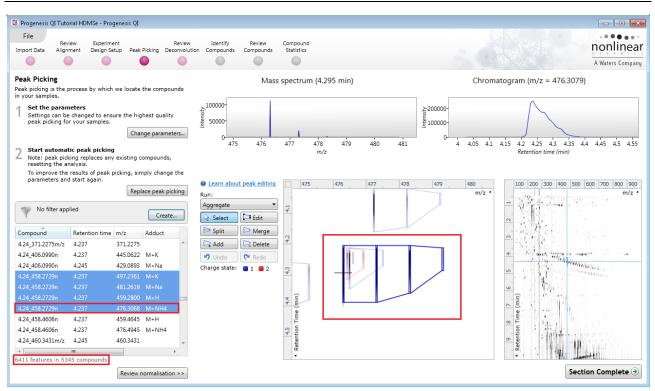
Before looking at the Normalisation if you have been using the Peak Picking Editing tools to explore their functionality you should redo the Peak Picking by clicking on **Replace Peak Picking**

🔲 Progenesis QI Tutorial HDMSe - Progenesis QI - 8 × File Review Experiment Review Identify Review Compound Statistics nonlinear Import Data Alignment Design Setup Peak Picking Deconvolution Compounds Compounds A Waters Company Peak Picking Chromatogram (m/z = 476.3079) Mass spectrum (4.295 min) Peak picking is the process by which we locate the compounds in your samples. 100000 Set the parameters ≥200000-1 Settings can be changed to ensure the highest quality peak picking for your samples. 50000 بة 100000 ا Change parameters... 0 476 477 2 Start automatic peak picking Note: peak picking replaces any existing compounds, resetting the analysis. To improve the results of peak picking, simply change the parameters and start again. 475 478 480 481 4.5 479 4.1 4.2 4.3 4.4 Retention time (min) **QI** Progenesis QI × Clear existing compound ions? ? 0 100 200 300 400 500 600 700 800 900 Replace peak picking Peak picking has already been performed in this experiment. Restarting will clear all existing compound ions and any subsequent analysis. m/z Rur m/z Ag ili -No filter applied Are you sure you want to clear the existing compound ions and start again? Create... B 2 Compound Retention time m/z Adduct ŝ Clear compound ions and start again Cancel 4.24_406.0990n 4.237 445.0622 M+K 4 4.24_458.2729n 4.237 481.2619 M+Na 🌱 Undo 🛛 🍽 Redo With Sugar 4.24_458.2729n 4 237 459 2800 M+H Charge state: 👩 1 🛑 2 ſ 64 4.24 458.2729n 4.237 497.2361 M+K 4.24_458.4606n 4.237 459.4645 M+H 9 4.24_458.4606n 4.237 476.4945 M+NH4 4.4 (nin (uiu 4.24_460.3431m/z 4.245 460.3431 , me Time ω. 4.5 4.24_476.4216m/z 4.237 476.4216 σ. 4.24 476.5984m/z 4.245 476.5984 Reten ✓ III 6411 features in 5346 compounds Section Complete ightarrowReview normalisation >>

Click on clear compound ions and start again

When the Peak Picking has completed, the table will update the number of Compound ions (features) and Compounds reported to 6411 and 5345

Progenesis QI User Guide



Note: When you use the editing tools, the ions they are applied to automatically become tagged in the table (with a grey tag) indicating that the ion has been edited in some way.

W No filter ap	plied			(Create	
Compound	Retention time	m/z	Adduct	Max Abundance	Tag 💌	
4.23_194.9752m/z	4.230	194.9752		112.9617		*
4.23_177.9819m/z	4.230	177.9819		19.5474		
4.24_458.2729n	4.237	476.3068	M+NH4	1.1454e+004		
4.24_458.2729n	4.237	497.2361	M+K	1614.6910		
4.24_458.2729n	4.237	481.2619		931.5427		
4.24_458.2729n	4.237	459.2800	M+H	5283.5576		
4.24_406.0990n	4.245	429.0893	M+Na	126.9125		
4.24_406.0990n	4.237	445.0622	M+K	338.7607		

Note: the editing of a compound adduct results in all the compound adducts for the deconvoluted compound being assigned an edit tag. In addition the edited compound (M+NH₄) ion loses its Adduct assignment.

At this stage all editing tags are removed when you replace the Peak Picking.

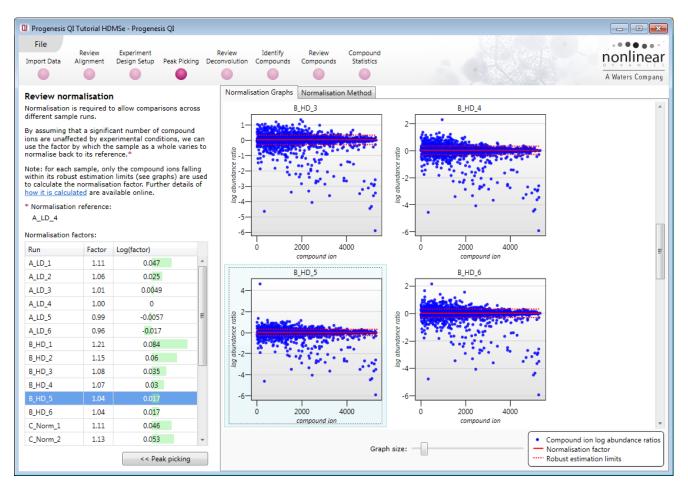
Normalisation of the data can be reviewed by clicking on **Review normalisation**.

Stage 6B: Reviewing Normalisation

The **Review Normalisation** page will open displaying plots for the normalisation of all the peaks 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.

The default method is to Normalise to all compounds.



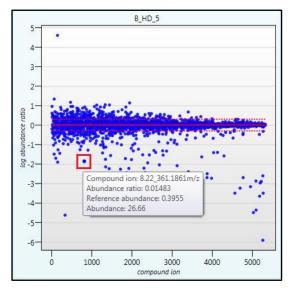
Normalisation factors are reported in the table to the left of the plots.

Calculation of Normalisation Factor:

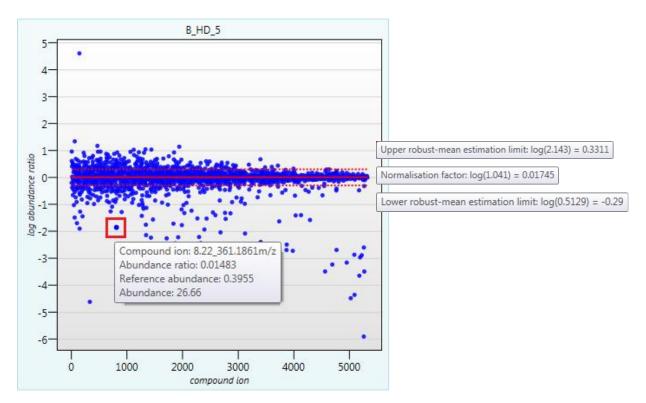
Progenesis QI will automatically select one of the runs that is 'least different' from all the other runs in the data set to be the 'Normalising reference'. The run used is shown above the table of Normalisation factors (in this example it is A_LD_4).

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

The details for individual Compound ions can be viewed as you hold the cursor over the dots on the plot. On the graph the compounds are shown ordered by ascending mean abundance.



The normalisation factor is then calculated by finding the mean of the log abundance ratios of the compounds that fall within the 'robust estimated limits' (dotted red lines). Compounds outside these limits are considered to be outliers and therefore will not affect the calculation of the normalisation factor.



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

Note: there are 5 Normalisation modes

Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics	nonlinear
			•	A Waters Company
Normalisatio	on Graphs No	ormalisation M	ethod	
Normalise to	o all compound	ls 🔻		
Normalise to	o all compound	ls		
Normalise u Normalise to	o a set of house sing total ion a o external stand ny normalisatio	bundance lard	oounds	

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

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

Now return to **Peak picking** by clicking on the button on the bottom left of the screen and the press **Section Complete** to move to the Review Deconvolution stage of the workflow.

Stage 7: Review Deconvolution

Following Peak Picking all the ions, including various adducts, for a compound are automatically recombined to provide accurate quantitation of each parent compound. Adducts of the same compound are grouped by a process called deconvolution. This stage of the workflow allows you not only to review the outcome of the deconvolution but also to add or remove adduct forms of the compound.



At this stage in the workflow you get the opportunity to review the process of Deconvolution for the Compounds. The table on the left displays the compounds ordered by the number of Adducts detected for each compound.

Note: the detected adducts are dependant on the list of expected adducts that you selected when you created the experiment

The total number of expected adducts is displayed as a montage at the top of the screen.

For an example: In **Compound 4.24_458.2729n**, 4 adducts have been detected: M+H, M+NH₄, M+Na and M+K (arrows) and displayed in the table. To find this compound easily, one may order the table by compound using the header or scroll down slightly on the Compounds table.

Tip: the 'About this' panel, top right summarises the information for the current compound. If not displayed click on the About this.

Grid size:

Grid size:

Grid size:

About this compound

Compound 4.24_458.2729n has the
following properties:

Retention time: 4.237 mins

Neutral mass: 458.2729

Adducts: 4 (M+H, M+NH4, M+Na,
M+K)

This compound's neutral mass and
adducts are based on mass differences
between its adduct forms.

The same details are available for the compound in the table to the bottom left.

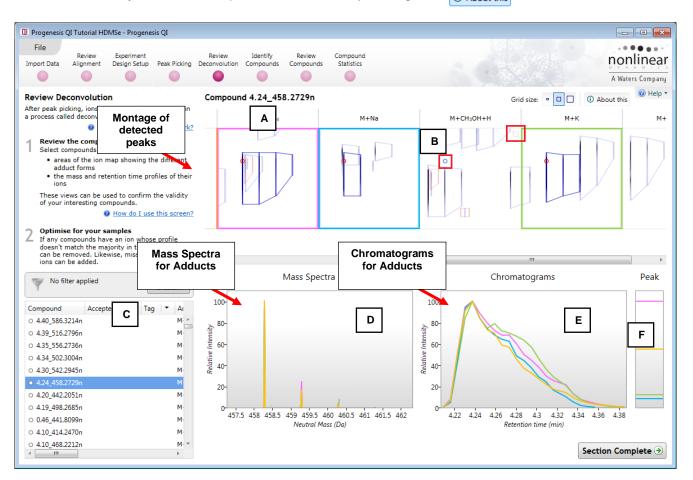
Use the Grid size to enlarge the montage view of the adducts.

Using **Compound 4.24_458.2729n** we can look at how the information is displayed on the multiple views following deconvolution.

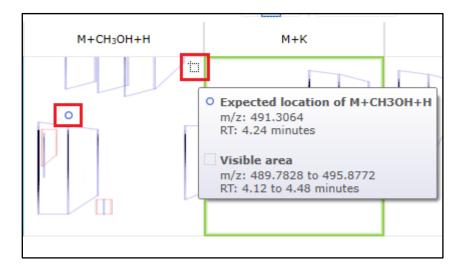
Waters

🕜 Help

Note: for clarity the 'About this' pane can be hidden by clicking on (i) About this



- (A) Montage view showing location of detected adducts for compound
- (B) Where an adduct has not been detected the expected location for the adduct is displayed on the relevant panel as a blue circle when you hover the cursor over the panel. In this case no Methanol adduct (M+CH₃OH + H) has been detected.

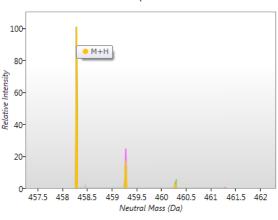


Note: hover the cursor over the icon on the top right of the panel and you will get information about the expected location and the area displayed.

(C) The table displays the information known about the current compound with regards to the number of adducts, identity of the adducts and where deconvolution has been successful the Compound's Neutral mass is displayed. Note: when you hover the cursor over the Neutral mass for the current compound it displays the adducts and their m/z values that were used to generate the the value. This is shown in an expanded view of the table below. Also shown is the 'tool tip' for the displayed m/z for each compound. An example is shown below for Compound 4.24_458.2729n.

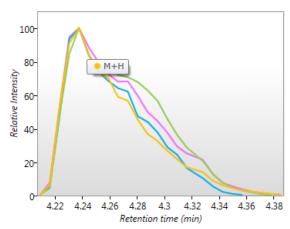
Compound	Accepted ID	Tag	•	Adducts	Adduct count	Neutral mass	m/z	Retention time	Fragmented?
O 4.34_502.3004n				M+H, M+NH4, M+Na, M+K	4	502.3004	520.3342		Yes
O 4.30_542.2945n				M+H, M+NH4, M+Na, M+K	4	542.2945	560.3283	4.302 m	u/z
• 4.24_458.2729n				M+H, M+NH4, M+Na, M+K	4	458.2729	476.3068		ne m/z of the most abundant compound n. To calculate this, each ion's abundance
O 4.20_442.2051n				M+H, M Based on these adducts:	4	442.2051	460.2389	4.202 is	defined as the sum of un-normalized
O 4.19_498.2685n				M+H, M • M+H at m/z=459.2800	4	498.2685	516.3024	4.187 al	oundances for that ion across all samples.
○ 0.46_441.8099n				M+H-2F • M+NH ₄ at m/z=476.3068	4	441.8099	406.7960	.463	Yes
O 4.10_414.2470n				• M+Na at m/z=481.2619 • M+K at m/z=497.2361	4	414.2470	432.2808	4.104	Yes
O 4.10_468.2212n				M+H, Mדוזווא, פודואם, פודוג	4	468.2212	486.2550	4.097	Yes

(D) This panel shows the Mass Spectra for the detected adducts of the Compound, colour coded by adduct, where the relative intensities are plotted against the **Neutral mass** scale.



Mass Spectra

Chromatograms



(E) This panel shows the Chromatograms for the detected adducts of the Compound, colour coded by adduct, where the relative intensities are plotted against the **Retention Time** scale.

Peak

● M+H

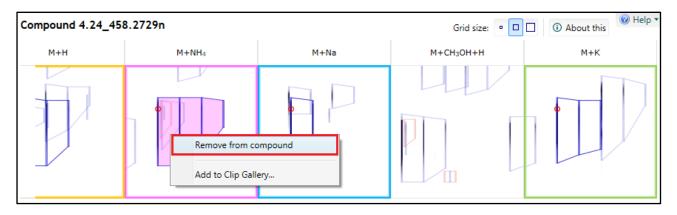
(F) This panel shows the actual Peak heights for the detected adducts of the Compound, colour coded by adduct.

Editing Compound adducts.

The number of adducts assigned to a compound by the process of deconvolution depends on the peak detection and the adducts selected at the beginning of the workflow when the experiment is created.

The process of deconvolution can only assign a Neutral mass for a compound if two or more adducts have been detected. The accurate quantitation of a compound is dependent on summing the intensities for a compounds adducts. In addition the accuracy of the quantitation for any compound is dependent not only on the detection of the adducts but also on the correct assignment during the process of deconvolution. In a complex sample there may be a need to add or remove adducts from a compound.

To remove an adduct right click on the assigned peak in the appropriate panel of the montage. So as an example we require to remove the $M+NH_4$ adduct for compound **Compound 4.24_458.2729n**, right click on the peak in the adduct panel and click **Remove from compound**.



In the table the number of adducts for the compound is reduced by 1 and a tag (grey) indicating that the adducts of this compound have been edited is created.



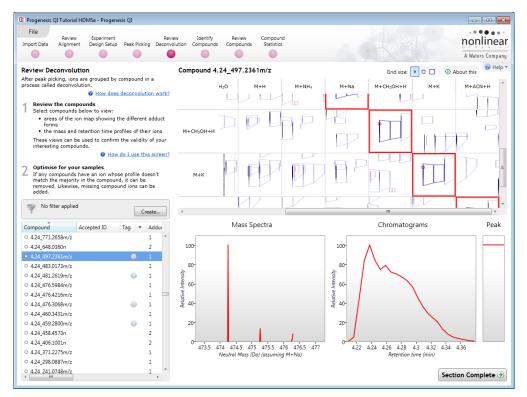
To add an adduct to a compound right click on a peak in the appropriate panel for the adduct and click **Add to compound**, again the table of compounds will update to reflect the change.

If you now remove 2 of the remaining 3 adducts the montage view will change to display the full matrix of all possible peak locations for expected second adducts dependant on the m/z values for these adducts. Also as there is only one adduct detected the Neutral mass for the compound can no longer be calculated.

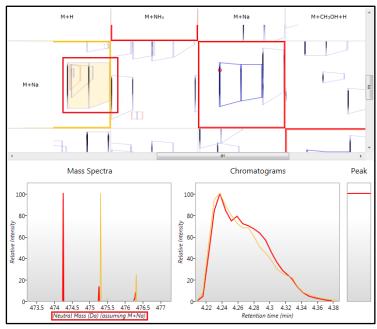
Try removing all the adducts from Compound 4.24_458.2729n to leave just the M+K.

The montage now turns into a matrix displaying all the possible locations for a charge state +1 Adduct in accordance with the list of expected adducts you selected at the beginning of the experiment. Also the individual Compound ions are displayed in the table i.e. **Compound 4.24_497.2361**

Note: because the peak is a charge state +1 this excludes M+2H and M+H+Na as being the compound's sole adduct on the basis of charge.

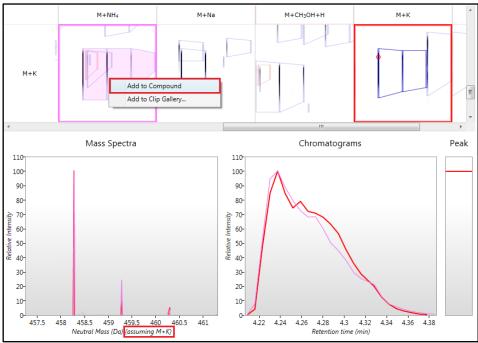


As an example to explore the possibility that **Compound 4.24_497.2361's** sole adduct is M+Na try hovering over the detected peaks for the expected adducts. In this case when you hover over M+H as the possible second adduct you can see that although the chromatograms elute at the same retention time the mass



spectra do not coincide. On this basis the second compound adduct cannot be M+H when M+Na is assumed to be the sole detected adduct.

Now try M+K as the assummed sole adduct of **Compound 4.24_497.2361**. This time when you hover over the cursor over the M+NH₄ adduct both the m/z on the mass spectra and RT on the chromatogram appear to coincide on the graphs with the corresponding location for the existing single adduct M+K.



To add the $M+NH_4$ adduct to the compound right click then add this second adduct to the compound. As you add the second adduct the table will update to indicate that the compound has 2 adducts, a **Neutral mass** is calculated (n appears after the compound in the table) and the matrix is replaced by a single row of adducts. This can be explored for the presence of additional adducts, in this case an additional 2 adducts can also be added (M+H and M+Na).

Exploring the expected location for the second adduct of a compound

After Peak picking there will be a list of compounds with a single adduct detected. For these compounds the review deconvolution montage will be displayed as a matrix where the process of deconvolution has been unable to assign a second adduct on the basis of any of the expected adducts displaying the correct m/z and retention time at the expected locations for any of the selected adduct ions.





Note: As you hover the cursor over any of the expected locations for the second adduct the Mass spec and Chromatogram for the second adduct are displayed in the bottom panels. If both the m/z and RT appear to coincide (in this example the 'double headed' arrows indicate that neither coincide) with the corresponding location for the existing single adduct then Add this second adduct to the compound by right clicking on it and selecting **Add to compound**.

By adding the second adduct the montage and table will update to reflect the addition of the adduct to the compound and a Neutral mass will be assigned based on the 2 adducts. The matrix collapses to the single row displaying the locations of the two detected adducts.

Using 'Tagging' to filter data for identification

In order to identify the compounds you can focus on data that is showing significant differences between one or more groups by first creating tags for particular 'subsets' of your data. Then filter the displayed data using the tags to provide a targeted list of compounds to identify.

Before creating new tags you can remove the automatically generated 'Edited' you created in the previous section. Right click on the table and select 'Edit tags' click **Delete tag** and then **OK**.

As an example of the use of tags we will set out to identify the compounds that show a significant (Anova p-value < 0.05), 2 fold or greater increase in abundance for one of the conditions (A_LD, B_HD or C_Norm).

Edit Tags

To create a tag for all compounds displaying an Anova p value <0.05, right click on the Compounds table and select Quick tags then select **Anova p-value...**

W No filter ap	oplied				Cre	ate	
Compound	Accepted ID	Tag	·	Adducts		Addu	
O 4.00_384.1998n	I			M+H, M+NH4, M+Na, M+K		4	*
O 4.33_486.2313n	i i i i i i i i i i i i i i i i i i i			M+H, M+Na, M+K		3	
o 5.00_440.1	No tags to assign	_		M+NH4, M+Na, M+K		3	
0.47_459.7				M+H, M+Na, M+K		3	
0 8.96_440.3	New tag			M+NH4, M+Na, M+K		3	
○ 9.57_582.4	Quick Tags	- •		Anova p-value		3	
0.46_195.8 🚰	Edit tags			Max fold change		3	
0.46_700.6	Add to Clip Gallery			Minimum CV		3	
0.55_331.0887m	I			Not identified		3	
0.46_405.7886n	I			Not fragmented		3	
○ 9.02_574.0949n				Separated by drift time		3	
0 8.96_396.2877n	I			Identified and separated by drift time		3	
0.47_353.7767m	i i i i i i i i i i i i i i i i i i i		_	M+H-H ₂ O, M+H, M+K		3	

🛛 New Quick Tag
Where a compound has: Anova p-value: < 0.05
Apply the following tag:
Anova p-value ≤ 0.05
Create tag Cancel

Accept the default value (≤ 0.05) and the offered name.

On pressing **Create tag** a tag appears in the table against all the compounds with an Anova p-value ≤ 0.05

💚 No fi	lter applied	ł			Crea	ate		
Compound		Accepted ID	Tag 💌	Adducts		Addu		
0 4.00_384.3	1998n			M+H, M+NH ₄ , M+Na, M+K		4		
0 4.33_486.2	2313n			M+H, M+Na, M+K		3		
o 5.00 440.	1694n		_	M+NH4, M+Na, M+K		3		
0.4: 🔴	Anova p-\	/alue ≤ 0.05	1	M+H, M+Na, M+K		3		
0 8.96	New tag			M+NH4, M+Na, M+K		3		
O 9.5	Quick Tag	ls ►	Ano	va p-value		3		
0.4	Edit tags		Max	fold change	1	3		
0.4	Add to Cli	p Gallery	Min	mum CV	1	3		
0.55_551.0	0.55_331.088/n			Not identified				
0.46_405.	7886n		Not	fragmented	I	3		
9.02_574.0	0949n		Sepa	Separated by drift time				
0 8.96_396.3	2877n		Iden	Identified and separated by drift time				
0.47_353.7	7767n			M+H-H2O, M+H, M+K	-	3		

Q New Quick Tag	×
Where a compound has: Max fold change: 2	
Apply the following tag: Max fold change ≥ 2	
Create tag Ca	ncel

To add a second 'Quick Tag' for those compounds with a Fold difference of 2 or greater, right click on the table to open the 'Tag' menu. Select Quick Tags and then Max fold change and accept the default value (\geq 2).

You will be offered a tag named **Max fold change \geq 2**.

On pressing **Create tag** a second tag appears in the table against all the compounds with a Max fold change \geq 2.

Now to get the table to display **only** those compounds which satisfy the required criteria, i.e. those compounds that display a significant 2 fold or greater increase in Mean abundance, apply a filter.

No filter ap	plied		Create	
Compound	Accepted ID	Tag 🔹	Adducts	
0 4.93_426.1888n			M+NH ₄ , M+Na, M+K	*
O 5.15_312.2079n		•	M+H-H ₂ O, M+H, M+NH ₄	
0 4.77_442.1479n			M+H-2H ₂ O, M+H-H ₂ O, M+H	
0 4.73_692.3840n			M+H, M+NH4, M+K	
○ 5.87_344.1838n		-	M+H-H ₂ O, M+Na, M+K	
0 6.22_310.0805n			M+H, M+Na, M+K	
0 6.67_342.1801n			M+H, M+Na, M+K	
0 7.36_470.2521n		-	M+NH4, M+Na, M+K	
0 7.64_484.2682n			M+NH4, M+Na, M+K	
0 8.61_668.3078n			M+H-H ₂ O, M+Na, M+K	
0 8.66_640.1120n		•	M+H, M+NH4, M+Na	
0 8.96_396.2877n		•	M+H, M+NH4, M+Na	*

Create.

No filter applied

To set up a filter click **Create** to open the Filter dialog and drag the new tags on to the **Show** compounds with all of these tags.

Q Filter the compounds	×	Q Filter the compounds	
Create a filter Show or hide compounds based on a se to create the filter. For more guidance,	election of their tags. Move tags to the appropriate boxes please see the <u>online reference</u> .	Create a filter Show or hide compounds based on a s to create the filter. For more guidance,	election of their tags. Move tags to the appropriate please see the <u>online reference</u> .
Available tags: Anova p-value ≤ 0.05 (1342 compo Max fold change ≥ 2 (488 compound Max fold change ≥ 2 (488 compound)		Available tags:	Show compounds that have all of thes: Anova p-value ≤ 0.05 (1342 compo Max fold change ≥ 2 (488 compounds) Show compounds that have at least or these tags: Hide compounds that have any of these
<u>Clear the filter</u>	OK Cancel		ОК Са

On clicking **OK** the table now displays only those compounds with the 2 tags. This example is ordered by number of adducts (adduct count)

Compound	Accepted ID	Tag 💌	Adducts
• 7.13_276.1366n		۵	M+H-2H ₂ O, M+H-H ₂ O, M+H, M+Na, M+
○ 7.11_308.1061n		۷	M+H-H ₂ O, M+H, M+Na, M+K
○ 8.02_322.2744n		۲	M+H, M+Na, M+K
○ 6.22_308.0834n		۲	M+H, M+Na, M+K
○ 7.11_308.1991n		۷	M+H, M+Na, M+K
○ 6.50_270.1040n			M+H, M+Na, M+K
○ 8.01_322.1212n		۹	M+H, M+Na, M+K
○ 8.21_322.1212n		۵	M+H, M+Na, M+K
○ 8.20_322.2747n		۲	M+H, M+Na, M+K
○ 0.81_619.2019n		۲	M+H+Na, M+K
○ 8.02_200.0836n		۵	M+H-2H ₂ O, M+Na

Note: the Tag filter panel has changed, informing you that a filter is currently applied.

With the filter still active click Section Complete to move to the Identify Compounds stage.

Stage 8: Identify Compounds

The user guide now describes how to identify compounds from **all or a subset** of compounds.

Note: that the filter you applied at the previous stage is currently active.

To display all the compounds in the table click **Edit** and then click on **Clear the filter** and then **OK**.

To Find **Compound 3.77_412.0365n**, which has a neutral mass of 412.0365, either scroll to the right in the table or expand the table by dragging the icon to the right of the table. Then order the table on descending number of adducts (Adduct count).

Search	for identification	s	1
	Create		egend:
pted ID	Tag 💌 1		

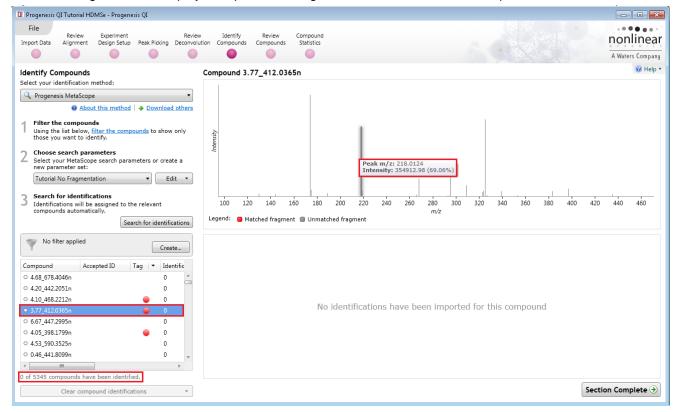
Edit.

Tag filter applied

compounds may be hidden

T									Create	:
Compound	Accepted ID	Tag 🔻	Identifications	Fragmented?	Neutral mass	m/z	Drift time	Adduct count		
○ 4.15_428.2268n			0	Yes	428.2268	4	• • •	-	_	
○ 7.13_276.1366n		4	0	Yes	276.1366	2 The m/	z of the mo	ost abundant c	ompound	
○ 4.60_634.3787n			0	Yes	634.3787	6 ion. To	calculate tl	his, each ion's sum of un-norm	abundance	
○ 4.68_678.4046n			0	Yes	678.4046			at ion across a		
○ 4.20_442.2051n			0	Yes	442.2051	460.2389	4.16	4		
○ 4.10_468.2212n			0	Yes	468.2212	486.2550	4.59	4		
• 3.77_412.0365n		-	0	Yes	412.0365	413.0437	3.73	4		
○ 6.67_447.2995n			0	Yes	447.2995	448.3068	4.81	4		
○ 4.05_398.1799n			0	Yes	398.1799	416.2137	3.83	4		
○ 4.53_590.3525n			0	Yes	590.3525	608.3863	5.67	4		
○ 0.46_441.8099n			0	Yes	441.8099	406.7960	2.92	4		
○ 4.19_498.2685n			0	Yes	498.2685	516.3024	4.75	4		
○ 0.41_217.9763n			0	Yes	217.9763	182.9624	1.94	4		
0 9.95 412.3191n			0	Yes	412.3191	435.3083	4.97	4		Ŧ

At this stage the view displays the pattern of fragmentation for the selected compound



Note: the m/z displayed in the table is for the most abundant compound ion (in this case the M+H adduct). Currently there are 5345 compounds of which none have been identified.

Identification of compounds using the search method Progenesis MetaScope

This identification method is designed to support identifications from a number of different databases.

Depending on the available information in the database and the type of search being performed you can create and save re-usable **search parameter sets**.

For example: searching a database using compound values for Mass and Retention Time.

🔲 Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment Review Ide Import Data Alignment Design Setup Peak Picking Deconvolution Comp		nonlinear
	MetaScope search parameters	A Waters Company
Identify Compounds Compou	Define a set of MetaScope parameters that can be saved for later reuse. Learn more in the <u>online reference</u> .	i Help ▼
Select your identification method:	Name:	
Progenesis MetaScope	Tutorial No Fragmentation	
About this method + Download others	Compound database	
1 Filter the compounds	C:\Program Files (x86)\Nonlinear Dynamics\Prc Browse	
Using the list below, filter the compounds to show only those you want to identify.	Data format: Auto-detect	
Choose search parameters		
Z Select your MetaScope search parameters or create a	Search parameters	
new parameter set:	Precursor tolerance: 12 ppm 🔻	
Tutorial No Fragmentation Edit	Retention time within: 0.1 minutes	
2 Search for identifications	CCS within: 2.5 %	
 Identifications will be assigned to the relevant 100 compounds automatically. 		300 320 340 360 380 400 420 440 460
Search for identifications	Additional compound properties source	
	Read additional compound properties from this file	
Vo filter applied	<no database="" selected=""> Browse</no>	
Compound Accepted ID Tag 🔻 Identific	Fragment search method	
○ 4.68_678.4046n 0	On not use fragmentation data	
0 4.20_442.2051n 0	Perform theoretical fragmentation	
○ 4.10_468.2212n 😑 0	Fragment tolerance: 12 ppm	
• 3.77_412.0365n 🥚 0	Perform fragment database search	nported for this compound
0 6.67_447.2995n 0	<no database="" selected=""> Browse</no>	
○ 4.05_398.1799n 😑 0	Fragment tolerance: 12 ppm 🔻	
• 4.53_590.3525n 0		
0.46_441.8099n 0		
< m >>	Save search parameters Cancel	
0 of 5345 compounds have been identified.		
Clear compound identifications		Section Complete 🥑

First select the identification method from the top left of the screen. In this example for the Tutorial data use Progenesis MetaScope, a flexible search engine which is designed to work with databases where you can set thresholds for Mass, Retention time, Collisional Cross Section and Theoretical Fragmentation depending on the format and content of the database.

Set the Precursor tolerance to 12ppm, in this case you can use the **Tutorial No Fragmentation** search parameters as shown above.

Browse to locate the database file.

Location of example Structure Data Format (SDF) files after install:

C:\Program Files (x86)\Nonlinear Dynamics\Progenesis QI\Plugins\MetaScopeSearch\SDF Databases\

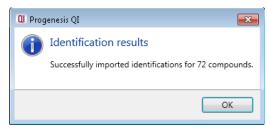
Select: Tutorial Fragmentation.sdf (if not already selected)

Progenesis QI User Guide

There are a number of Data Formats available including **Auto-detect**, which we will use here.

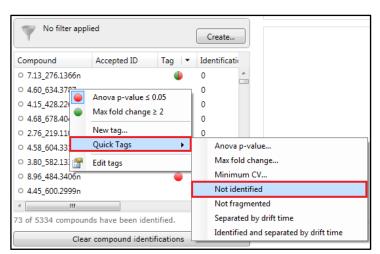
Save the Search Parameters and click Search for identifications.

A dialog will tell you the number of compounds that you have imported identifications for.



To focus only on the identified compounds, hide all the unidentified compounds. To do this you must first tag all the unidentified compounds. Right click on the table and select the Quick Tag for **Not identified.**

Then create a filter that hides these compounds by selecting Create and dragging the **Not identified** tag on to the **Hide** panel.



Click **OK** to apply the filter.

Having imported the search results you can see there are a number of possible identifications for the identified compounds.

No filter applied

Possible identifications for the current compound are ranked on an overall Score based on, Mass error, Retention Time error (if available), Isotope Similarity (calculated from the comparison of the measured isotope distribution for the compound vs the theoretical based on the compound formula, Fragmentation Score and the compound's Collisional Cross-Section (CCS) if available.

f their tags. Move tags to the appropriate boxes to he <u>online reference</u> .
Show compounds that have all of these tags:
Show compounds that have at least one of these tags:
Hide compounds that have any of these tags:
Not identified (5273 compounds)

🚭 Tag filter app	lied			Pos	sible identific	ations: 4							
Compounds m	ay be hidden	Edit			Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation score	Mass error (ppm)	0
Compound	Accepted ID	Tag 💌		*	46506142	Paracetamol	M+H-H	C ₈ H ₉ NO ₂	2.90	52	0	0.49	
4.68_179.0954n				\Rightarrow	HMDB02210	2-Phenylglycine	M+H-H	$C_8H_9NO_2$		39.2	0	0.49	
9.48_382.2750n				☆	HMDB12219	Dopamine quinone	M+H-H	$C_8H_9NO_2$		39.2	0	0.49	HN
2.96_314.1423n		4		\Rightarrow	HMDB29703	Methyl 2-aminobenzoate	M+H-H	C ₈ H ₉ NO ₂		39.2	0	0.49	
5.18_151.0634n		4											
5.18_179.0953n		4											Í
• 2.91_151.0634n		4											
3.63_179.0946n													\sim
0.07_413.2657m/z			-										
< <u> </u>		•											I OH
All 72 filtered compo	unds have been i	dentified.		4		III						•	

Find **Compound 2.91_151.0634n.** The solid grey icon to the left of the **Compound** column indicates possible identifications have been made for this compound. With the available Retention time data for Paracetamol this would indicate that Paracetamol is a stronger match.

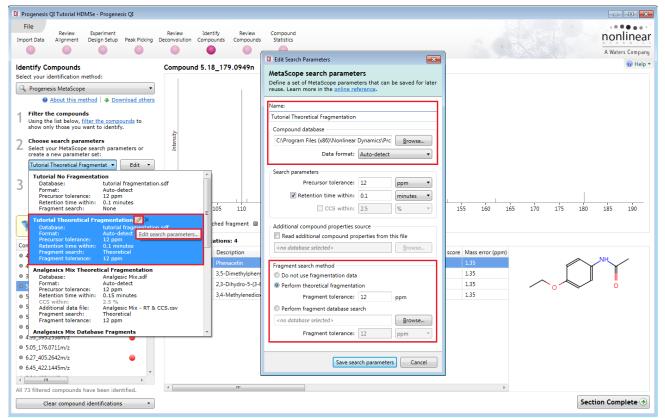
Tag filter app compounds m		Edit		Pos	sible identific	ations: 4							
 compounds m 	ay be nidden	Contin		☆	Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation score	Mass error (ppr	0
Compound	Accepted ID	Tag 💌		\star	46506142	Paracetamol	M+H-H	C ₈ H ₉ NO ₂	2.90	38.5	0	3.51	Ī
• 7.11_308.1061n			*	☆	HMDB02210	2-Phenylglycine	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.51	[
6.22_308.0834n		۱		\Rightarrow	HMDB12219	Dopamine quinone	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.51	HN
8.53_382.2724n				\pm	HMDB29703	Methyl 2-aminobenzoate	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.51	1
6.50_270.1040n		4											
• 2.96_314.1423n													
• 2.91_151.0639n		4											
4.30_162.0684n													
4.68_179.0954n													\checkmark
3.63_179.0946n			-										
<		÷.		_									он
All 73 filtered compo	unds have been i	dentified.		۲.								+	

Note: when the Retention time data is **not** included as part of the search scoring (by unticking the option on the search parameters) then there is no difference in the scoring for the 4 identification candidates in the table to the right.

Incorporating Theoretical Fragmentation

To improve the confidence in the compound identification you can choose to perform theoretical fragmentation of a candidate list of compounds and then match the resulting 'in silico' fragmentation against the measured/observed fragments for a compound. The candidate molecules are selected from a compound database based on the exact mass (within a specified error range given in (ppm).

The compound database is in SDF format providing the structural information for each candidate compound. Using this list of candidates, the fragmentation algorithm generates all possible fragments for a candidate compound in order to match the fragment mass with the measured peaks.

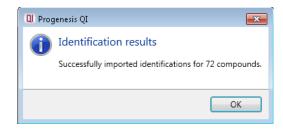


To perform theoretical fragmentation using the **tutorial fragmentation.sdf** as the source of the candidate compounds select the second Search Parameter Set (Tutorial Theoretical Fragmentation) from the drop down

Note: you can review/edit the parameter set, either by clicking on the 'pencil' icon in the drop down or using the **Edit** option to the right of the drop down.

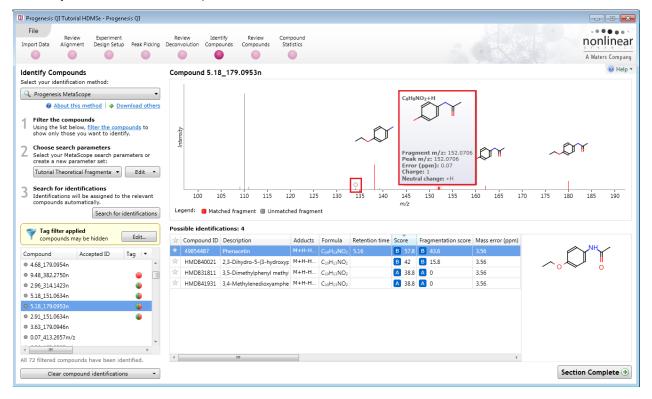
Ensure the **Perform theoretical fragmentation** from the Fragment search method is selected and then save search parameters. To perform the Theoretical Fragmentation search click on **Search for identifications**

As before search 72 results will be returned.

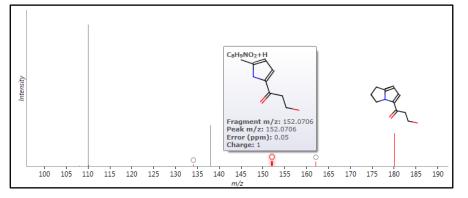


 About this method A Downlos Filter the compounds Using the list below, filter the compounds show only those you want to identify. Choose search parameters Select your MetaScope search parameter set: Tutorial Theoretical Fragmenta Iterational State 	
Using the list below, filter the compoun show only those you want to identify. Choose search parameters Select your MetaScope search parameter create a new parameter set:	<u>ds</u> to
show only those you want to identify. 2 Choose search parameters Select your MetaScope search parameter create a new parameter set:	<u>ds</u> to
2 Choose search parameters Select your MetaScope search parameter create a new parameter set:	
Select your MetaScope search parameter create a new parameter set:	
create a new parameter set:	ers or
Tutorial Theoretical Fragmenta 💌 🛛 🛛	
	dit 🔻
Search for identifications	
3 Identifications will be assigned to the re-	levant
compounds automatically.	act and

Where fragments have been measured for a compound and a match is made with the theoretical fragments from each candidate compound a solid coloured line appears on the Fragmentation graph. This is accompanied with either an image of the fragment displayed (if there is room) or an open symbol at the top of the line which enables display of the fragment when the cursor is held over it (as shown below for **Compound 5.18_179.0953n**).



The fragmentation score for each possible identity appears in the table. When you click on one of the possible identifications with a lower fragmentation score the Fragmentation graph displays which fragments contribute to that score for example the **2nd ranked compound** Fragmentation matches (see below).



Note: a fragmentation score of 0 indicates either that **no** match was achieved between the theoretical fragments (of a candidate compound) and the measured fragments or in fact the compound has no fragmentation data. For example, in **Compound 5.84_448.3061m/z**:

Q Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment Import Data Alignment Design Setup Peak Picking	Review Identify Review Compound Deconvolution Compounds Compounds Statistics	nonlinear A Waters Company
Identify Compounds Select your identification method:	Compound 5.84_448.3061m/z	iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii
Progenesis MetaScope About this method Download others		
Filter the compounds Using the list below, filter the compounds to show only those you want to identify.		
2 Choose search parameters Select your MetaScope search parameters or create a new parameter set:		This compound has no fragmentation data
Tutorial Theoretical Fragmentat Edit Search for identifications		
 Identifications will be assigned to the relevant compounds automatically. Search for identifications 	Possible identifications: 2	
Search for identifications	Compound ID Description Adduct	ts Formula Retention time Score Fragmentation score
Tag filter applied	* HMDB40994 2-(2-Methylbutanoyl)-9-(3- M+NH.	
Compounds may be hidden	🚖 HMDB41990 piritramide M+NH	4 C ₂₇ H ₃₄ N ₄ O
Compound Accepted ID Tag 👻		
5.18_179.0949n 5.18 151.0637n		
• 5.84_448.3061m/z		
 5.96_422.1466m/z 6.00_163.0395m/z 		FY
< + +		
All 73 filtered compounds have been identified.	< III	, , , ,
Clear compound identifications -		Section Complete 🕥

Note: where multiple searches have been performed the search yielding the highest overall score for each candidate is displayed. Also the search parameters giving rise to these scores are displayed as tooltips when the cursor is held over the Search Parameter indicator.

Q Progenesis QI Tutorial HDMSe - Progenesis QI			
File			
Review Experiment Import Data Alignment Design Setup Peak Picking	Review Identify Review Compound Deconvolution Compounds Compounds Statistics		nonlinear
			A Waters Company
Identify Compounds	Compound 2.91_151.0634n		🛞 Help 🔻
Select your identification method:	compound 2.91_151.0054m		
Q Progenesis MetaScope	0		
About this method + Download others			
 Filter the compounds 			
Using the list below, filter the compounds to show only those you want to identify.	-ti-		
	Intensity		
2 Choose search parameters Select your MetaScope search parameters or			
create a new parameter set:			
Tutorial Theoretical Fragmentar Edit	° .		Ŷ
3 Search for identifications Identifications will be assigned to the relevant	100 105 110 115	120 125 130	135 140 145 150 155 160
compounds automatically.		m/z	155 140 145 150 155 160
Search for identifications	Legend: 🧧 Matched fragment 🗐 Unmatched frag	Iment	
Tag filter applied	Possible identifications: 4		
compounds may be hidden	📩 Compound ID Description Addu	cts Formula Retention time Score Frag	gmentation score Mass error (ppm)
Compound Accepted ID Tag 👻	* 46506142 Paracetamol M+H-		73.7 0.49
• 4.68_179.0954n	* HMDB12219 Dopamine quinone M+H-	Singhoz Si	Search configuration: B
© 9.48_382.2750n	HMDB02210 2-Phenylglycine M+H-		Method: Progenesis MetaScope
© 2.96_314.1423n 🌒	HMDB29703 Methyl 2-aminobenzoate M+H-	m C ₈ H ₉ NO ₂	Database: tutorial fragmentation.sdf
© 5.18_151.0634n			Precursor tolerance: 12 ppm Retention time within: 0.1 minutes
• 5.18_179.0953n			Theoretical fragment tolerance: 12 ppm
• 2.91_151.0634n			
 3.63_179.0946n 0.07_413.2657m/z 			Ý
(
All 72 filtered compounds have been identified.	٠		, ОН
Clear compound identifications			Section Complete)
cical compound identifications			

The table of compounds can be reduced to displaying only those compounds that have been identified, show a significant difference between the conditions, and are altered by at least two-fold, by applying a 'Tag Filter' based on the assigned 'Quick tags.

	on a selection of their tags. Move tags to the appropriate boxes to the please see the <u>online reference</u> .
Available tags:	Show compounds that have all of these tags:
	Anova p-value ≤ 0.05 (1342 compounds)
	Max fold change ≥ 2 (488 compounds)
	Show compounds that have at least one of these tags:
	Hide compounds that have any of these tags:
	Not identified (5273 compounds)

The table now displays only information for 21 identified compounds.

File New With States Service X Service X <th>0 Progenesis QI Tutorial HDMSe - Progenesis QI</th> <th></th> <th></th>	0 Progenesis QI Tutorial HDMSe - Progenesis QI		
Tedentification with a submatched is advanced by a submatched of the relevant compound a trepted in the relevant of the relevant compound a trepted in the relevant of the relev	Review Experiment	Deconvolution Compounds Compounds Statistics	D Y N A M I C S
create a new parameter set: Γutorial Theoretical Fragmenta • Edit • 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 μegend: • Matched fragment • Unmatched fragment • • • • • • • • • • • • •	Select your identification method:		@ Help •
trigg max points may be hidden Edit * 4702 Warfarin M+H, C10 Hu/CO4 7.12 B 63.5 B 47.7 4.05 Compound Accepted ID Tag •	Create à new parameter set: Tutorial Theoretical Fragmenta Edit Search for identifications Identifications will be assigned to the relevant compounds automatically.	100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 Legend: Matched fragment Unmatched fragment Possible identifications: 1	
Clear compound identifications	compounds may be hidden Edit Compound Accepted ID Tag • 7.11_308.1061n • • • 6.50_270.1040n • • • 6.22_308.0834n • • • 2.91_151.0634n • • • 2.96_314.1423n • • • 5.18_179.0953n • •	* 4702 Warfarin M+H, C ₁₃ H ₁₆ O ₄ 7.12 B 63.5 B 47.7 4.06	

Creating and using databases

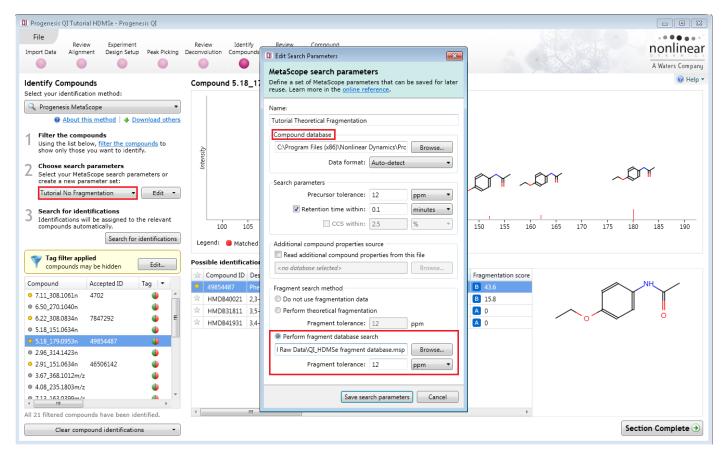
The identification data can be stored and used to perform fragment database searches. This allows you to build up your own (local) databases of fragment information. These can then be used to perform future identifications based on fragment matches when using the same compound database. To create and subsequently export data to such a database you must indicate compounds to export by accepting their identity.

	File				Review	Ide	tify Review	Compound	d							
-	Save			D	econvolution											nonlinea
1	Close				•										7	A Waters Compan
	Export compoun	d measurements			Compour	nd 5.1	8_179.0953n (Phenacetin	1)							🕜 Help
	Export adduct m	easurements d identifications				QI Exp	oort fragment datal	pase				• ×				
	Export compoun Export isotope in								a you want to expo	rt.						
	Export fragment	database				0 C	ompounds with ac	cepted IDs 🔘	All compounds							
_		compound prop	erties	_	2	V	Compound	Accepted ID	Adducts							
	Export to pathwa				Intensity	V	2.91_151.0634n	46506142	M+H-H ₂ O, M+H							
	Import compoun	, id identifiers as ta	q		Int		3.69_195.0881m/z		M+H				<u> </u>		<pre>///</pre>	
	Export to EZinfo.		-			V	5.18_179.0953n	49854487	M+H-H ₂ O, M+H						~~ "	
	Open EZinfo pro						6.22_308.0834n	7847292	M+H, M+K, M+N							
		·				V	7.11_308.1061n	4702	M+H, M+K, M+N	a, M+H-H ₂ O						
	Export inclusion	list														
	Experiment prop	erties			10						/ 5 compounds	colocted	160 165	170	175 180	185 190
	Show Clip Gallery	y			Legend:						5 compounds	selected.				
×	Exit				Possible i		bout fragment dat	<u>abases</u>			Export	Cancel				
_				_	🚖 Comp	ound IE	Description	Ado	ducts Formula	Retention ti	me Score F	Fragmentation	score			
	mpound	Accepted ID	Tag 🔻		* 49854		Phenacetin	M+	H-H C10H13NO2	5.16	B 57.8	B 43.6			NH	
	7.11_308.1061n	4702	4	*	🚖 HMD	B40021	2,3-Dihydro-5-(3-	-hydroxyp M+	H-H C ₁₀ H ₁₃ NO ₂		B 42	B 15.8			r T	T
	6.50_270.1040n				🚖 HMD	B31811	3,5-Dimethylpher	nyl methy M+	H-H C ₁₀ H ₁₃ NO ₂		A 38.8	A 0				
- 1	6.22_308.0834n 5.18 151.0634n	7847292		=	🚖 HMD	B41931	3,4-Methylenedic	xyamphe M+	H-H C ₁₀ H ₁₃ NO ₂		A 38.8	A 0	-	0	~	Ŭ
_	5.18_151.0634n 5.18 179.0953n	49854487	() ()													
	2.96 314.1423n	49604467	.													
- 1	2.90_314.1423h 2.91 151.0634n	46506142	-													
- 1	3.67 368.1012m/;															
- 1	4.08_235.1803m/;															
	7 12 162 0200m/-	-		Ŧ												
٦ ا		unds have been	,		4			_								

Using the Filter set created for the 21 compounds in the previous section, accept the identity of a compound by clicking on the 'star' icon to the left of the Compound ID, it will turn yellow. Do this for a number of compounds displaying an acceptable fragmentation score.

Then select **Export fragment database...** from the file menu as shown above. Waters When you export these accepted identifications the file is saved as a Mass Spectrometry Profile (msp) file.

i.e. QI_HDMSe fragment database.msp (example: in the folder you restored the Tutorial Data Set.zip).



You can add to this file as you gather additional information from other experiments.

Now as an example re-perform the Tutorial Theoretical Fragmentation but this time instead of selecting the **Perform theoretical fragmentation** option select the **Perform fragment database search** option. Then locate the .msp file you have just created.

Then save the Search parameters and perform the Fragment database search by clicking on **Search for identifications**.

Note: you must use the same Compound database that was used when the original msp file was generated. This is because the initial step of Progenesis QI using a fragment database is to find all possible identifications based on neutral mass (RT and CCS used if available). Then for each possible identification (candidate) Progenesis QI searches the fragment database for a record with a matching compound ID. It then compares the stored fragment data to the current experimental fragmentation data.

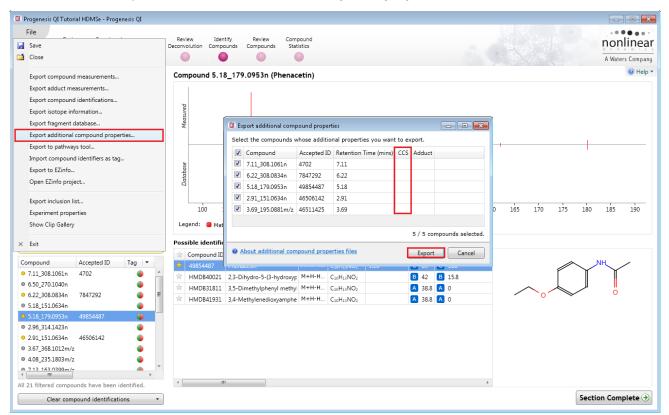
On performing the search you will get a 'mirror plot' where the upper half represents the measured/observed fragmentation and the lower half shows the currently stored fragmentation in the database. This gives rise to a **Fragmentation score** describing how good the match is between the observed fragmentation and the stored fragmentation in the database.

The fragmentation score is based on the number of fragments matched and their intensity. This excludes any contribution from the matching of the precursor.

The current search with the highest score is designated C in this example (Compound 5.18_179.0953n).

File Review Experim nport Data Alignment Design S		Review Iden Deconvolution Compo		Compound Statistics										nonline
														A Waters Comp
lentify Compounds		Compound 5.1	8_179.0953n (Pher	nacetin)										🕜 He
elect your identification method:														
🔍 Progenesis MetaScope	•													
Ø About this method	Download others	Irred												
Filter the compounds		Measured												
Using the list below, filter the	compounds to	<				1								
show only those you want to it	lentity.													
Choose search parameters Select your MetaScope search														
 Select your MetaScope search create a new parameter set: 	parameters or	Database												
Tutorial No Fragmentation	▼ Edit ▼	atab												
Easych for identifications														
Search for identifications Identifications will be assigned	to the relevant		, , , , , , , , , , , , , , , , , , , ,											
 Identifications will be assigned compounds automatically. 			05 110 115	120 125	5 130	135 140	145	150 155	160	165	170	175	180	185 190
 Identifications will be assigned compounds automatically. 	to the relevant	100 1	05 110 115 ched fragment Dumm			135 140	145 m/z	150 155	160	165	170	175	180	185 190
Identifications will be assigned compounds automatically. Search		100 1 Legend: Mat	ched fragment 🔲 Unm			135 140		150 155	160	165	170	175	180	185 190
 Identifications will be assigned compounds automatically. 		100 1 Legend: Mat	ched fragment 🔲 Unm ations: 4	atched fragme	nt		m/z			165	170	175	180	185 190
Identifications will be assigned compounds automatically. Search Tag filter applied compounds may be hidden	for identifications	100 1 Legend: Mat Possible identific	ched fragment Unm ations: 4 Description	atched fragme Adducts	nt Formula	Retention time	m/z Score	Fragmentation		165	170	175	180	185 190
Identifications will be assigned compounds automatically. Search Tag filter applied compounds may be hidden Compound Accepted ID	for identifications	100 1 Legend: ● Mat Possible identific ☆ Compound ID ☆ 49854487	ched fragment Unm. ations: 4 Description Phenacetin	Adducts	nt Formula C ₁₀ H ₁₃ NO ₂		m/z Score C 69	Fragmentation		165	170	175	180	185 190
Identifications will be assigned compounds automatically. Tag filter applied compounds may be hidden Compound Accepted ID 9 7.11_308.1061n 4702	for identifications	100 1 Legend: ● Mat Possible identific ☆ Compound ID ☆ 49854487 ★ HMDB40021	ched fragment Unm. ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro	Adducts M+H-H xxyp M+H-H	nt Formula C10H111NO2 C10H113NO2	Retention time	m/z Score C 69 B 42	Fragmentation C 100 B 15.8		165	170	175	180	185 190
Identifications will be assigned compounds automatically. Tag filter applied compounds may be hidden Compound Accepted ID P.11_308.1061n 4702 © 6.50_270.1040n	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me	Adducts Adducts M+H-H oxyp M+H-H thyi M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170	175	180	
Identifications will be assigned compounds automatically. Tag filter applied compounds may be hidden icompound Accepted ID 7.11_308.061n 4702 6.50_270.1040n 9.622_308.0834n 7847292	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ched fragment Unm. ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro	Adducts Adducts M+H-H oxyp M+H-H thyi M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42	Fragmentation C 100 B 15.8 A 0		165	170	175	180	
Identifications will be assigned compounds automatically. Search Tag filter applied compounds may be hidden Compound Accepted ID 7.11_308.1061n 4702 6.50_27.0140n 6.50_270.1440n 9.518_151.0634n 7847292	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me	Adducts Adducts M+H-H extyp M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170	175	180	
Identifications will be assigned compounds automatically. Search Tag filter applied compounds may be hidden Compound Accepted ID 7.11_308.1061n 4702 6.50_270.1040n 6.50_270.1040n 6.52_308.0834n 7847292 5.18_151.0634n 5000000000000000000000000000000000000	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me	Adducts Adducts M+H-H extyp M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170	175	180	
Identifications will be assigned compounds automatically. Search Tag filter applied compounds may be hidden Compound 7.11_308.1061n 4702 9.52.720.1040n 6.622.308.0834n 6.622.308.0834n 7847292 9.518_151.0634n 49854487	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me	Adducts Adducts M+H-H extyp M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170	175	180	185 190
Identifications will be assigned compounds automatically. Tag filter applied compounds may be hidden Tog filter applied compound Accepted ID 7.11_308.1061n 4702 6.50_270.1040n 5.02,70.1040n 5.18_151.0634n 5.18_151.0634n 9 5.18_150.0553n 49854487 9 2.96_314.1423n	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me	Adducts Adducts M+H-H extyp M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170	175	180	185 190
Identifications will be assigned compounds automatically. Tag filter applied compounds may be hidden Compound Accepted ID > 7.11_308.1061n 4702 > 6.50_270.1040n 5.18_179.0953n > 5.18_179.0953n 49554487 > 2.96_31.1423n 46506142	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me	Adducts Adducts M+H-H extyp M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170	175	180	
Identifications will be assigned compounds automatically. Tag filter applied compounds may be hidden Compound Accepted ID 7.11_308.1061n 4702 6.50_270.1040n 6.22_308.0834n 7847292 5.18_173.0953n 49854487 9.296_314.1423n 9.291_151.0634n 4.08_235.1803m/z 3.13_68.012m/z 3.13_68.2032m/z	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me	Adducts Adducts M+H-H extyp M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170	175	180	
Identifications will be assigned compounds automatically. Search Tag filter applied compounds may be hidden Search Tog filter applied compound Accepted ID 7.11_308.1061n 4702 6.50_270.1040n 6.62_308.0834n 6.22_308.0834n 7847292 5.18_151.0634n 49554487 2.96_314.1423n 2.91_151.0634n 3.67_368.1012m/z 46506142 3.67_368.1012m/z 46506142	for identifications	100 1 Legend: Mat Possible identific Compound ID 49854487 HMDB40021 HMDB31811	ched fragment Unm. ations: 4 Description Phenacetin 2,3-Dihydro-5-(3-hydro 3,5-Dimethylphenyl me 3,4-Methylenedioxyam	Adducts Adducts M+H-H extyp M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	Retention time	m/z Score C 69 B 42 A 38.8	Fragmentation C 100 B 15.8 A 0		165	170 0	175	180	

You can also manage and search against additional compound properties. The advantage of this is that you can create a database of compound data that retains information on an identified compound for m/z, retention time. Export and save this **Additional Compound properties** file.



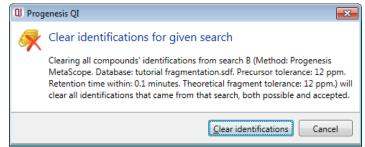
Note: if the data format contains Drift time data and it has been calibrated then the value for a compounds Collisional Cross Section (CCS) may also be stored as an additional compound property and used in the process of compound identification.

Deleting Compound Identifications

Before using the **additional compounds properties** file, you can explore how to clear compound identifications. If more than one Search Configuration has been used then you can remove identifications based on the Configuration. To do this click on **Clear compound identifications** and select the configuration.

QI Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment Import Data Alignment Design Setup Peak Picking	Review Identify Review Compound Deconvolution Compounds Compounds Statistics	nonlinear A Waters Company
Identify Compounds Select your identification method: Progenesis MetaScope • About this method ◆ Download others • About this method ◆ Download others • Jitter the compounds • Using the list below, filter the compounds to show only those you want to identify. • Choose search parameters Select your MetaScope search parameters or create a new parameter set: • Tutorial No Fragmentation • • Select for identifications • Identifications will be assigned to the relevant compounds submatically.	Compound 5.18_179.0953n (Phenacetin)	W Help ▼
	Possible identifications: 4	
Tag filter applied compounds may be hidden Edit	☆ Compound ID Description Adducts Formula Retention time Score Fragmentation score	A NU
Compound Accepted ID Tag 🔻		
• 7.11 308.1061n 4702	Image: Microsoft Micros	
● 6.50_270.1040n	☆ HMDB31811 3,5-Dimethylphenyl methyl M+H-H C10H13NO2 ▲ 38.8 ▲ 0	
6.522,308.0834n 7847292 5.18,151.0634n Clear identifications from search A (Progenesis Met Clear identifications from search B (Progenesis Met Clear identifications from search C (Progenesis Met Clear all identifications	Scope)	
Clear compound identifications 👻		Section Complete 🏵

When deleting identifications by Search configuration (in this example B) you will be warned which ones will be cleared.

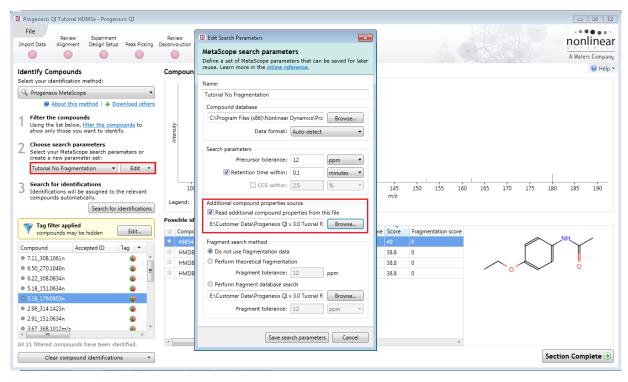


Following deleting the identifications based on Search configuration B there will be 3 possible remaining identifications for **Compound 5.18_179.0953n**

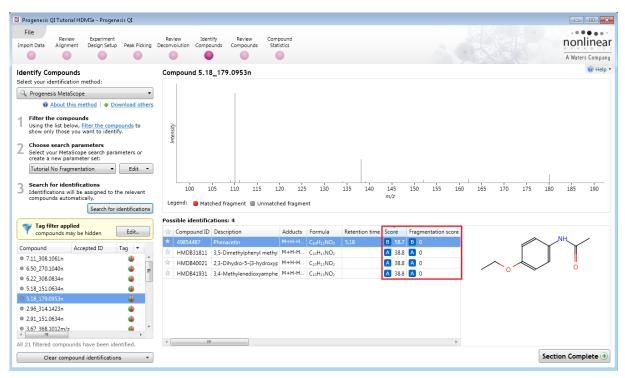
Now clear all the compound identifications and re-perform the **Tutorial No Fragmentation** search as described on page 45.

🐨 Tag filter applied 🦳	Pos	sible identific	ations: 3				_		
compounds may be hidden	Edit	Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation score	
	*	49854487	Phenacetin	M+H-H	C10H13NO2	5.16	C 69	C 100	NH
	ag 💌 👷	HMDB31811	3,5-Dimethylphenyl methyl	M+H-H	C10H13NO2		A 38.8	A 0	
0 7.11_308.1061n	2 🕺 📩 🖄	HMDB41931	3,4-Methylenedioxyamphe	M+H-H	C10H13NO2		A 38.8	A 0	
⊃ 6.50_270.1040n	• E								✓ 0 ♥ 0
6.22_308.0834n	•								
5.18_151.0634n	•								
• 5.18_179.0953n	4								
0 2 96 31/ 1/23n									
A Clear identifications from search A (Pr	ogenesis MetaScope	e)							
Q Clear identifications from search C (Pr	ogenesis MetaScope	≥)							
X Clear all identifications			·						
									Section Complete
Clear compound identifications	•								Section Complete

To make use of the **Additional compound properties file**, you created on page 46, either edit, copy or make a new set of search parameters using this file as the source. In the example below we have edited the existing search parameters for the **Tutorial no Fragmentation** search parameters.



When you re-perform the search a higher overall score is achieved using the additional scored compound data.



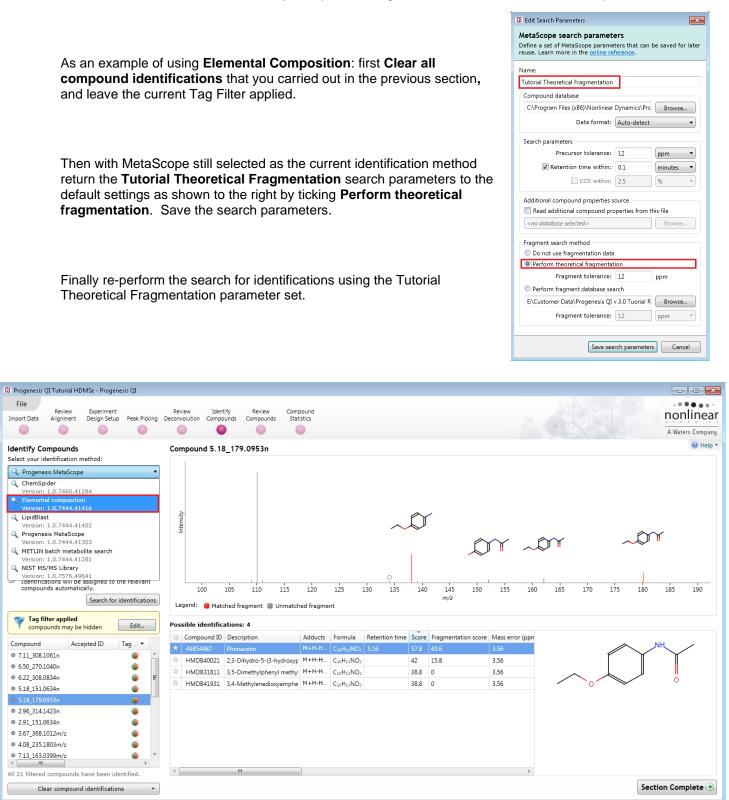
Note: the saved **Additional Compound Properties** file is the source of the 'in house' saved compound properties (Retention time and CCS) while the Compounds database file provides the neutral mass and structure details.

Each matching component (where available) contributes 20% to the overall score: Mass error, Retention time, Isotope similarity, Fragmentation score and Collisional Cross-Section (CCS).

i.e. if only mass error and isotope similarity are available then the maximum score will be out of 40.

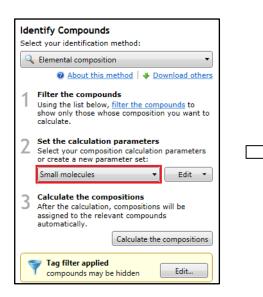
Additional Identification Methods

In addition to Metascope and Metlin, Progenesis QI v3.0 has identification methods for Elemental Composition, LipidBlast, ChemSpider and NIST. With the exception of NIST, these methods are installed by default and are available at the Identify Compounds stage from the identification method drop down.



Then select Elemental composition as the Identification methods.

There are 3 default parameter sets: Small molecules, Lipids and CHNO each containing a list of Elements to be considered and thresholds for **Precursor tolerance (ppm/Da)** and **Isotope similarity (%)**.



	tify Compounds t your identification met	hod:	
Q	Elemental composition		•
	About this method	od 🖊 Download	l others
	Filter the compounds Using the list below, filte show only those whose c calculate.		
L	Set the calculation par Select your composition or create a new paramet	calculation param er set:	
	Small molecules	▼ Ed	it 🔻
3	Small molecules 🖌 Precursor tolerance: Isotope similarity: Elements:	5 ppm	
	Lipids Precursor tolerance: Isotope similarity: Elements:	3 ppm 95.00% C, H, N, O, P	itions
Con	CHNO Precursor tolerance: Isotope similarity: Elements:	3 ppm 95.00% C, H, N, O	•

Select the **Small molecules** parameter set and click **Calculate the compositions**.

Click OK to import the 10 identifications.

🔲 Prog	enesis QI
()	Identification results
	Successfully imported identifications for 10 compounds.
	ОК

The **Possible identifications** table for the current compound

updates to include any compositions that meet the current calculation parameters.

	Deconvolution Compounds Compounds	Compound Statistics			A Waters Co	i c s	
Identify Compounds Select your identification method:	Compound 5.18_179.0953n					Help *	
C Elemental composition							
About this method Download others							
Filter the compounds Using the list below, filter the compounds to show only those whose composition you want to calculate.	ntensity						
2 Set the calculation parameters Select your composition calculation parameters or create a new parameter set: Small molecules	Int	I					
3 Calculate the compositions After the calculation, compositions will be assigned to the relevant compounds automatically. Calculate the compositions	Legend: 😑 Matched fragment 🗐 Unn	20 125 130 135 140 natched fragment	145 150 155 m/z	160 165 170 175	180 185 1	90	
😴 Tag filter applied	Possible identifications: 5					_	
compounds may be hidden Edit	☆ Compound ID Description ☆ 49854487 Phenacetin		Adducts Formula M+H-H2 C10H13NC				
Compound Accepted ID Tag 💌	+ HMDB40021 2,3-Dihydro-5-(3-hydr	ownronanoul) 14 purrelizing	M+H-H2 C10H13NC				
• 7.11_308.1061n	C10H13NO2 Formula C10H13NO2						
• 6.50_270.1040n	☆ HMDB31811 3,5-Dimethylphenyl m		M+H-H2 C10H13NC				
• 6.22_308.0834n	☆ HMDB41931 3,4-Methylenedioxyan	,	M+H-H2 C10H13NC				
© 5.18_151.0634n				No struct	ure data		
• 5.18_179.0953n							
2.96_314.1423n 4 2.91 151.0634n		Possible identifications: 5					
2.91_151.0634n 3.67_368.1012m/z		Score Fragmentation score	Mass error (ppm)	Retention time error (mins)	Isotope similarity	Link	Search Configuration
e 108 225 1802m/r		A 57.8 A 43.6	3.56	0.02	98.32	🔮 <u>pubche</u>	A Method: Progenesis MetaSco
< >	< <u> </u>	A 42 A 15.8	3.56		98.32	S nonlinea	. A Method: Progenesis MetaScor
All 21 filtered compounds have been identified.		B 38.8 B 0	3.56		98.32	-	. B Method: Elemental composition
Clear compound identifications -		A 38.8 A 0	3.56		98.32	nonlinea	_
		A 38.8 A 0	3.56		98.32	-	
		🔼 20.0 📉 U	00.0		90.32	🔇 <u>nonlinea</u>	. A Method: Progenesis MetaSco

Note: For the example shown an additional compound ID appears as you accept the identifications based on the calculation of the Elemental Composition.

Scrolling along to the right the actual Mass error and Isotope similarity values are displayed. The **B** indicates the type of search applied.



Note: in this case the only elemental composition returned (using the current thresholds) corresponds to the formula returned for the other possible identifications.

To edit/create/copy a parameter set select from the **Edit** drop down to the right.

2	Set the calculation parameters Select your composition calculation or create a new parameter set:	parameters
	Small molecules 🔹	Edit 🝷

As an example, create a new parameter set by first copying the existing default set for small molecules.

Select Create a copy... from the Edit drop down, the edit dialog for the calculation parameters will appear.

Vame: Small mo	lecules (2)		н																	He
Compositi	ion:			Li	Be											В	с	Ν	0	F	Ne
Element H	From 0	To 50	Remove?	Na	Mg											AI	Si	Р	s	CI	Ar
С	0	50	×	К	Ca	Sc	Ti	v	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
N	0	5	×	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Те	т	Xe
0	0	20	×	ND.	31	'	21	IND	IVIO		Nu	NI	Fu	Ag	cu	m	311	30	ie		~~
S	0	2	×	Cs	Ba		Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	TI	Pb	Bi	Po	At	Rn
				Fr	Ra		Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	FI	Uup	Lv	Uus	Uu
						La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Но	Er	Tm	Yb	Lu	
olerance	s: tolerance	e: 5	ppm 🔻			Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	

Rename the Parameter set (Small Mol plus P), then add Phosphorous (P) to the composition by clicking on it in the Periodic table.

Then set the maximum number of Phosphates for the composition by adjusting the value in the 'To' column (i.e. 4). Increase the Precursor tolerance to 10ppm.

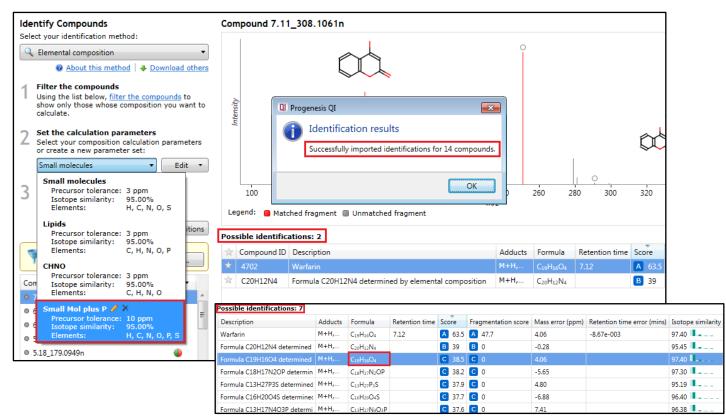
Elemen	tal con	npositio	on calculation p on calculati composition ca	on pa	rame		s that	can be	saved	for lat	er reus	se. Lea	rn moi	re in th	e <u>onlir</u>	ne refe	rence.				×
Name: Small Mc	ol plus P			н																	He
Composit				Li	Be											В	с	N	0	F	Ne
Element H	From 0	To 50	Remove?	Na	Mg											AI	Si	Р	s	CI	Ar
С	0	50	×	к	Ca	Sc	Ti	v	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
N	0	5	×	Rb	Sr	Y	Zr	Nb	Мо	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
O P	0	20 4	×	Cs	Ba		Hf	Та	w	Re	Os	Ir	Pt	Au	Hg	TI	Pb	Bi	Po	At	Rn
S	0	2	×	Fr	Ra		Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	FI	Uup	Lv	Uus	Uuo
												_			_			-			
Tolerance	s:					La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	
Precursor	toleranc	e: 10	ppm 🔻			Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	
Isotope si	milarity:	95	%																		
																		Save		Car	icel

Finally Save the new parameter set.

Note: on saving, the new parameter set becomes current and its parameters are listed in the drop down.

Make compound 7.11_308.1061n current in the table and click Calculate the composition.

When the compositions are calculated using the adjusted composition and tolerances an additional 5 possible ids appear in the list.



Note: Adding phosphorous to the composition parameter set and increasing the Precursor tolerance returns more possible identifications including the formula for Warfarin which had previously been excluded due to the previously low Precursor tolerance.

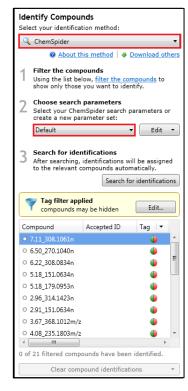
Running a ChemSpider Search

ChemSpider is a web-based chemical structure database with access to over 32 million structures from hundreds of data sources. This method makes use of those ChemSpider web services, automatically exporting data from Progenesis QI to ChemSpider for searching according to the parameters you select, importing the results, and assigning them against the correct compounds within the software.

Note: use of ChemSpider requires internet access and is a service provided to users with a valid Progenesis QI support contract.

Select ChemSpider as the identification method.

If there are already compound identifications present from the previous sections, before performing a search first click on **Clear compound identifications** and click **Clear all identifications**.



Progenesis QI User Guide

ChemSpider can access a large number of data sources when you Search for identifications.

With the search parameters set to Default Click Edit.

The first time you use ChemSpider, KEGG is selected as the default Data Source.

Untick the Perform theoretical fragmentation option

To add additional data sources or change the current data source Click **Select data sources...**.

The ChemSpider data source dialog opens. This enables you to search and select from the 276 data sources currently accessed by ChemSpider.

Add the most appropriate source(s) for your current search to the right hand panel.

Tip: since the ChemSpider search is based on accurate mass matching, searching a large number of data sources can take an appreciable amount of time.

The selected Data sources will be added to the Search Parameters. **Note**: if more than five are selected then a number is reported for those used in addition to the 5 displayed.

Then click Search for identifications.

ame: Default		
Required search parame	eters	
Precursor tolerance:	5	ppm ·
Precursor tolerances a minimum of 0.0001 Data Data source:		
	(Select data sources.
Optional search parame	ters	
Perform theoretical	fragmentation	
Fragment tolerance:	5	ppm
Filter by isotope sim	ilarity score	
Isotope similarity:	95	%
Filter by elemental c	omposition	
Elemental composition	P: 0-2 S: 0-2	50 N: 0-10 O: 0-30
	Select eler	nental composition.

Search 🔎		
276 data sources left to choose from: abcr Activate Scientific		data sources will be searched: KEGG
ACToR: Aggregated Computational Tox Adrian Hobson Advanced ChemBlocks AK Scientific AKos	< Remove << Remove all	
Alfa Aesar Alfa Chemistry Alichem		
Alinda Chemical Alkamid Amadis Chemical Ambeed		
AnalytiCon Discovery		

🛛 Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment Import Data Alignment Design Setup Peak Pickin	Review Identify Review Compound Deconvolution Compounds Compounds Statistics	nonlinea
		A Waters Company
Identify Compounds Select your identification method: ChemSpider About this method > Download other I Using the list below, filter the compounds to show only those you want to identify. Choose search parameters Select your ChemSpider search parameters or create a new parameter set: Default Edit Search for identifications After searching, identifications will be assigned to the relevant compounds automatically.	Searching for identifications (86/119 masses)	Cancel
Tag filter applied compounds may be hidden Edit Compound Accepted ID Tag • 7.11.308.1061n • • • 6.50.270.1040n • • • 6.52.308.0834n • • • 5.18.151.0634n • • • 2.90.51.0634n • • • 2.91.51.0634n • • • 0 450.314.1423n • • • 0 518.1170m/r • • • 0 518.1170m/r • •	No identifications have b	Identification results Successfully imported identifications for 19 compounds OK
Clear compound identifications		Section Complete 🤿

🔲 ChemSpider Data Sou

Choose ChemSpide

The progress bar reports the number of masses being searched. When the search completes the number of compounds with imported identifications is displayed.

Note: although Phenacetin is identified as one of the possible identities for **Compound 5.18_179.0953**, based on a mass error tolerance of 5ppm there are 9 other possibilities with the same mass error.

Progenesis QI Tutorial HDMSe - Progenesis QI												
File Review Experiment port Data Alignment Design Setup Peak Picking I	Review Identi Deconvolution Compo.											nonlin
												A Waters Co
entify Compounds lect your identification method:	Compound 5.18	3_179.0953n										0
ChemSpider -												
About this method												
Filter the compounds Using the list below, <u>filter the compounds</u> to show only those you want to identify.	ntensity											
Choose search parameters Select your ChemSpider search parameters or create a new parameter set: Default • Edit •	Inte											
									1			
Search for identifications After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications		105 110 115 hed fragment I Unmatched f		125 13	30 135		145 150 v/z	155 160	165	170	175 180	0 185 190
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications	Legend: 🗧 Match	ched fragment 🔲 Unmatched f	ragment			m	o/z				175 180	0 185 190
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications	Legend: Match Possible identifica	thed fragment Unmatched f ations: 10 Description	Adducts For	rmula Rete	ention ti Score	Fragmentation	Mass error (ppm)	155 160 Retention time error	Isotope sim		175 180	0 185 190
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit.	Legend: Match Possible identifica Compound ID CSID13491	thed fragment Unmatched f ations: 10 Description 4-amino-3-phenylbutyric acid	Adducts For M+H-H2 C108	rmula Rete H ₁₃ NO ₂	ention ti Score 38.8	Fragmentation	Mass error (ppm) 3.52		Isotope sim 98.32		175 180	D 185 190
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag V	Legend: Match Possible identifica Compound ID CSID13491 CSID16141	thed fragment Unmatched f ations: 10 Description	Adducts For	rmula Rete H ₁₃ NO ₂	ention ti Score	Fragmentation 0 0	Mass error (ppm)		Isotope sim 98.32		175 180	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag <	Legend: Match Possible identifica Compound ID CSID13491	thed fragment Unmatched f ations: 10 Description 4-amino-3-phenylbutyric acid	Adducts For M+H-H2 C108	mula Rete H ₁₃ NO ₂	ention ti Score 38.8	Fragmentation	Mass error (ppm) 3.52		Isotope sim 98.32			
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag • Accepted ID Tag • Accepted ID Tag •	Legend: Match Possible identifica Compound ID CSID13491 CSID16141	thed fragment Unmatched f atlons: 10 Description 4-amino-3-phenylbutyric acid MPMC	Adducts Form M+H-H2 C10H M+H-H2 C10H	mula Rete H ₁₃ NO ₂ H ₁₃ NO ₂ H ₁₃ NO ₂	ention ti Score 38.8 38.8	Fragmentation 0 0	Mass error (ppm) 3.52 3.52		Isotope sim 98.32		0	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit T1_3081061n 622_308.0834n	Legend: Match Possible identifica Compound ID CSID13491 CSID16401 CSID16606	hed fragment Unmatched f ations: 10 Description 4-amino-3-phenylbutyric acid MPMC XMC	Adducts For M+H-H2 C10H M+H-H2 C10H M+H-H2 C10H	mula Rete H ₁₃ NO ₂ H ₁₃ NO ₂ H ₁₃ NO ₂ H ₁₃ NO ₂	ention ti Score 38.8 38.8 38.8 38.8	Fragmentation 0 0 0	Mass error (ppm) 3.52 3.52 3.52		Isotope sim 98.32			
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag 7.11_308.106.1n 650_270.1040n 652_2308.0834n 518_151.0634n	Legend: Match Possible identifica Compound ID TC CSID13491 CSID16141 CSID16606 CSID2006639	ted fragment Unmatched f tions: 10 Description 4-amino-3-phenylbutyric acid MPMC XMC L-Homophenylalanine	Adducts Form M+H-H2 C10F M+H-H2 C10F M+H-H2 C10F M+H-H2 C10F M+H-H2 C10F M+H-H2 C10F	rmula Reter H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2	ention ti Score 38.8 38.8 38.8 38.8 38.8 38.8	Fragmentation 0 0 0 0	Mass error (ppm) 3.52 3.52 3.52 3.52 3.52		Isotope sim 98.32 98.32 98.32 98.32			
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit. mound Accepted ID Tag • Accepted ID Tag •	Legend: Matcl Possible identifica CSID13491 CSID13491 CSID16046 CSID2006639 CSID2006639 CSID20083	hed fragment Unmatched f atoms: 10 Description 4-amino-3-phenylbutyric acid MPMC XMC L-Homophenylalanine IPC	Adducts For M+H-H2 C10 ^b	rmula Reter H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2 H13NO2	ention ti Score 38.8 38.8 38.8 38.8 38.8 38.8 38.8 38.	Fragmentation 0 0 0 0 0 0 0	Mass error (ppm) 3.52 3.52 3.52 3.52 3.52 3.52		Isotope sim 98.32 98.32 98.32 98.32 98.32 98.32 98.32			
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit Tag 1000 Accepted ID Tag V 7.11_308.1051n 6.22_308.0834n 5.18_15.10534n 2.18_1290923n 296_314.1423n	Legend: Matcl Possible identification Compound ID	hed fragment Unmatched f ations: 10 Description 4-amino-3-phenylbutyric acid MPMC XMC L-Homophenylalanine IPC Fusaric acid	Adducts Form M+H-H2 C10 ¹	mula Rete H ₁₃ NO ₂ H ₁₃ NO ₂	ention ti Score 38.8 38.8 38.8 38.8 38.8 38.8 38.8 38.	m Fragmentation 0 0 0 0 0 0	Mass error (ppm) 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52		Isotope sim 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32			
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit. mpound Accepted ID Tag • Accepted ID Tag • Accepte	Legend: Matcl Possible identifica ☆ Compound ID ☆ CSID13491 ☆ CSID16401 ☆ CSID16401 ☆ CSID2006639 ☆ CSID23083 ☆ CSID23242	Here fragment Unmatched f Unmatched f Description 4-amino-3-phenylbutyric acid MPMC XMC L-Homophenylalanine IPC Furaric acid phenacetin	Adducts Form M+H-H2 C10 ¹	mula Rete H13NO2 H13NO2	ention ti Score 38.8 38.8 38.8 38.8 38.8 38.8 38.8 38.	m Fragmentation 0 0 0 0 0 0 0 0 0	Mass error (ppm) 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52		Isotope sim 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32			
After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit. Tag offilter applied compounds may be hidden Edit. Tag offilter applied Compounds may be hidden Edit. Tag of the compound of the compound for the compound of the compound for the compound of the compound for the compound of the compound of the compound for the compound of the compound of the compound for the compound of the compound of the compound of the compound for the compound of t	Legend: ● Matcl Possible identifica ☆ ☆ CSD13491 ☆ CSD13491 ☆ CSD1541 ☆ CSD206639 ☆ CSD2324 ☆ CSD3324 ★ CSD4590 ☆ CSD4990	hed fragment Unmatched f attorns: 10 Description 4-amino-3-phenylbutyric acid MPMC XMC L-Formophenylalanine IPC Fusaric acid phenprobamate	Adducts For M+H-H2 Cub M+H+H2 Cub M+H+H2 Cub	rmula Rete H13NO2 H13NO2 H13NO2 H13NO2	ention t Score 38.8 38.8 38.8 38.8 38.8 38.8 38.8 38.	m Fragmentation 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Vz Mass error (ppm) 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52		Isotope sim 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32			
After searching, identifications will be assigned to the relevant compounds audomatically. Search for identifications Tag fifter applied compounds may be hidden Edit mpound Accepted ID Tag • Accepted ID Tag	Legend: Matcl Possible identification Compound ID ☆ Compound ID ☆ CSD13491 ☆ CSD16461 ☆ CSD16606 ☆ CSD206639 ☆ CSD206639 ☆ CSD206639 ☆ CSD20663 ☆ CSD4066 ☆ CSD6906	hed fragment Unmatched f atoms: 10 Description 4-amino-3-phenylbutyric acid MPMC XMC L-Homophenylalanine IPC Fusaric acid phenprobamate Risocaine	Adducts Form M+H-H2 C10 ^k	rmula Rete H13NO2 H13NO2 H13NO2 H13NO2	ention ti Score 38.8 38.8 38.8 38.8 38.8 38.8 38.8 38.	Fragmentation 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Mass error (ppm) 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52		Isotope sim 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32 98.32			

The ChemSpider search parameters includes 3 additional options that can be selected to modify the search:

Theoretical fragmentation can be performed on the search results returned by the ChemSpider search using a specified Fragment tolerance.

Isotope similarity filtering discards ChemSpider results where the isotope similarity score falls below a specified cut-off. **Note**: the score is calculated by comparing theoretical and measured isotope patterns as described for the MetaScope search engine.

Select elemental composition which compares the formula returned by ChemSpider to a specification based on a list of elements and their allowed count ranges.

QI Edit ChemSpider Search	Parameters	×								
ChemSpider search parameters Define a set of ChemSpider parameters that can be saved for later reuse. Learn more in the <u>online reference</u> .										
Name:										
Default										
Required search parame	ters									
Precursor tolerance:	5	ppm 🔻								
Precursor tolerances an minimum of 0.0001 Da										
Data source:	KEGG.									
	Se	elect data sources								
Optional search parame	ters									
Perform theoretical f	ragmentation									
Fragment tolerance:	5	ppm								
Filter by isotope simi	ilarity score									
Isotope similarity:	95	%								
Filter by elemental co	omposition									
Elemental composition	: C: 0-100 H: 0-150 P: 0-2 S: 0-2	N: 0-10 O: 0-30								
	Select eleme	ental composition								

Performing Theoretical Fragmentation on ChemSpider Results

As an example, tick the **Perform theoretical fragmentation** option with the **Fragment tolerance** set to 5ppm.

Save the parameter set and re-perform the search.

When the ChemSpider search is re-performed, using the additional information provided by performing theoretical fragmentation on the results this now provides a level of order to the 10 possible identifications for **Compound 5.18_179.0953** making Phenacetin to be the more likely compound identification.

File												
Review Experimen port Data Alignment Design Setu		Review Ident Deconvolution Compo										nonli
												A Waters (
entify Compounds		Compound 5.18	170.09535								The second s	
lect your identification method:		compound 5. re	5_179.095511									
ChemSpider	•											
About this method	Download others											
Filter the compounds												
Using the list below, filter the con		2					\sim					
show only those you want to ider	ntify.	Intensity				\sim						
Choose search parameters		In I					×.	~	~ •	\sim	~	
Select your ChemSpider search p create a new parameter set:	parameters or									J		•••
Default	▼ Edit ▼											
												1
						0						
Search for identifications	di ka andarad	100	105 110 115	120	125		140	145 150	155 160	165	170 175	180 185 1
Search for identifications After searching, identifications wi to the relevant compounds auton	vill be assigned matically.	100	105 110 115	120	125	130 135	140	145 150 m/z	155 160	165	170 175	180 185 1
After searching, identifications wi to the relevant compounds auton	vill be assigned matically. or identifications		105 110 115 ched fragment II Unmatched f		125				155 160	165	170 175	180 185 1
After searching, identifications wi to the relevant compounds auton Search fo	matically.		ched fragment 🔳 Unmatched f		125				155 160	165	170 175	180 185 1
After searching, identifications wi to the relevant compounds auton Search fo	matically.	Legend: 📒 Mato	ched fragment 🔲 Unmatched f		125 Formula					165	170 175	180 185 1
After searching, identifications wi to the relevant compounds auton Search for Tag filter applied compounds may be hidden	matically. or identifications Edit	Legend: 🗧 Mato	ched fragment 🔲 Unmatched f	fragment Adducts		130 135 Retention ti Score		m/z		165	170 175	180 185 1
After searching, identifications wi to the relevant compounds auton Search for Tag filter applied compounds may be hidden mpound Accepted ID	Edit	Legend: Mate	thed fragment Unmatched f ations: 10 Description	Adducts	Formula	130 135 Retention ti Score B 47.6	Fragmentatio	m/z on Mass error (ppm		165	170 175	180 185 1
After searching, identifications wi to the relevant compounds auton Search for compounds may be hidden mpound Accepted ID 7.11_308.1061n	matically. or identifications Edit	Legend: ■ Mate Possible identifica ☆ Compound ID ★ CSID4590	thed fragment Unmatched f ations: 10 Description phenacetin	Adducts M+H-H2 M+H-H2	Formula C10H13NO2	130 135 Retention ti Score B 47.6 B 41.9	Fragmentatic B 43.6	m/z on Mass error (ppm 3.52		165	170 175	HN
After searching, identifications wi to the relevant compounds autor compounds may be hidden mpound Accepted ID 7.11_308.1061 6.50_270.1040n	Edit	Legend: ■ Mate Possible identifica ☆ Compound ID ★ CSID4590 ☆ CSID82699	ched fragment Unmatched f ations: 10 Description phenacetin Salsolinol	Adducts M+H-H2 M+H-H2	Formula C10H13NO2 C10H13NO2	Retention ti Score B 47.6 B 41.9 B 41.8 B 41.8	Fragmentatic B 43.6 B 15.5	m/z on Mass error (ppm 3.52 3.52		165	170 175	180 185 1
After searching, identifications wi to the relevant compounds autor Compounds sumary be hidden mpounds (Accepted ID 7.11_3016in 6.22_308.0834n	Edit	Legend: ■ Mato Possible identifico ☆ Compound ID ★ CSID4590 ☆ CSID82699 ☆ CSID324	thed fragment Unmatched f ations: 10 Description phenacetin Salsolinol Fusaric acid	Adducts M+H-H2 M+H-H2 M+H-H2	Formula C10H13NO2 C10H13NO2 C10H13NO2	Retention ti Score B 47.6 B 41.9 B 41.8 B 41.8 B 41.8 B 41.8	Fragmentation B 43.6 B 15.5 B 14.6	m/z on Mass error (ppm 3.52 3.52 3.52		165	170 175	HN
After searching, identifications wi to the relevant compounds autor Compounds autor Tag filter applied compounds may be hidden mpound Accepted ID 7.11_308.1061n 6.50_270.1040n 6.52_2_308.084n 5.18_151.0634n	matically. or identifications Edit Tag • Edit Egit Edit Edit Edit	Legend: Mate Possible identific: Compound ID CSID4590 CSID4590 CSID4599 CSID4599 CSID4599 CSID4599 CSID4590 CSID4590	the d fragment Unmatched f ations: 10 Description phenacetin Salsolinol Fusaric acid Risocaine	Adducts Adducts M+H-H2 M+H-H2 M+H-H2 M+H-H2	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	Retention t Score B 47.6 G 41.9 G 41.9 G 41.8 G 41.8 G 41.8 G 43.8	Fragmentatic B 43.6 B 15.5 B 14.6 B 14.6	m/z Mass error (ppm 3.52 3.52 3.52 3.52		165	170 175	HN 0
After searching, identifications with to the relevant compounds auton Search for Compounds may be hidden mpound Accepted ID 7.11_308.1061n 6.50, 270.1040n 6.22_308.0834n 5.138_151.0634n 5.138_179.0953n	Edit	Legend: Mate Possible identific: Compound ID CSID4590 CSID4590 CSID4590 CSID4590 CSID4590 CSID4090 CSID491	hed fragment Unmatched f ations: 10 Description phenacetin Satsolinol Fuseric acid Risocaine 4-amino-3-phenylbutyric acid	Adducts M+H-H2 M+H-H2 M+H-H2 M+H-H2 M+H-H2 M+H-H2 M+H-H2 M+H-H2	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	Retention t Score B 47.6 G 41.9 G 41.9 G 41.8 G 41.8 G 41.8 G 43.8	Fragmentation B 43.6 B 15.5 B 14.6 A 0 A 0	m/z Mass error (ppm 3.52 3.52 3.52 3.52 3.52 3.52		0	170 175	HN 0
After searching, identifications with to the relevant compounds auton Search for Compounds may be hidden mpound Accepted ID 7.11_308.1061n 6.22_308.0834n 5.18_151.0634n 5.18_151.0634n 2.96_314.1423n	matically. or identifications Edit Tag • Edit Egit Edit Edit Edit	Legend: ● Mato Possible identification ☆ ☆ Compound ID ★ CSID4590 ☆ CSID4590 ☆ CSID42699 ☆ CSID5060 ☆ CSID5090 ☆ CSID4324 ☆ CSID5090 ☆ CSID13491 ☆ CSID1411	hed fragment Unmatched f ations: 10 Description Satsolinol Fusaric acid Risocaine 4-amino-3-phenylbutyric acid MPMC	Adducts M+H-H2	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	130 135 Retention t Score B 47.6 B 41.8 B 41.8 B 43.8 A 38.8 A 38.8	Fragmentation B 43.6 B 15.5 B 14.6 A 0 A 0	m/z Mass error (ppm 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52		0	170 175	HN
After searching, identifications wi to the relevant compounds autor compounds autor mpound Accepted ID 7.11_08.061n 650_270.1040n 6.52_2.086.0834n 5.18_1510.0954n 5.18_1510.0954n 2.96_314.1422n 2.91_151.0634n	Edit	Legend: ● Matc Possible identification Compound ID ★ CSID4590 ★ CSID4590 ★ CSID62699 ★ CSID324 ★ CSID3491 ★ CSID16401 ★ CSID16066	hed fragment Unmatched f ations: 10 Description phenacetin Salsolinol Futaric acid Risocaine 4-amino-3-phenylbutyric acid MFMC XMC	Adducts M+H+H2 M+H-H2 M+H-H2	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	130 135 Retention t Score 6 47.4 5 41.8 6 41.8 6 43.8 A 38.8 A 38.8 A 38.8 A 38.8	Fragmentatic B 43.6 B 15.5 B 14.6 B 14.6 A 0 A 0 A 0	m/z 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52		0	170 175	HN
After searching, identifications with to the relevant compounds auton Search for compounds may be hidden compound Accepted ID 7.11,308.1061n 6.50,270.1040n 6.50,270.1040n 6.50,270.1040n 5.18,151.0634n 5.18,151.0634n 5.18,179.0933n 2.96,314.1423n 2.91,151.0634n 3.67,368.1012m/z	Edit	Legend: ● Matc Possible identification ○ ☆ CSID4590 ☆ CSID324 ☆ CSID6906 ☆ CSID6906 ☆ CSID6111 ☆ CSID13491 ☆ CSID13491 ☆ CSID1606 ☆ CSID16006539	hed fragment Unmatched f ations: 10 Description phenacetin Satsolinol Fusaric acid Risocaine 4-amino-3-phenylbutyric acid MPMC XMC L-Homophenylalanine	Adducts M+H-H2	Formula C10H13NO2 C10H13NO2	130 135 Retention t Score 6 47.4 5 41.8 6 41.8 6 43.8 A 38.8 A 38.8 A 38.8 A 38.8	Fragmentatic B 43.6 B 15.5 B 14.6 A 0 A 0 A 0 A 0	m/z Mass error (ppm 3.52 3.52 3.52 3.52 3.52 3.52 3.52 3.52		0	170 175	180 185 1
After searching, identifications with to the relevant compounds auton Search for Compounds may be hidden ompound Accepted ID 7.11,308,1061n 6.50,270,1040n 6.622,308,0384n 5.18,151,0634n 5.18,151,0634n 2.96,314,1423n 2.91,151,0634n 3.67,368,1012m/z	ratically. " or identifications Edit Tag V E	Legend: Matc Possible identification Matc compound ID CSID4590 ☆ CSID4590 ☆ CSID324 ☆ CSID324 ☆ CSID345 ☆ CSID344 ☆ CSID13491 ☆ CSID16061 ☆ CSID16061 ☆ CSID200639 ☆ CSID200639	hed fragment Unmatched f ations: 10 Description phenacetin Satisfinal Fusaria acid Risocaine 4-amino-3-phenylbutyric acid MPMC XMC L-Homophenylalanine IPC	Adducts M+H-H2	Formula C10H11NO2 C10H11NO2 C10H11NO2 C10H11NO2 C10H11NO2 C10H11NO2 C10H11NO2 C10H11NO2 C10H11NO2 C10H11NO2	130 135 Retention t Score D 47.6 Q 41.9 Q 41.8 Q 41.8 Q 41.8 Q 43.8 Q 38.8 Q 38.8	Fragmentatic B 43.6 B 15.5 B 14.6 A 0 A 0 A 0 A 0	m/z		0	170 175	180 185 1

Filtering using ChemSpider Results using Elemental composition

To apply a filter based on a specified elemental composition to ChemSpider you must first set up the filter as a new or edited parameter set.

Tip: a simple way to create a new search parameter set is to **Create a copy**, rename and edit an existing set.

With the Default set as current, select **Create a copy** from the Edit menu. In the parameters dialog rename this set (i.e. HCN Filter) and tick **Filter by elemental composition** and untick Perform theoretical fragmentation.

Identify Compounds Select your identification method:	Compound 5.
🔍 ChemSpider 🔹	
About this method + Download others	
 Filter the compounds Using the list below, <u>filter the compounds</u> to show only those you want to identify. Choose search parameters Select your ChemSpider search parameters or create a new parameter set: 	Intensity
Default 🔹 Edit 👻	
3 Search for identifications	
 After searching, identifications will t to the relevant compounds automat Create r 	new
Search for ic Create a	а сору
Tag filter applied Delete	tit

CN Filter		
Required search param	eters	
Precursor tolerance:	5	ppm •
Precursor tolerances a minimum of 0.0001 D		
Data source:	KEGG.	
		Select data sources
Optional search param	eters	
Perform theoretical		i
Fragment tolerance:	5	ppm
Filter by isotope sin	nilarity score	
		%
Isotope similarity:	95	
sotope similarity: Filter by elemental	composition	: 0-100 N: 0-10 O: 0-30
sotope similarity: 7 Filter by elemental	composition	:: 0-100 N: 0-10 O: 0-30 I-2

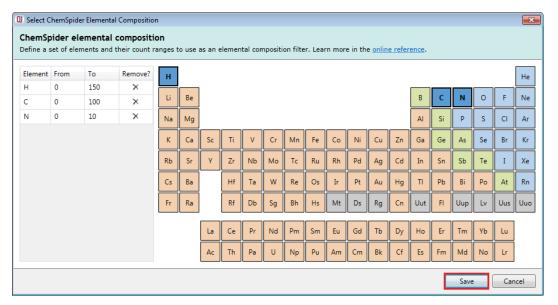
Click **Select elemental composition...**to edit the current Elemental composition.

The existing composition is displayed.

Element	From	То	Remove?	н																	Не
н	0	150	×																		
с	0	100	×	Li	Be											В	C	N	0	F	Ne
N	0	10	×	Na	Mg											AI	Si	Р	s	CI	Ar
0 p	0	30 2	×	к	Ca	Sc	Ti	v	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Р S	0	2	×	Rb	Sr	Y	Zr	Nb	Мо	Тс	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
				Cs	Ba		Hf	Та	w	Re	Os	Ir	Pt	Au	Hg	TI	Pb	Bi	Po	At	Rn
				Fr	Ra		Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	FI	Uup	Lv	Uus	Uu
						La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	
						Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	

Remove elements either by clicking on the x in the table or on the element(s) in the periodic table. Then adjust the ranges (From and To) of atoms for each element in the formula using the table.

Finally **Save** the adjusted composition.



The amended composition appears in the HCN Filter parameter set. Save this to make it current.

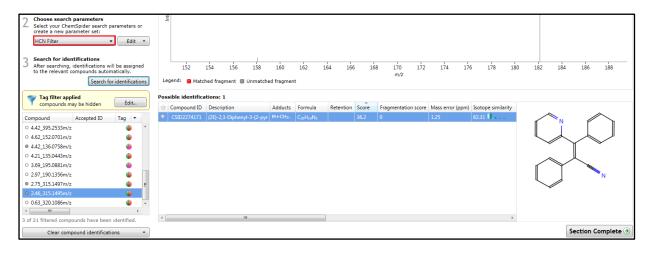
The new set now displays the Elemental Composition Filter

L	Choose search parameters Select your ChemSpider search p create a new parameter set:	ara	ameter	s or
	HCN Filter	•	Edi	t 🔻
3	Default / X Precursor tolerance: Data sources: Theoretical fragmentation: Fragment tolerance: Isotope similarity filter: Elemental composition filter:	Ki or 5 of	ppm ff	ned tions
Con • 7 • 6	HCN Filter Precursor tolerance: Data sources: Theoretical fragmentation: Fragment tolerance: Isotope similarity filter: Elemental composition filter:	or 5 of	ppm ff	

As an example, using the Default search parameters ChemSpider returns 10 possible identities for **Compound 2.46_315.1495**

Choose search parameters Select your ChemSpider search p create a new parameter set: Default	Edit •	Int												
Search for identifications After searching, identifications w to the relevant compounds autor Search fit		Leç		154 156 158 ned fragment 🔳 Unmatche	160 d fragmer	162 16	64 166	5 16	8 170 172 <i>m/z</i>	174 17	6 178 :	180 182	184 18	6 188
🜍 Tag filter applied		Poss	ible identifica	tions: 10										
compounds may be hidden	Edit	☆	Compound ID	Description	Adducts	Formula	Retention	Score	Fragmentation score	Mass error (ppm)	Isotope similarity	1		
Compound Accepted ID	Tag 💌	*	CSID2274171	(2E)-2,3-Diphenyl-3-(2-pyr	M+CH3	C ₂₀ H ₁₄ N ₂		36.2	0	1.25	82.31		2 M	
4.42_395.2533m/z	•	\Rightarrow	CSID10128107	vomilenine	M+H-2	$C_{21}H_{22}N_2O_3$		35.9	0	0.95	80.78 📕 🗕 💷 🗕		ĩ	
4.62_152.0701m/z		☆	CSID26332999	(1R,10S,12R,13R,14S,16S,18	M+H-2	C21H22N2O3		35.9	0	0.95	80.78 📕 🗕 💷		\checkmark	
4.42_136.0758m/z	۷	☆	CSID391595	Cathenamine	M+H-2	C21H22N2O3		35.9	0	0.95	80.78 📕 🗕 💷 🗕	\checkmark	Ý	\sim
4.21_135.0443m/z		☆	CSID391596	19-epi-Cathenamine	M+H-2	C21H22N2O3		35.9	0	0.95	80.78	~		
3.69_195.0881m/z	٠	☆	CSID4445235	Methyl (15alpha,19E)-16-fc	M+H-2	C21H22N2O3		35.9	0	0.95	80.78		\checkmark \land	
2.97_190.1356m/z	•	☆	CSID4573616	Methyl (19E)-16-formylsar;	M+H-2	C21H22N2O3		35.9	0	0.95	80.78			N
2.75_315.1497m/z	🕘 E	\$	CSID2272523	(E)-Ranitidine	M+H	C13H22N4O3S		35.8	0	3.16	82.96		1	
2.46_315.1495m/z	۵.	☆	CSID77849	Aphistar	M+H	C13H22N4O3S		35.8	0	3.16	82.96			
0.63_320.1086m/z	*	☆	CSID10128463	Testosterone, 9-chloro-17-	M+H-2	C20H27CIO3		31.5	0	-4.28	62.32			
of 21 filtered compounds have bee	en identified.	۰ 📃										Þ.		
Clear compound identificat													Sectio	n Complete

Clear the compound identifications and change the search parameters to HCN Filter and click Search for identifications.



Application of the filter to the ChemSpider results reduces the possible identities to one based on restricting the composition to HCN.

Running a LipidBlast Search

As the samples for this tutorial were prepared from Urine, using LipidBlast as your primary method of compound identification is **not** appropriate.

For details on how to setup and use LipidBlast refer to Appendix 6 (page 104)

Running a NIST MS/MS Search

Select NIST MS/MS library from the Identifications Methods.

Note: to use this compound search method you will need to purchase the NIST MS/MS library. Hence using this section is dependent on you having purchased and downloaded the NIST MS/MS library plugin.

The NIST library search plugin bundles the NIST MS/MS libraries and performs a combination of neutral mass and MS/MS based searches. It uses the same search and scoring algorithms as <u>Progenesis</u> <u>MetaScope</u>.

Firstly, delete all your compound identifications using "Clear compound identifications".

Set the required search parameter for the both the Precursor and the Fragment mass error tolerances for matching each compound and its observed fragments, the default tolerance is 5ppm.

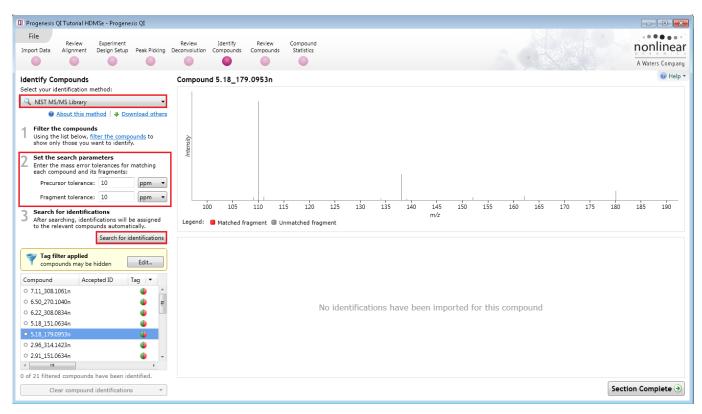
Precursor tolerance

This value, in ppm or Da, is the maximum difference between the mass for the compound in the NIST library and the observed mass.

Fragment tolerance

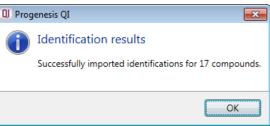
This threshold (in ppm or Da) is the maximum difference between the mass of a fragment in the NIST MS/MS library and the observed mass of the fragment.

For this example, set both tolerances to 10ppm.



Then click Search for identifications.

The progress bar reports the number of masses being searched. When the search completes the number of compounds with imported identifications is displayed.



You can visualise the matching of measured to database-predicted fragments for each potential ID using the 'mirror plot' on the same screen, after clicking on a search result. Matched fragments are shown in red, unmatched in grey; and hovering the cursor over a fragment match will display relevant information in a tooltip.

Q Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment	Review Identify Review Compound	
	Deconvolution Compounds Compounds Statistics	nonlinear
		A Waters Company
Identify Compounds	Compound 5.18_179.0953n	🔞 Help 🔻
Select your identification method:	-	
NIST MS/MS Library	G	
About this method + Download others	Weasure	
Filter the compounds	Peak m/z: 110.0608	
Using the list below, filter the compounds to show only those you want to identify.	Intensity: 100.0601	0%)
Set the search parameters	Δ m/z: 0.0007	
Enter the mass error tolerances for matching each compound and its fragments:	pase	
Precursor tolerance: 10 ppm •	Database	
Fragment tolerance: 10 ppm 🔹		
2 Search for identifications	50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 13 m/z	5 140 145 150 155 160 165 170 175 180 185 190 195
After searching, identifications will be assigned to the relevant compounds automatically.	Legend: 🗧 Matched fragment 🗐 Unmatched fragment	
Search for identifications	Possible identifications: 1	
Tag filter applied	🖈 Compound ID Description Adducts Formula Retention ti Score Fragmenta	
compounds may be hidden	★ 62-44-2 Phenacetin M+H-H ₂ C ₁₀ H ₁₃ NO ₂ 58 95.8	
Compound Accepted ID Tag 💌		Ŭ Î Î
• 7.11_308.1061n		
● 6.50_270.1040n 🕒 ≡		
• 6.22_308.0834n		
 5.18_151.0634n 5.18 179.0953n 		
• 2.96_314.1423n		
• 2.90_914.142311 • • • • • • • • • • • • • • • • • •		
<		
17 of 21 filtered compounds have been identified.	4 <u> </u>	
Clear compound identifications •		Section Complete 🏵

For all the identification methods described in this section additional information and guidance can be obtained **online** in the form of FAQs accessed by the links imbedded in the software.

Q Progene	esis QI Tutorial HI	OMSe - Progene	sis QI		
File	Review	Experiment		Review	v Identify
Import Dat	ta Alignment	Design Setup	Peak Picking	Deconvolu	
•				•	•
				_	
	Compounds ur identification i	method:		Comp	ound 5.18_17
Select you	-	method:	•	Comp	oound 5.18_17
Select you	ur identification i MS/MS Library	method: nethod 🔶 Dor	vnload others		oound 5.18_17

For each identification method the **About this method** link will take you to relevant page(s) for the current method.

Stage 9: Review Compounds

The Review Compounds stage allows you to examine the behaviour of all or subsets of compounds based on tag filters.

Note: having explored the various methods for identifying compounds, before you use **Review Compounds** return to the Identify Compounds stage and click **Clear compound identifications (if not already clear),** then select Progenesis MetaScope as the identification method, select Tutorial Theoretical Fragmentation as the current search parameters and check they are set as shown below using the **Edit** drop down.

Q Progenesis QI Tutorial HDMSe - Progenesis QI			
File Review Experiment Import Data Alignment Design Setup Peak Picking Dec	Review Identify Review convolution Compounds Compounds	QI Edit Search Parameters MetaScope search parameters	nonlinear
		Define a set of MetaScope parameters that can be saved for later	A Waters Company
	Compound 5.18_179.0953n	reuse. Learn more in the <u>online reference</u> .	🛞 Help 🔻
Select your identification method:		Name:	
Progenesis MetaScope		Tutorial Theoretical Fragmentation	
		Compound database	
Filter the compounds Using the list below, filter the compounds to	2	C:\Program Files (x86)\Nonlinear Dynamics\Prc Browse	
show only those you want to identify.	ntensity	Data format: Auto-detect 🔹	
2 Choose search parameters Select your MetaScope search parameters or	5	Search parameters	
create a new parameter set:		Precursor tolerance: 12 ppm •	
Tutorial Theoretical Fragmentat 🔹 Edit 🝷		Retention time within: 0.1 minutes	
2 Search for identifications		CCS within: 2,5 %	
Identifications will be assigned to the relevant compounds automatically.	100 105 110 11		155 160 165 170 175 180 185 190
Search for identifications	Legend: 🧧 Matched fragment 🗏 Un	Additional compound properties source	
Tag filter applied		Read additional compound properties from this file <pre></pre>	
compounds may be hidden Edit		<no aatabase="" selectea=""> Drowse</no>	
Compound Accepted ID Tag 🔻		Fragment search method	
• 7.11_308.1061n		Do not use fragmentation data	
○ 6.50_270.1040n		Perform theoretical fragmentation	
○ 6.22_308.0834n		Fragment tolerance: 12 ppm	this compound
o 5.18_151.0634n		Perform fragment database search	
• 5.18_179.0953n		E:\Customer Data\Progenesis QI v 3.0 Tuorial R Browse	
 2.96_314.1423n 2.91_151.0634n 		Fragment tolerance: 12 ppm *	
0 3.67_368.1012m/z			
< +		Save search parameters Cancel	
0 of 21 filtered compounds have been identified.			
Clear compound identifications			Section Complete $ e e e e e e e e e e e e e $

Click **Search for identifications.** Click **Edit** on the applied tag filter the tags created earlier should still be applied, if not set them up as described in tagging page 42. The display should look as below.

QI Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment	Review Identify Review Compound	
	Deconvolution Compounds Compounds Statistics	nonlinear
		A Waters Company
Identify Compounds	Compound 5.18_179.0953n (Phenacetin)	Q Filter the compounds
Select your identification method:		Create a filter
Progenesis MetaScope		Show or hide compounds based on a selection of their tags. Move tags to the appropriate boxes
About this method Download others		to create the filter. For more guidance, please see the <u>online reference</u> .
Filter the compounds Using the list below, filter the compounds to		Available tags: Show compounds that have all of these tags:
Show only those you want to identify.		Anova p-value ≤ 0.05 (1342 compounds)
Choose search parameters	Inte	Max fold change ≥ 2 (488 compounds)
Select your MetaScope search parameters or create a new parameter set:		Show compounds that have at least one of these tags:
Tutorial Theoretical Fragmentat Edit		
	0	
3 Search for identifications Identifications will be assigned to the relevant	100 105 110 115 120 125 130 135	Hide compounds that have any of these tags:
compounds automatically.	Legend: 🧧 Matched fragment 🔳 Unmatched fragment	Not identified (5273 compounds)
Search for identifications	Legend. Matched fragment Wonnatched fragment	
Tag filter applied	Possible identifications: 4	
compounds may be hidden Edit	Compound ID Description Adducts Formula Retent	t Clear the filter OK Cancel
Compound Accepted ID Tag 💌	49854487 Phenacetin M+H-H ₂ C ₁₀ H ₁₃ NO ₂ 5.16	
• 7.11_308.1061n	☆ HMDB40021 2,3-Dihydro-5-(3-hydroxypro; M+H-H ₂ C ₁₀ H ₁₃ NO ₂	42 15.8
● 6.50_270.1040n	HMDB31811 3,5-Dimethylphenyl methylcar M+H-H2 C10H13NO2	38.8 0
• 6.22_308.0834n	HMDB41931 3,4-Methylenedioxyamphetan M+H-H2 C10H13NO2	38.8 0
• 5.18_151.0634n		ů .
2.96_314.1423n 2.91 151.0634n		
• 3.67 368.1012m/z		
< >		
All 21 filtered compounds have been identified.	III III	÷
Clear compound identifications		Section Complete 🤿

Finally to accept Phenacetin as **Compound 5.18_179.0953n's** identity, click on the 'star' to the left of the Compound ID.



Now move to **Review Compounds** by clicking on the Workflow.

File Review Experiment port Data Alignment Design Setup Peak Picking D	Review Identify Deconvolution Compoun			pound istics								D Y	nlir	(T
	• •	•									100	AV	Vaters Co	
eview Compounds	Find a compound: Sea	arch	,	P	Filter compou	ids 🔻 Filter is	active						۷) He
ing this screen, you can find the compounds of erest in your experiment.	Compound	Neutral mass	m/z	z Rete	ntion time Drift	time Peak Width	Tag 💌	Accepted ID	Identifications	Anova (p)	q Value	Max fold change	Highes	t m
Create a shortlist to review	4.08_235.1803m/z	<unknown></unknown>	235.1803	1 4.08	2.54	0.18	۲		1	< 1.1E-16	< 1.1E-16	Infinity	B_HD	
In the table, sort and filter the compounds based on their measurements, to generate a shortlist	6.22_308.0834n	308.0	309.0907	1 6.22	3.13	0.30			1	< 1.1E-16	< 1.1E-16	1.99E+05	B_HD	
for further review.	• 5.18_179.0953n	179.(A	180.1025				۵	49854487				6.48E+04	B_HD	
Bow are the measurements calculated?	7.13_163.0399m/z	<unknown></unknown>	163.0399	1 7.13	1.62	0.17			1	< 1.1E-16	< 1.1E-16	3.75E+04	B_HD	
To sort the table by a given value, simply click the relevant column header.	4.42_136.0758m/z	<unknown></unknown>	136.0758	1 4.42	1.57	0.14	۲		1	< 1.1E-16	< 1.1E-16	4.82E+03	B_HD	
the relevant column header.	● 7.11_308.1061n	308.1061	309.1134	1 7.11	2.97	0.18	۲		1	< 1.1E-16	< 1.1E-16	2.72E+03	B_HD	
Review the compounds For each compound of interest, inspect the ions'	● 6.50_270.1040n	270.1040	271.1112	1 6.50	2.75	0.14	۲		1	< 1.1E-16	< 1.1E-16	1.53E+03	B_HD	
ali nd peak picking:	• 3.69_195.0881m/z	<unknown></unknown>	195.0881	1 3.69	1.94	0.15	۲		3	< 1.1E-16	< 1.1E-16	1.28E+03	B_HD	
B Review selected compound	• 2.96_314.1423n	314.1423	315.1496	1 2.96	3.02	0.27	۲		1	< 1.1E-16	< 1.1E-16	872	B_HD	
You can also double-click to review a compound.	4													•
Choose the correct identifications	Compound 5.18_179.0	0953n:												
For each compound, select one of its possible	Compound abundance	ce Possible id	lentification	ns 3D M	ontage Drift ti	ne montage								
identifications as the accepted one.			C N	lorm			A	LD				B HD		
To speed this up, you can <i>automatically</i> accept identifications in compounds where only one of the possible identifications has:	10.866667 -		_					-						
	L						-	₩ł.						
Score ≥ 50.0 Accept identifications														
	- pur													
Export data for further processing By exporting your data to external tools, there's	Abundance													
Export data for further processing							-							
Export data for further processing By exporting your data to external tools, there's						C	1							
Export data for further processing By exporting your data to external tools, there's no limit to your analysis.						С]							
Export data for further processing By exporting your data to external tools, there's no limit to your analysis.						c]							
Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	70		Ŕ	×		C]							
Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	ArcSinh Normalised.		Ŕ	M		С]							
Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	ArcSinh Normalised.		×	a		С]	111						•

Window A displays the main table of Compounds with identifications as well as those that remain **unknown** after the identifications have been imported.

Window B Provides tools to set thresholds for the acceptance of identifications based on a score.

Window C displays either: a Compound abundance plot, list of Possible identifications, 3D Montage, Drift Time Montage (for data collected with drift time), for the current compound highlighted in Window A.

Using **Compound 5.18_179.0953n** as an example the 4 views for Window C show:

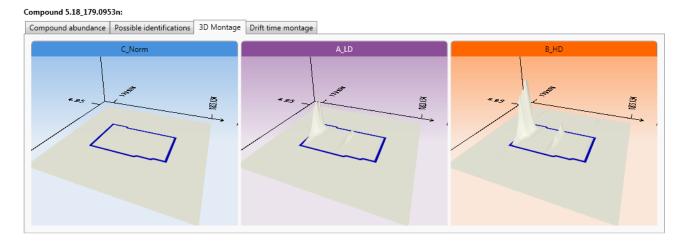
npound abundance	Possible identifications 3D I	Montage Drift time monta	ge		
	C_Norm		A_LD	B_HD	
10.866667 — - - -			101	I ∳I	
- 5.533333 — - -					
0.200000 -	A				

The **Compound Abundance** displaying normalised values for each run, a mean value and 3 standard deviations.

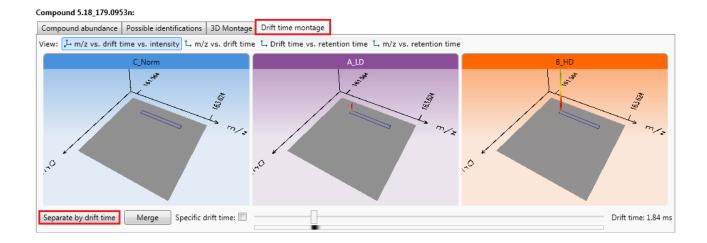
Possible identifications for current compound with Overall Score based on, Mass error, Retention Time (if available), Isotope Similarity based on the comparison of the measured isotope distribution for the compound vs the expected based on the compound formula, Fragmentation Score and CCS if available.

Com	pound abunda	nce Possible identification:	5 3D Mon	tage Drift	time montage				
oss	ible identific	ations: 4							
\$	Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation score	Mass e	
☆	49854487	Phenacetin	M+H-H	C ₁₀ H ₁₃ NO ₂	5.16	57.8	43.6	3.56	NH
☆	HMDB40021	2,3-Dihydro-5-(3-hydroxyp	M+H-H	C ₁₀ H ₁₃ NO ₂		42	15.8	3.56	
☆	HMDB31811	3,5-Dimethylphenyl methyl	M+H-H	C ₁₀ H ₁₃ NO ₂		38.8	0	3.56	
☆	HMDB41931	3,4-Methylenedioxyamphe	M+H-H	C10H13NO2		38.8	0	3.56	

The **3D montage** is based on the first run in each group.



The **Drift time montage** tab is **only** shown for HDMSe data.



Setting a Compound's identity

If you have not already set or accepted a compounds identity at the **Identify Compounds** stage, or wish to amend it as a result of reviewing the data displayed here then you have the same functionality available to you in the **Possible Identifications** tab

For example, to accept Warfarin as the identity for Compound 7.11_308.1061n click on the 'star'.

This updates the compound's 'Accepted ID' and Description.

File Review Experiment mport Data Alignment Design Setup Peak Picking	Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics	đ						ņ	online
											A	Waters Comp
eview Compounds	Find a com	pound: Search		Q	Filte	r compounds 🔻	Filter is active					🕜 He
sing this screen, you can find the compounds of terest in your experiment.	Tag 💌	Accepted ID	Identifications	Anova (p)	q Value	Max fold change	Highest mean	Lowest mean	Isotope distribution	Max Abundance	Min CV%	Description
Create a shortlist to review	۲		1	< 1.1E-16	< 1.1E-16	Infinity	B_HD	C_Norm		8767.2876	2.71	
In the table, sort and filter the compounds based on their measurements, to generate a shortlist	۲		1	< 1.1E-16	< 1.1E-16	1.99E+05	B_HD	A_LD	-	17312.0551	4.73	
for further review.	۲	49854487	4	< 1.1E-16	< 1.1E-16	6.48E+04	B_HD	C_Norm		12216.2532	2.05	Phenacetin
How are the measurements calculated?	۲		1	< 1.1E-16	< 1.1E-16	3.75E+04	B_HD	A_LD		3747.6004	1.22	
To sort the table by a given value, simply click	۲		1	< 1.1E-16	< 1.1E-16	4.82E+03	B_HD	C_Norm		276.6853	2.11	
the relevant column header.	۵	4702	1	< 1.1E-16	< 1.1E-16	2.72E+03	B_HD	A_LD		18770.1608	3.08	Warfarin
Review the compounds	۲		1	< 1.1E-16	< 1.1E-16	1.52E+03	B_HD	C_Norm		1172.3795	3.75	
 For each compound of interest, inspect the ions' alignment and peak picking: 	۵		3	< 1.1E-16	< 1.1E-16	1.28E+03	B_HD	C_Norm		6200.9573	2.86	
Review selected compound	۵		1	< 1.1E-16	< 1.1E-16	872	B_HD	C_Norm		11817.3731	4.46	
You can also double-click to review a compound.			1	< 1.1E-16	< 1.1E-16	713	B_HD	A_LD		127.2337	6.26	
tou can also double-click to review a compound.			4	< 1.1E-16	< 1.1E-16	88.1	B HD	C Norm		1293.4101	1.81	
Choose the correct identifications For each compound, select one of its possible	4								m			•
identifications as the accepted one.	Compound	d 7.11_308.106	iln:									
To speed this up, you can automatically accept identifications in compounds where only one of	Compour	nd abundance	Possible ident	ifications	BD Montage	e Drift time mon	tage					
the possible identifications has:	Possible	identification	ns: 1									
Score ≥ 50.0 Accept identifications	🚖 Com	pound ID Des	cription	Ac	iducts Fo	rmula Retent	tion time Score	Fragmentatio	n score			
	* 470	2 Wa	rfarin	M	FH, C₁	9H16O4 7.12	63.5	47.7		/	\sim	
Export data for further processing By exporting your data to external tools, there's										ſ		
no limit to your analysis.										он 🍬)
											Ϋ́	
Export to pathways tool										~ ~ .	人ノ	
										Υ	\sim	
										\sim	≥_	
									Þ		0	
operiment design												
xperiment design eview your data from a different perspective:										_		Complete

Rather than review all the possible compound identifications you can set a threshold for the score. This allows the automatic acceptance of compound identity where **only** one of the possible identifications is greater than the defined score.

As an example, set the threshold as 48 and then order on Accepted ID and Description.

File Review Alignment	Experiment Design Setup Peak Picking	Review Deconvolutio	Identify n Compound	Review s Compounds	Compou Statistic							~	nonline
			•	•									A Waters Comp
eview Compounds		Find a cor	npound: Sear	ch	2	Fil	ter compounds	 Filter is act 	ive				🕜 He
ing this screen, you car terest in your experime	n find the compounds of ot	ſag ▼	Accented ID	Identifications	Anova (n)	g Value	Max fold change	Highest mean	Lowest mean	Isotope distribution	Max Abundance	Min CV%	Description
Create a shortlist t		a a a a a a a a a a a a a a a a a a a	14708992	1	< 1.1E-16		-	B HD	C Norm	Botope distribution	276.6853	2.11	Acetanilide
In the table, sort and filter the compounds based		7847292	1		< 1.1E-16		B HD	A LD		17312.0551	4.73	Alprazolam	
on their measuremer for further review.	nts, to generate a shortlist	4	46511425	3	< 1.1E-16	< 1.1E-16	1.28E+03	B HD	C Norm		6200.9573	2.86	Caffeine
How are the	e measurements calculated?		855682	1	< 1.1E-16			B HD	C Norm		8767.2876	2.71	Lidocaine
	a given value, simply click		46506142	4		< 1.1E-16		B HD	C Norm		1293.4101	1.81	Paracetamol
the relevant column	header.		49854487	4		< 1.1E-16		B HD	C Norm		12216.2532	2.05	Phenacetin
Review the compo			7847488	1	< 1.1E-16	< 1.1E-16	872	BHD	C Norm		11817.3731	4.46	Ranitidine
For each compound of alignment and peak	of interest, inspect the ions' picking:		9357	1	< 1.1E-16	< 1.1E-16	1.52E+03	B HD	C Norm		1172.3795	3.75	Tolbutamide
Review selected compound		4702	1	< 1.1E-16	< 1.1E-16	2.72E+03	B HD	A LD		18770.1608	3.08	Warfarin	
Marco and a darable	click to review a compound.			1	0.00016	0.00121	2.89	C Norm	B HD		12.6884	26.24	
tou can also double-	click to review a compound.	4		•			***	0.00			407.0202	1.00	
Choose the correct For each compound, identifications as the	select one of its possible		nd 3.69_195.0	881m/z:									r
	u can automatically accept	Compou	Compound abundance Possible identifications 3D Montage Drift time montage										
	pounds where only one of	Possible	e identificati	ons: 3									
the possible identifica	Accept identifications	🚖 Cor	mpound ID D	escription	Adducts	Formula	Retention time	Score Fragm	entation score	Mass error (ppm)		0	
Score > 49.0	Acceptioenuncations	🛨 465	511425 C	offeine	M+H	C8H10N4C	3.69	71.2 74.4		2.46			1
Score ≥ 48.0	Export data for further processing		IDB12488 1,	2,3,4-tetrahydro	M+Na	$C_{11}H_{12}N_2$		37.9 0		-6.62		L	
Export data for fur			IDB37295 3-	[(2-Mercapto-1-	M+H	$C_8H_{18}OS_2$		37 0		4.88	_N		- N
Export data for fur	ata to external tools, there's	× HN											
Export data for fur By exporting your da	ata to external tools, there's ysis.	× HN											
Export data for fur By exporting your da	ata to external tools, there's	× HN											- N
Export data for fur By exporting your da	ata to external tools, there's ysis.	× HN									0	N	^N
Export data for fur By exporting your da	ata to external tools, there's ysis.	× HN									0	N	Ň
Export data for fur By exporting your da	sta to external tools, there's ysis. Export to pathways tool	× HM								Þ	0	N	~ N″

Note: accepting a score threshold causes the Accepted ID and Description fields to be populated when there is a single ID with a score \geq 48 for the compound.

In this example Paracetamol (2.91_151.0634n) was accepted manually. Also the identification of Aspirin (7.12_163.0399) was rejected by clicking on the star icon for possible identifications.

Compound Validation View

Further information about a compound and it's adduct forms, which enables you to confirm the validity of the measurements, is available in the Compound Validation View.

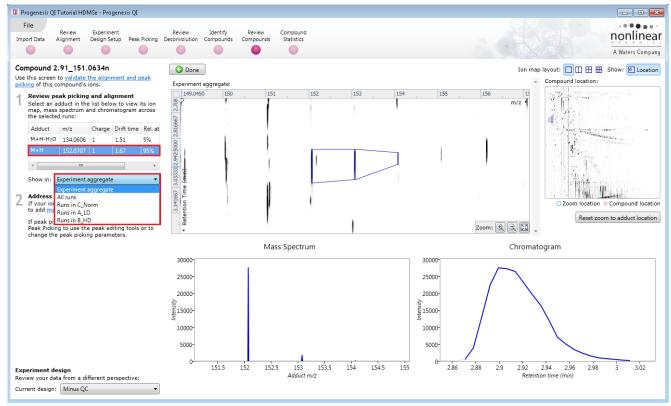
In this view:

- A montage of ion maps helps to validate the ions' peak picking and alignment
- Graphs of mass spectra and chromatograms provide further confirmation of correct peak picking

To open the view either double click on the Compound in the Review Compounds screen or click on Review selected compound

Try double clicking on the compound **2.91_151.0634n** (note: you can accept this compound's identity, as Paracetamol by clicking on the 'star' in the table)

As the **Compound Validation** view opens you can select which adduct of the compound to review first, in this example (M+H).



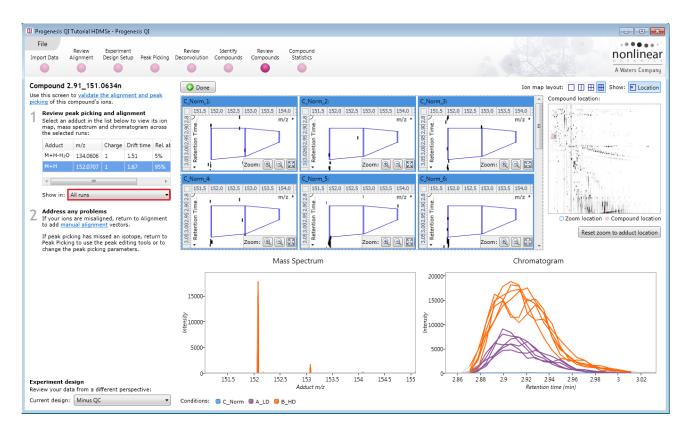
Use the **Show Location** option at the top right of the screen to view the ion's location within the full ion map. **Note**: you can 'toggle' this view on and off and reset the zoom for the adduct location.

Below the list of adducts, a drop-down list lets you select whether you want to see the selected adduct:

- on the experiment aggregate
- across all runs, or
- only on runs within a single experiment condition

The **Experiment aggregate** option is most useful for validating the ions' peak picking, as the aggregate is generated immediately prior to peak picking and is used as the input to that process.

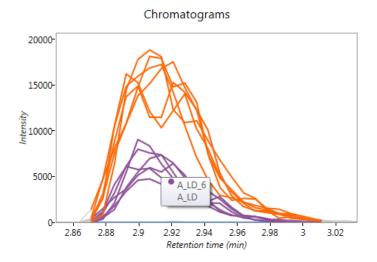
To validate the alignment you will find it best to select to select the **All runs** option and then adjust the ion map layout to display multiple runs



Use the lon map layout and zoom tools to set up the views as shown above. The multi-panels allow you to confirm the alignment across all the runs.

Note: that when the **All runs** are selected the mass spectra and chromatograms are shown for all the runs in the experiment.

In the chromatograms view each line corresponds to a single run, coloured according to the experiment condition to which it belongs. If you hover the cursor over a line, the name of the run which generated it will appear. When showing all runs, these graphs give further validation of how the adduct form's abundance is changing between experimental conditions.



Compounds separated by Drift Time

To view compounds separated by drift time, first clear the existing Filter by clicking on **Filter compounds** and then **Clear the filter.**

Then right click on the Review Compounds Table and create a Quick Tag for those compounds Separated by drift time.

Review Compounds	Find a compound: Se	arch		P		Filter cor	npounds	•						(Ø) H	elp
Using this screen, you can find the compounds of interest in your experiment.	Compound	Neutral mass	m/z	z	Rete	ntion time	Drift time	Peak Width	Tag 💌	Accepted ID	Identifications	Anova (p)	q Value	Max fold ch	ha
Create a shortlist to review	• 4.42_136.0758		Anova p-value ≤ 0.05		4.42		1.57	0.14	۲	14708992	1	< 1.1E-16	< 1.1E-16	4.82E+03	*
In the table, sort and filter the compounds based on their measurements, to generate a shortlist	• 6.22_308.0834				6.22		3.13	0.30	۲	7847292	1	< 1.1E-16	< 1.1E-16	1.99E+05	j
for further review.	• 3.69_195.0881	Max fold cha			3.69		1.94	0.15	۵	46511425	3	< 1.1E-16	< 1.1E-16	1.28E+03	
How are the measurements calculated?	• 4.08_235.1803	Not identifie	d		4.08		2.54	0.18	۲	855682	1	< 1.1E-16	< 1.1E-16	Infinity	1
To sort the table by a given value, simply click	• 2.91_151.0639	New tag			2.91		1.67	0.14		46506142	4	< 1.1E-16	< 1.1E-16	88.1	
the relevant column header.	• 5.18_179.0949	Quick Tags		→		Anova p-	value			49854487	4	< 1.1E-16	< 1.1E-16	6.48E+04	
Review the compounds	• 2.96_314.1423 🖀	Edit tags				Max fold change				7847488 1	< 1.1E-16	< 1.1E-16	872		
For each compound of interest, inspect the ions' alignment and peak picking:	6.50_270.1040	Add to Clip	Gallery			Minimun	n CV			9357	1	< 1.1E-16	< 1.1E-16	1.52E+03	
Review selected compound	• 7.11_308.1061n	308.1061	309.1134	1	1	Not iden				4702	1	< 1.1E-16	< 1.1E-16	2.72E+03	
You can also double-click to review a compound.	0 4.99_316.2118m/z	<unknown></unknown>	316.2118	1		Not fragr					0	0.00772	0.0299	1.33	
	0 4.98_633.3678m/z	<unknown> 633.3678 2</unknown>				l by drift tir				0	0.175	0.249	1.32		
Choose the correct identifications For each compound, select one of its possible	o 5.00_148.0763m/z	<unknown></unknown>	148.0763	1		Identified	and separa	ted by drift ti	ime		0	0.271	0.31	1.5	1
identifications as the accepted one.	•													•	1

Once you have created the tag then filter the table to show only those compounds Separated by Drift Time.

Q Filter the compounds	— ×
Create a filter Show or hide compounds based on a selection o create the filter. For more guidance, please see t	f their tags. Move tags to the appropriate boxes to the <u>online reference</u> .
Available tags:	Show compounds that have all of these tags:
 Not identified (5273 compounds) Anova p-value ≤ 0.05 (1342 compounds) Max fold change ≥ 2 (488 compounds) 	Separated by drift time (158 compounds) Show compounds that have at least one of these tags: Hide compounds that have any of these tags:
Clear the filter	OK Cancel

To do this click on **Filter compounds**... and drag the new Tag on to the **Show** panel and click OK.

The table now shows the only the compounds that are separated by drift time. Click on **Separate by drift time** to view the detected compound ions with the same m/z and RT.

File Review Experiment nport Data Alignment Design Setup Peak Picking De	Review Identify econvolution Compoun	Review ds Compound	Comp ds Stat										nonlir	••
	• •	•											A Waters Co	mpai
eview Compounds	Find a compound: Sea	irch		ø	Filter cor	npounds	▼ Filter is a	ctive					0	Help
sing this screen, you can find the compounds of terest in your experiment.	Compound	Neutral mass	m/z	z	Retention time	Drift time	Peak Width	Tag 💌	Accepted ID	Identifications	Anova (p)	g Value	Max fold change	Hi
Create a shortlist to review	0 3.82_867.1741n	867.1741	885.2079	1	3.82	8.26	0.13			0	0.888	0.549	1.03	A,
In the table, sort and filter the compounds based	o 3.82_885.2068m/z	<unknown></unknown>	885.2068	2	3.82	5.13	0.10			0	0.149	0.226	1.09	В
on their measurements, to generate a shortlist for further review.	0 3.82_901.1805m/z	<unknown></unknown>	901.1805	1	3.82	8.10	0.13			0	0.63	0.471	1.06	B
How are the measurements calculated?	• 3.82_901.1809m/z	<unknown></unknown>	901.1809	2	3.82	5.13	0.11			0	0.106	0.185	1.07	В
To sort the table by a given value, simply click	O 3.82_905.1798m/z	<unknown></unknown>	905.1798	2	3.82	5.13	0.08			0	0.269	0.309	1.25	В
the relevant column header.	0 3.84_455.0922m/z	<unknown></unknown>	455.0922	2	3.84	2.86	0.10			0	0.416	0.384	1.23	B
Review the compounds	0 3.84_594.1359m/z	<unknown></unknown>	594.1359	2	3.84	3.67	0.13			0	0.425	0.388	1.04	c
For each compound of interest, inspect the ions' alignment and peak picking:	0 3.84_614.1083m/z	<unknown></unknown>	614.1083	2	3.84	3.73	0.13			0	0.786	0.519	1.03	В
Review selected compound	0 3.86_284.0927m/z	<unknown></unknown>	284.0927	2	3.86	2.11	0.15			0	0.396	0.375	1.04	В
You can also double-click to review a compound.	4													Þ.
	Compound 3.82_901.3	809m/z												
Choose the correct identifications For each compound, select one of its possible	Compound abundant		and the state		2D Masters D	rift time m	ontana							
identifications as the accepted one.	Compound abundant	e Possible id		ns	50 Montage	inc une m	ontage							_
		n			1.10.11									
To speed this up, you can <i>automatically</i> accept identifications in compounds where only one of the possible identifications has:	View: ↓→ m/z vs. dri	ft time vs. inte C_Norm	ensity L, r	m/z	vs. drift time L	Drift time	e vs. retention A_LD	time 🗅	m/z vs. reten	tion time		B_HD		
identifications in compounds where only one of	View: ↓ m/z vs. dri		ensity L, r	m/z	vs. drift time 1	Drift time		time L	m/z vs. reten	tion time		B_HD		
identifications in compounds where only one of the possible identifications has:	View: J- m/z vs. dri		m/2	m/z	vs. drift time L	Drift time		time L → m/ł	m/z vs. reten	tion time				
identifications in compounds where only one of the possible identifications has: Score ≥ 48.0 Accept identifications Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	View: [J-m/z vs. dri			m/z	vs. drift time L	Drift time			m/z vs. reten	tion time	s J		. m/t	
identifications in compounds where only one of the possible identifications has: Score ≥ 48.0 Accept identifications Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	View: [J-m/2 vs. dri	C.Norm	m/2	ł	vs. drift time L	Drift time			m/z vs. reten	tion time	e de la construction de la const		n/t	13 r

Waters

Exporting compound data

Compound data can be exported in a csv file format. You can either export the **compound measurements** or the **compound identifications.** As an example of data export, first clear any existing filters and then order the table on **Description** and highlight all the compounds which have a description.

File Review Experiment Alignment Design Setup Peak Picking	Review Deconvolution	Identify Compounds	Review Compounds	Compoun Statistics								A Waters Comp
Review Compounds Jsing this screen, you can find the compounds of	Find a com	pound: Search		٩	Filte	er compounds 💌	Filter is activ	e				🔞 He
nterest in your experiment.	Tag 🛛 💌	Accepted ID Id	entifications	Anova (p)	q Value	Max fold change	Highest mean	Lowest mean	Isotope distribution	Max Abundance	Min CV%	Description
Create a shortlist to review	4	14709000 1		<u>- 1 15</u> -16	< 1.1E-16	4.82E+03	B_HD	C_Norm		276.6853	2.11	Acetanilide
In the table, sort and filter the compounds based on their measurements, to generate a shortlist	4		value ≤ 0.05	-16	< 1.1E-16	1.99E+05	B_HD		—	17312.0551	4.73	Alprazolam
for further review.	4	Max fold o	-	-16		1.28E+03	B_HD			6200.9573	2.86	Caffeine
How are the measurements calculated?	4	-	nea by drift time	-16		Infinity	B_HD	C_Norm	 _	8767.2876	2.71	Lidocaine
To sort the table by a given value, simply click the relevant column header.	4	· ·		-16	< 1.1E-16	88.1	B_HD	C_Norm		1293.4101	1.81	Paracetamol
the relevant column header.	4	New tag		-16	< 1.1E-16	6.48E+04	B_HD	C_Norm	 _	12216.2532	2.05	Phenacetin
Review the compounds	4	Quick Tag	s	-16		872	B_HD			11817.3731	4.46	Ranitidine
 For each compound of interest, inspect the ions' alignment and peak picking: 	4	🚰 Edit tags		-16			B_HD					Tolbutamide
Review selected compound	4	Add to Cli	p Gallery	-16			B_HD	A_LD				Warfarin
You can also double-click to review a compound.		1		0.00016	0.00121	2.89	C Norm	B HD		12.6884	26.24	

Then right click on the highlighted compounds and create a New tag... called 'Confirmed Compounds'

Create	new tag			×
	Confirmed Compounds]
		ОК	Cance	el

I Filter the compounds × Create a filter Show or hide compounds based on a selection of their tags. Move tags to the appropriate boxes to create the filter. For more guidance, please see the <u>online reference</u>. Available tags: Show compounds that have all of these tags: Not identified (5273 compounds) Confirmed Compounds (9 compounds) Anova p-value ≤ 0.05 (1342 compounds) Max fold change ≥ 2 (488 compounds) Show compounds that have at least one of these tags: Separated by drift time (158 compounds) Hide compounds that have any of these tags: Clear the filter OK Cancel

Click on Filter Compounds then drag the new tag onto the Show panel of the **Filter compounds** dialog and click **OK**.

Click on the File menu and select **Export compound identifications**, adjust the properties to be included in the export and click OK.

When you have saved the file, a dialog opens allowing you to open the file if required:

Q Export Compound Identifications		
	Export complete	
	Open File Open Folder Close	

Choose properties to be include	ed in exported file
🔽 Compound	
📝 Compound ID	
Accepted?	
Adducts	
📝 Formula	
🔽 Score	
Fragmentation Score	
📝 Mass Error (ppm)	
📝 Isotope Similarity	
📃 Theoretical Isotope Distribu	ution
📃 Link	
Description	
📝 Neutral mass (Da)	
📃 m/z	
📃 Charge	
Retention time (min)	
📝 Anova (p)	
📝 q Value	
📝 Max Fold Change	
🔽 Tags	

Exporting Compound Data to Pathways Tool(s)

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

Currently Progenesis QI supports the export to:

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

Export data for further processing

no limit to your analysis.

By exporting your data to external tools, there's

Export to pathways tool

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

Support for the export of compound identity and expression data to either pathways tool is provided as standard.

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

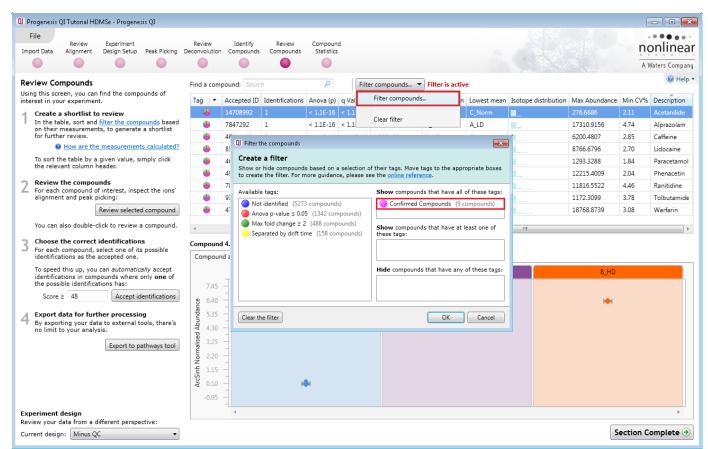
The following section describes an example of the processes involved in exporting to and using **IMPaLA**. For additional details on the export of data to IPA refer to Appendix 8 (page 106)

Using IMPaLA

Typically, only accepted identifications are exported to pathway analysis tools. This is the default and recommended option.

First select the compound data to export to the pathways tool using tag filtering to 'focus' the set to export. For this example you can use the '**Confirmed Compounds'** set described in the previous section.

To ensure that the 9 Confirmed Compounds are the only ones displayed in the table set up the filter as shown below:



Then click Export to pathways tool, then select IMPaLA from the Pathways tools and click Next.

Waters

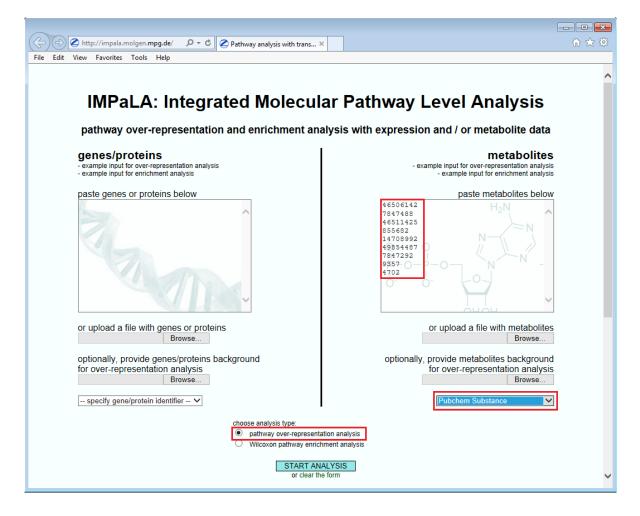
Ensure that Accepted IDs only is selected.

Select either Pathway over-representation analysis or Wilcoxin pathway enrichment test.

Make sure the **Open IMPaLA in my browser** is ticked and then click **Copy compounds to clipboard.**

Export Pathways Information		-2
Configure your export Choose which identifications to	export and the type of analysis you want to perform.	
Select which identifications to e	export:	
Accepted IDs only (recommended)	ended) 🔘 All possible IDs	
Select the type of analysis to p	erform:	
Pathway over-representation an	alysis	
Pathway over-representation an		
Wilcoxon pathway enrichment a	nalysis	
🔽 Open IMPaLA in my browser		
	•	

When IMPaLA opens, paste the metabolites into the right hand panel.



Select specific metabolite identifier

Click Start Analysis

Analysis results list is returned:

11 pathways found. Results per page: 50 v Go to page (previous) 1 of 1 (next)					
pathway name	pathway source	overlapping metabolites	all metabolites	Pmetabolites	Q _{metabolites}
Warfarin Pathway, Pharmacokinetics	PharmGKB	1	1 (1)	0.000759	1
Warfarin Pathway, Pharmacodynamics	PharmGKB	1	1 (1)	0.000759	1
Warfarin Action Pathway	SMPDB	1	5 (5)	0.00379	1
Benzodiazepine Pathway, Pharmacokinetics	PharmGKB	1	11 (11)	0.00832	1
Kenobiotics	Reactome	1	52 (52)	0.0389	1
Cytochrome P450 - arranged by substrate type	Reactome	1	119 (121)	0.0873	1
Bile secretion - Homo sapiens (human)	KEGG	1	162 (162)	0.117	1
Phase I - Functionalization of compounds	Reactome	1	189 (193)	0.136	1
Phase I - Functionalization of compounds	Wikipathways	1	190 (195)	0.137	1
Biological oxidations	Reactome	1	309 (321)	0.215	1
Metabolism	Reactome	1	1281 (1384)	0.672	1

Click to explore returned analysis.

Note: the success of a Pathways analysis is dependent on the organism under study being available to search using **IMPaLA**. This data set is not truly a suitable example but returns results that at least demonstrate the process.

Now move to the Compound Statistics section by clicking section complete.

Stage 10: Compound Statistics

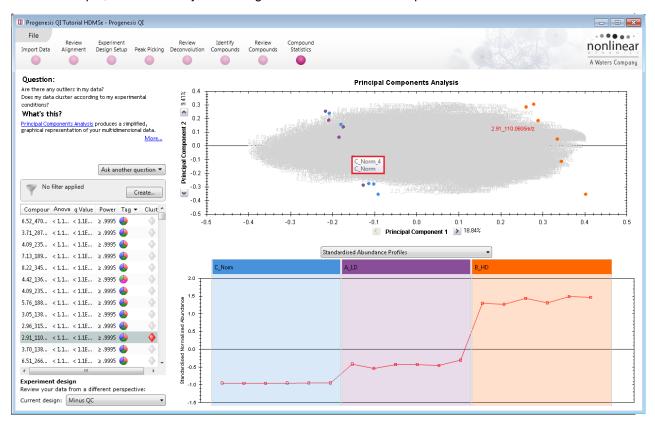
The user guide now describes the functionality of the Multivariate Statistics.

Compound Statistics opens calculating the Principal Components Analysis (PCA) for the active 'tag' filter if one exists.

Note: make sure you have cleared any Tag Filters before performing Principal Components Analysis (PCA)



As an example, we will start by examining the PCA for all 5345 compounds.



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

Note: the 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.

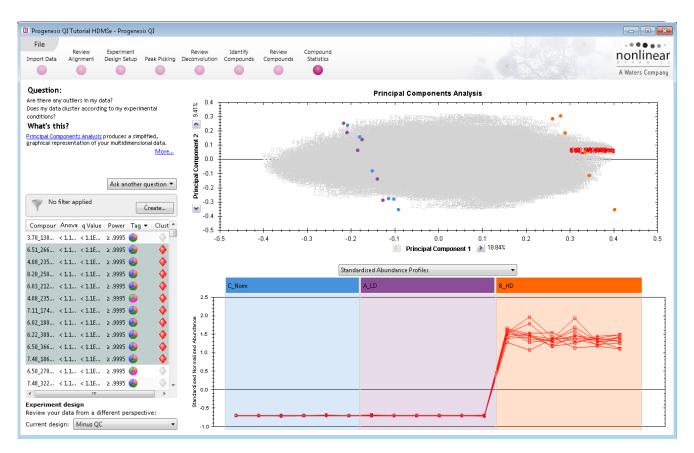
Principal Component Analysis (PCA)

In **Compound Statistics** 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 compounds in the table will highlight the compounds on the 'Biplot' and their abundance profiles will appear in the lower panel.



Note: the table in the Statistics 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 compound, using the abundance 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 8 (page 109)

Correlation Analysis

Use the tags created in Review Compounds to filter the compounds displayed in the table. We are going to explore the Correlation Analysis for all the Compounds (with possible identifications) that display a significant 2-fold or greater difference in abundance.



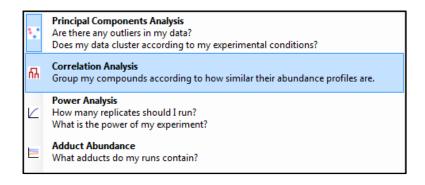
To filter the data click Create

Set up the filter as shown below

QI Filter the compounds Create a filter Show or hide compounds based on a selection of to create the filter. For more guidance, please s	of their tags. Move tags to the appropriate boxes ee the <u>online reference</u> .
Available tags: Not identified (5273 compounds) Anova p-value ≤ 0.05 (1342 compounds) Separated by drift time (158 compounds) Confirmed Compounds (9 compounds)	Show compounds that have all of these tags:
Clear the filter	OK Cancel

On pressing OK the PCA will recalculate.

To set up the **Correlation Analysis** using this filtered data set click on **Ask another question** (above the table). A selection of 4 tools will appear in the form of questions



Select the second option to explore 'feature correlation based on similarity of abundance profiles'



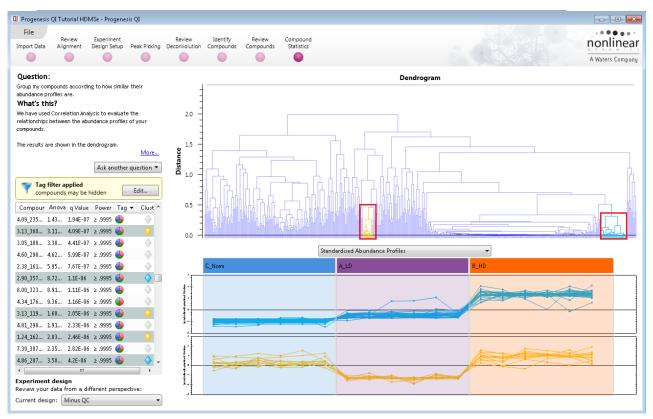
This time the statistically based question(s) being asked is:

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

The question is answered by:

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

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



Correlation Analysis enables the grouping of compounds together according to how similar their abundance profiles are.

Clicking on a branch on the Dendrogram selects the compounds on the table. You can then tag this group, of 'potentially' related compounds, by right clicking and creating a 'New tag' for them. Then use the tag to focus on them at the **Review Compounds** stage.

Exporting to EZinfo

To export data to EZinfo select the option from the file menu.

Note: EZinfo only appears in the Progenesis QI File menu if it has been installed. Therefore, this section is only relevant to users who have access to EZinfo.

	Progenesis QI Tutorial HDMSe - Pi	rogenesis QI
	File	
	Save	nt tup Peak Picking
	Close	
	Export inclusion list	
	Export to EZinfo	imilar their
	Open EZinfo project	-
		aluate the
	Experiment properties	ofiles of your
	Show Clip Gallery	
×	Exit	ı. <u>More</u>

For this example, before exporting to EZinfo, please ensure that you have created an Experimental Design Setup to compare the A_LD and B_HD conditions. To do this, return to Experiment Design Setup set up and create an additional design for 'Low vs High' as shown below.

QI Progenesis QI Tutorial HDMSe - Prog File Review Experime Import Data Alignment Design Se	ent Rev		Compound Statistics	Maximize nonlinear A Waters Company
All conditions Minus	QC Low vs	High I 🗙 🗋 New		Ø Help ▼
Setup conditions	Runs	Add Selected Runs to Condition	▼ Search ♀	
Setup the conditions that you want t (e.g., control, drug A, etc), and then your samples to the correct condition	assign each of	C_Norm_1	C_Norm_2	C_Norm_3
Low	Delete			
	A_LD_1 <u>Remove</u>			
	A_LD_2 <u>Remove</u>			
	A_LD_3 Remove	C_Norm_4	C_Norm_5	C_Norm_6
	A_LD_4 Remove			S. S
	A_LD_6 Remove			139.00
📒 High	Delete			
	B_HD_1 Remove	D_QC_1	D_QC_2	D_QC_3
	B_HD_2 Remove		Sugar,	See.
	B_HD_3 Remove	10 J & 10 J		3,7 37
	B_HD_4 Remove			
	B_HD_5 Remove	D_QC_4		
A LL Pri	B_HD_6 Remove			•
Add condition				Section Complete 🏵

Return to **Compound Statistics,** make sure that there is no Filter applied, then select **Export to EZinfo.** You will be asked to save the project before EZinfo opens.

QI Progenesis QI Tutorial HDMSe - Progenesis QI	
File	
Save tup Peak Picking I	Review Identify Review Compound Compound Statistics nonlinea
Close	A Waters Compan
Export inclusion list	Principal Components Analysis
Export to EZinfo	QI Save As
Open EZinfo project	Co Customer Data > Progenesis QI v2.4 > • 49 Search Progenesis QI v2.4 P
Experiment properties s a simplified,	Organize 🔻 New folder 🔢 👻 🕐
Show Clip Gallery More	Name Date modified 6.22_309.3142m2
× Exit	Libraries Progenesis QI Tutorial HDMSe.Analysis 19/10/2018 11:39
	Documents
Ask another question 🔻	© Protoco
No filter applied Create	Videos
Compour Anova q Value Power Tag - Clust *	The Computer
7.76_486 < 1.1 < 1.1E ≥ .9995 🌑 💧	▲ System (C:) 0.2 0.3 0.4 0.5
6.22_409 < 1.1 < 1.1E ≥ .9995 🥚 💧	DATABASE (E) + 4 III +
6.23_309 < 1.1 < 1.1E ≥ .9995 🌰 🧄	File name Progenesis QI Tutorial HDMSe.usp
7.48_322 < 1.1 < 1.1E ≥ .9995 🧶 🔷	Save as type: UMetrics EZinfo project (".usp)
6.22_309 < 1.1 < 1.1E ≥ .9995 🍊 🔶	
6.22_311 < 1.1 < 1.1E ≥ .9995 🍊 🔶	Hide Folders Save Cancel
7.11_174 < 1.1 < 1.1E ≥ .9995 🌑 🔷	
7.12_370 < 1.1 < 1.1E ≥ .9995 🍊 🔷	
7.11_308 < 1.1 < 1.1E ≥ .9995 🍊 🔷	
6.50_366 < 1.1 < 1.1E ≥ .9995 🤚 🔷	
7.48_186 < 1.1 < 1.1E ≥ .9995 🦀 🔷	
7.48_322 < 1.1 < 1.1E ≥ .9995 🍐 🗸 🔻	9 -0.5
Experiment design	
Review your data from a different perspective: Current design: Low vs High	
current design: Low vs high	-1.5

Save the EZinfo project in the same folder as your Progenesis QI experiment.

A dialog for **Collecting data from EZinfo** opens in Progenesis QI and EZinfo opens displaying the exported data.

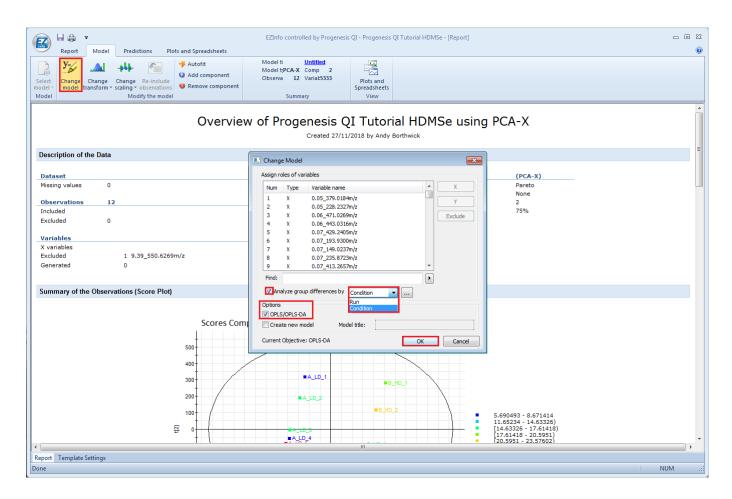
Note: the dialog may be obscured by the EZinfo application as it opens.

Ensure that Pareto Scaling is currently selected, if it is not current, then use **Change Scaling** in the 'Model tab' to set it to Pareto.

As we have specifically exported data to EZinfo for the comparison of 2 conditions we can generate an S-plot from a discriminate analysis (OPLS).

QI Collecting data from EZinfo	- • •
Select the compounds you are interested in on the EZinfo Loadings 'Transfer Loadings data' link to return a batch of compounds to Pro	
The compound batch will appear below where you can tag all the c for identification or further investigation.	ompounds in it ready
Right-click on a batch to assign a tag to them.	
Batch Compounds (count) Tag 💌	
Import	tags Cancel

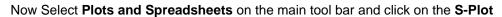
To do this in EZinfo with the exported data, first change the model of the data by selecting Change Model

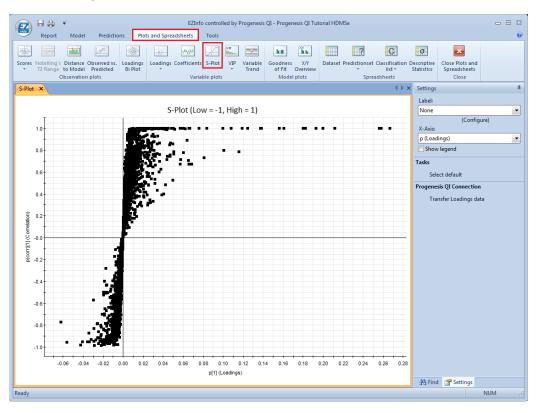


Select analyse group differences by Condition.

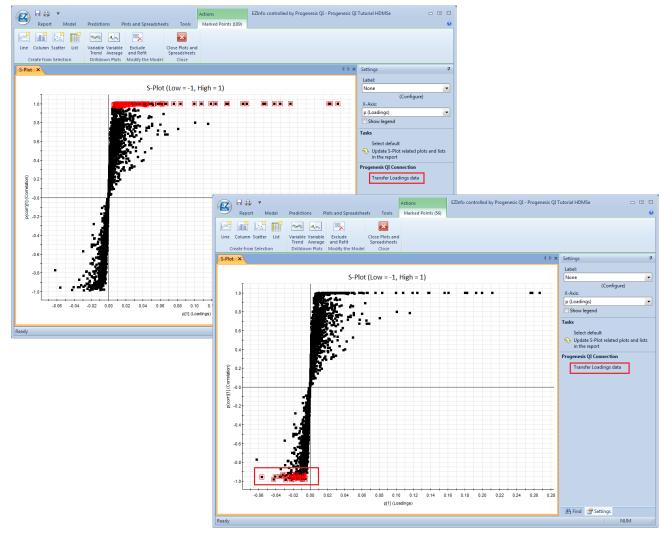
Set the options to OPLS/OPS-DA and click OK

The PCA will update the new model.





Select a set of points and click on Transfer Loadings data. Repeat this process for other sets of data.



As you transfer each group of selected data from the loadings plot a new batch appears in the **Collecting data from EZinfo** dialog.

Select and right click on each batch in turn and create a **New tag...**, naming it appropriately.

Create new tag	×
S_Plot_A	
	OK Cancel

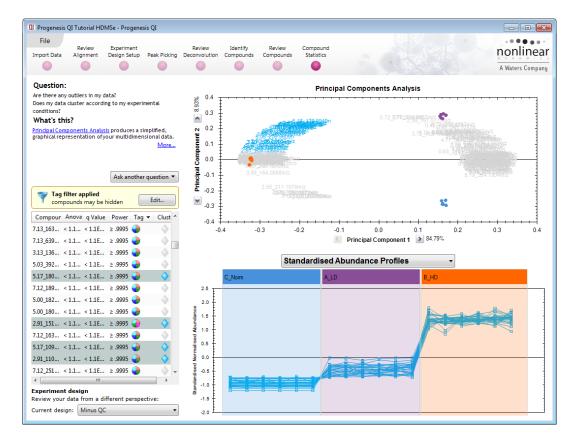
🔲 Collecting data from EZinfo 📃 🗖 💌						
Select the compounds you are interested in on the EZinfo Loadings plot, then click the 'Transfer Loadings data' link to return a batch of compounds to Progenesis QI.						
The compound batch will appear below where you can tag all the compounds in it ready for identification or further investigation.						
Right-click on a batch to assign a tag to them.						
Batch Compounds (count) Tag 💌						
1 194 🔵						
2 72 🥥						
Import tags Cancel						

Click Import tags

Close the dialog and EZinfo. The tags will now be available in to use in Progenesis QI.

	Filter the compounds Create a filter Show or hide compounds based on a selection of their tags. Move tags to the appropriate box to create the filter. For more guidance, please see the online reference.			
new Tag Filter in Compound Statistics to Show bounds that have at least one of these tags.	Available tags: ● Not identified (5273 compounds) ● Anova p-value ≤ 0.05 (1342 compounds) ● Max fold change ≥ 2 (488 compounds) ● Separated by drift time (158 compounds) ● Confirmed Compounds (9 compounds) ● Clear the filter	Show compounds that have all of these tags: Show compounds that have at least one of these tags: Show compounds that have at least one of these tags: Splot_A (194 compounds) Splot_B (72 compounds) Hide compounds that have any of these tags: OK Cancel		

Click OK to create the new PCA based on these tags.



Note: as you close Progenesis QI, EZinfo will also close.

Appendix 1a: Stage 1 Data Import and QC review of data set

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

To create a new experiment with your files select **New** give your experiment a name and then select a location to store the experiment files. Click **Next**.

Q Progenesis QI			- • ×
File			
Experiments			nonlinear
•			A Waters Company
Recent experiments		Search	Getting started G
Th	Create a new ESI experiment Create a new ESI experiment Analyse samples run using electrospray injection (ESI) to generate the charge on compounds. Enter a name for the new experiment: Progenesis QL Tutorial HDMSe Select the location to store the experiment files: Save experiment in the same folder as the run data Choose an experiment folder Esck Next Cancel	nts	EZinfo version information Version 3.0.3.0 of EZinfo is installed. UEP reader plugin update A new version of the UNIFI Export Packages reader plugin is available, allowing Progenesis Q1 to read runs with a single IMS channel from .UEP files. You can <u>download the plugin here</u> . Quickly go to an ion map location Wath to quickly validate your sample running by zooming to a known ion? <u>Go Ta Location</u>
Other experiments		•	Uump to a specific m/2 and RT using the Gor to Location to bit in the top-left corner of the ion maps. Latest blog posts What a ConFirencel Explore. Dream, Discover, When is a Biomarker not a biomarker? (Data 1) When is a Biomarker not a biomarker? A struct 1)

Now select the machine type that was used to run the samples.

QI Create New Experiment Set the analysis parameters The type of machine, and the settings used, determine analysed.	ie how your runs must be
What type of machine was used to run the samples?	
High resolution mass spectrometer	
Select the runs' data format:	High resolution mass spectrometer e.g. Waters SYNAPT G2/G2-S, AB SCIEX TripleTOF, Agilent QTOF, Bruker Maxis, Thermo LTQ Orbitrap Thermo LTQ Iontrap in Enhanced mode.
Resolution (full width at half maximum): 50000	Low resolution ion trap e.g. Bruker HCT, Bruker HCT Ultra, Thermo LTQ XL
Select the ionisation polarity for this analysis:	Thermo FT-ICR
Operative	
Negative	Direct sample analysis Pre-processed data from Progenesis Bridge
Back	Next Cancel

Note: if you have converted or captured the data as centroided then select Centroided data and enter the Resolution (full width at half maximum) for the MS machine used.

In this example the data was captured in **Profile** mode.

QI Create New Experiment	×
Set the analysis parameters The type of machine, and the settings used, determine how your runs must be analysed.	
What type of machine was used to run the samples?	
High resolution mass spectrometer 🔹	
Select the runs' data format: Orofile data Centroided data Resolution (full width at half maximum):	
Select the ionisation polarity for this analysis: Positive Negative 	
Back Next Canc	el

Finally select the ionisation polarity, either **positive** or **negative**, this will determine which list of possible adducts is available on the next page.

0 Create New Experiment	Q Create New Experiment
Select the possible adducts To correctly identify the compounds in your samples, we need to know how the ions are formed.	Select the possible adducts To correctly identify the compounds in your samples, we need to know how the ions are formed.
Adducts in library: Adducts in this experiment: M+H-2H2O >> M+H-H2O < M+H M+H4 M+H4 M+CH3OH+H M+CH3OH+H M+CH3OH+H M+CH3OH+H M+CH3OH+H M+CH3OH+H M+CH3OH+H M+CH3OH+H Edit adduct library	Adducts in library: M+Cl M+K-2H M+FA-H M-FA-H M-2H M-H M-H E Move adducts into this list to add them to your experiment. Edit adduct library
Back Create experiment Cancel	<u>Back</u> <u>Create experiment</u> Cancel

To add an adduct to the library select **Edit adduct library...** and then click **Add**. Now provide the Mass (monomer, dimer or trimer), Adduct, Adduct Mass and Charge of the new adduct and then add it to the library.

Q Adduct Library				- • ×	Q Create Ne	w Adduct		×
The adducts listed t experiments. Show: All Po 				vhen creating new	Enter the na Masses:	ame and det Adduct:	ails of the new ad Adduct Mass:	duct. Charge:
Adduct	Adduct Mass	Masses	Charge	<u>A</u> dd	2M -	+K	38.9632	1 日
🔁 М+Н	1.0073	1	1	Remove				
🚼 M+NH₄	18.0338	1	1			ſ		Constal
🚼 M+Na	22.9892	1	1	E		l	Add	Cancel
🚼 M+CH3OH+H	33.0335	1	1					
🔂 М+К	38.9632	1	1					
H+ACN+H	42.0338	1	1					
🔁 M+2Na-H	44.9712	1	1					
.M.J.A.CALLINS	CA 04.50							
			<u>S</u> ave	Cancel				

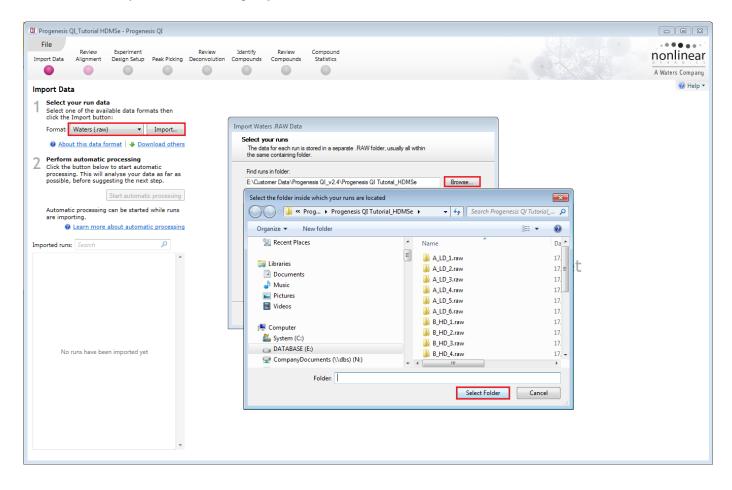
Now select the expected adducts based on your knowledge of your experimental conditions. In this example the polarity was positive and a basic set of Adducts was selected as shown on the right panel below.

dducts in library:			Adducts in this experime	nt:
🗄 M+2Na-H		>>	M+H-2H ₂ O	-
🗄 M+ACN+Na			H+H-H ₂ O	
🛨 M+2Na		<<	🔁 M+H	
🛨 М+ЗН	=		H+NH4	
🗄 M+2H+Na			🔁 M+Na	
🗄 M+2Na+H			M+CH3OH+H	
🛨 2M+H			М+К	
🛨 2M+NH4			M+ACN+H	
🗄 2M+Na			M+2H	
2M+3H ₂ O+2H	~		M+H+Na	

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

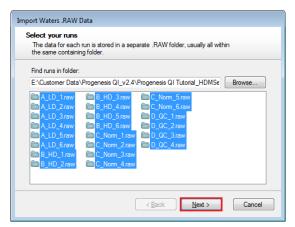
Select the 'Import Data file format', in this example they are Waters (.raw) files

Then locate your data files using Import...



Use Browse to locate and select the folder containing all the Data files, in this example (A_LD_1 to D_QC_4).

Select the required **.raw** folders for your experiment and click **Next.**



Import Waters .RAW Data
Lock mass calibration All of your runs contain lock mass calibration information. Please provide the calibration m/z.
V Perform lock mass calibration
Lock mass m/z: 556.2766
Note: If you have already calibrated your data externally, you should not perform calibration here.
< Back Next > Cancel

Import Waters .RAW Data						
Ready to import						
Please review the information below before starting the import process.						
Your runs are ready to be imported. Please review the options below.						
22 runs selected for import.						
Lock mass calibration: Yes Lock mass m/z: 556.2766						
< Back Import Cancel						

The folders will be examined and the Dialog will ask you for Lock Mass Calibration details.

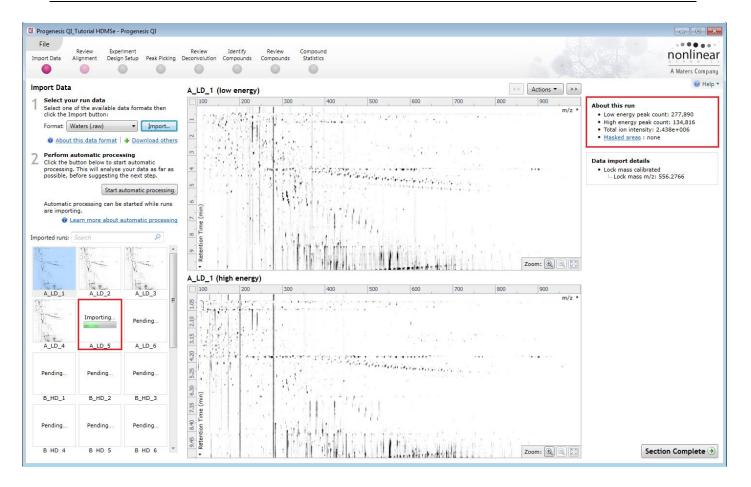
If this has already been performed externally then untick this option.

Note: for the Tutorial data Lock mass calibration is required.

Finally a summary page appears. Click Import

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



Each data file appears as a 2D representation of the run. At this stage, assuming you have selected 'profile data' (page 79) you will be warned if any of the data files have been 'centroided' during the data acquisition and conversion process.

Note: as each data file is loaded the progress is reported in the **Import Data** list. The dialog below the run reports on the QC of the imported Data files. In this case 'No problems found' with this data file.

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

Note: as the loading process starts you can also **set up and start the Automatic Processing** before the loading has completed. This is a 4 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 followed by setting up of an Experiment Design and Peak Picking.

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

Appendix 1b: Stage 1 Data Import from UNIFI (Waters Only)

LCMS data that has been acquired fron Waters Vion and Xevo machines where the acquisition is managed through the UNIFI Scientific Information System, is exported **Uncompressed** from UNIFI as a Unifi Export Package (UEP) format file.

To create a new experiment in Progenesis QI, using a .UEP file, select **New** give your experiment a name (Progenesis UEP example) and then select a location to store the experiment files. Click **Next**.

🛛 Progenesis QI				
File				nonlinea
Experiments				D V N A M T C
•				A Waters Compan
Recent experiments		Search	Q	Getting started Here are some resources to help you get started with Progenesis QI: • Fundamental concepts of the Progenesis QI workflow
OI Creat	e New Experiment			User guide and tutorial data
Create	a new ESI experiment e samples run using electrospray injection (ESI) to generate th			Frequently-asked questions EZinfo version information
Enter a	name for the new experiment:			Version 3.0.3.0 of EZinfo is installed.
	nesis UEP example			
Select t	the location to store the experiment files:			Quickly go to an ion map location Want to quickly validate your sample
	we experiment in the same folder as the run data			running by zooming to a known ion?
There 💿	noose an experiment folder	Browse	5	500
	Back	Cancel		<u>80</u>
	Dack	Cancer		Jump to a specific m/z and RT using the <u>Go To Location tool</u> in the top-left corner of the ion maps.
				Latest blog posts
thar avpariments				<u>6 ways Progenesis QI can help with</u> your compound identification
ther experiments				Season's Greetings from all at
		2.2.5	826.42898	Nonlinear! Barking up the right tree:

Now select the machine type, polarity that was used to run the samples. Then select the expected adducts based on you knowledge of your experimental conditions.

QI Create New Experiment	Q Create New Experiment	
Set the analysis parameters The type of machine, and the settings used, determine how your runs must be analysed.	Select the possible adducts To correctly identify the compounds in your samples are formed.	s, we need to know how the ions
What type of machine was used to run the samples? High resolution mass spectrometer Select the runs' data format: Profile data Centroided data Resolution (full width at half maximum): 18000 Select the ionisation polarity for this analysis: Positive Negative	Adducts in library: M+2Na-H M+ACN+Na M+2Na M+3H M+3H M+2H+Na M+2H+NA M+	Adducts in this experiment: M+H-2H ₂ O M+H-H ₂ O M+H M+NH ₄ M+NH ₈ M+CH ₃ OH+H M+CH ₃ OH+H M+K M+2H M+2H M+H+Na X
<u>B</u> ack Cancel	Back	Create experiment Cancel

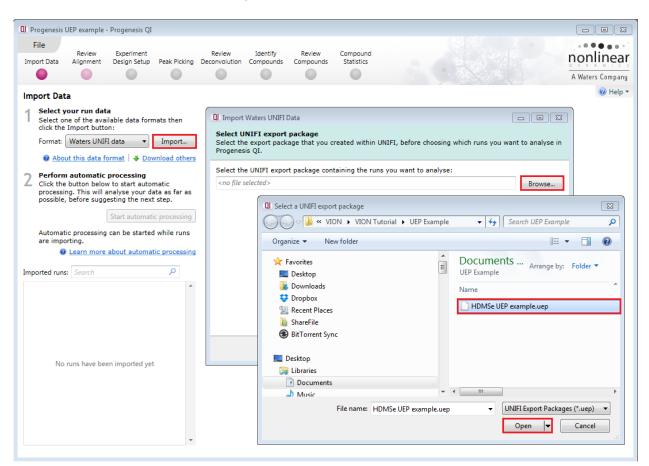
Click Create experiment to generate the experiment.

When Import data opens select **Waters UNIFI data** as the import data format.

Then click Import.

Import Data									
1	Select your run data Select one of the available data formats then click the Import button:								
	Format:	Waters UNIFI data 🔹 👻	Import						
	🔞 Abo	Waters UNIFI data Version: 1.0.6744.42923	vnload others						
2	Perforn Click the processi possible	Version: 1.0.6901.37225 SCIEX (.wiff) Version: 1.0.6680.30256	matic ata as far as t step.						
	Automat	mzXML files	ic processing						
		Thermo (.raw) Version: 1.0.6680.30349	tic processing						
Imp	orted runs	Thermo FT-ICR (.raw) Version: 1.0.6680.30349							

Browse to the location of the UEP and open the file



Depending on the data in UNIFI the exported package may contain 1 or more sample sets. In this example there is only one, containing 19 runs.

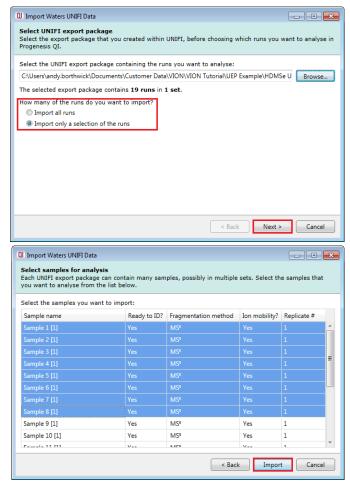
You have the option of importing all the samples or a selection.

In this example we will import a selection of the runs.

Click Next.

Select the runs to import.

Click Import.



As importing from the UEP starts you can set up the **automatic processing** to run on the data as described in Stage 2A of this guide (page 8).

QI Progenesis UEP example - Progenesis QI		
File Review Experiment	Review Identify Review Compound	
	Review Identify Review Compound Deconvolution Compounds Compounds Statistics	nonlinear
		A Waters Company
Import Data	Sample 1 [1] (low energy)	Ø Help ▼
1 Select your run data 2 Select one of the available data formats then click the Import button: Format: Waters UNIFI data ▼ Import Import Import Import	□ 100 200 300 400 500 500 700 800 900 m/z •	About this run • Low energy peak count: 347,874 • High energy peak count: 246,544 • Total ion intensity: 2.372+4005 ked areas : none
2 Perform automatic processing Click the button below to start automatic processing. This will analyse your data as far as possible, before suggesting the next step.	Select an alignment reference To compensate for drifts in retention time, all runs in the experiment must be alig to a single reference run.	nport details oss section area calibration exists.
Start automatic processing Automatic processing can be started while runs are importing.	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: Sample 1 [1]	
Sample 4 [1] Sample 5 [1] Sample 6 [1]	For information on choosing the alignment reference, and why you might want to select your own candidates, please see the <u>online guidance</u> .	
Sample 7 [1] Sample 8 [1] Sample 8 [1]		Section Complete 🤿

Once the importing is complete, the Automatic processing of the data proceeds.

Progenesis UEP example - Progenesis QI		
File Review Experiment	Review Identify Review Compound	
	Deconvolution Compounds Compounds Statistics	nonlinear
		A Waters Company
Import Data	Sample 1 [1] (low energy)	🔞 Help 🔻
1 Select your run data 2 Select one of the available data formats then click the Import button: Format Waters UNIFI data ▼ Import Import Import About this data format Import Import Click the button below to start automatic processing. This will analyze your data as far as far and the selection.	m/z •	About this run • Low energy peak count: 347,874 • High energy peak count: 246,544 • Total ion intensity: 2.327e+005 • Masked areas : none
processing. This will analyse your data as far as possible, before suggesting the next step.	Automatic processing	run is being used as the experiment's nment reference.
C Restart automatic processing	Current step: Analysing	ou want to use a different run as the nment reference, you'll need to discard
Automatic processing can be started while runs are importing. Learn more about automatic processing Imported runs: Search P	Importing runs: 8 of 8 processed Selecting reference: Sample 1 [1] Aligning runs: 7 of 7 processed Creating design: Created Peak picking: Picking Cancel	analysis and restart the automatic ressing: Restart automatic processing a import details Cross section area calibration exists.
Sample 1 [1] Sample 2 [1] Sample 3 [1] Sample 4 [1] Sample 5 [1] Sample 6 [1] Sample 7 [1] Sample 8 [1] -	A Retation Time (min) 10 8 6 4	Section Complete Э

Once the analysis completes refer to page 12 in the guide.



Appendix 2: Stage 2 Processing warnings and failures

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

For example, a run that fails to automatically align will trigger a warning, although analysis will continue; however, the automatic processing dialog will warn you that runs failed to align. On completion of the Automatic processing go to the **Review Alignment** stage to investigate the problem.

Time taken: 9 minutes 58 see	conds	
✓ Importing runs: 22	2 of 22 processed	
 Selecting reference: D 	_QC_3	
Aligning runs: 21	1 of 21 processed	
<u>A</u>	1 run failed to align - continuing without it	
 Creating design: Cr 	reated	
✓ Peak picking: 60	069 compound ions found	

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



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

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

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

Q Processing Complete			×
Automatic process Time taken: 10 minutes 2	ing complete (with 21 seconds	warnings).	
A Importing runs:	22 of 22 processed 1 failed to import		
 Selecting reference: 	D_QC_3		
 Aligning runs: 	20 of 20 processed		
 Creating design: 	Created		
 Peak picking: 	6069 compound ions fo	ound	
			Close

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

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

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

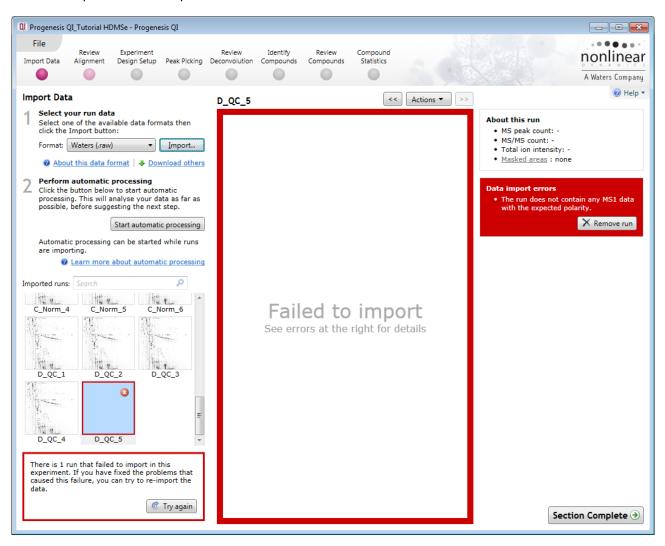
Q Processing Complete		×
Automatic process Time taken: 28 seconds	ing failed.	
A Importing runs:	19 of 19 processed 1 failed to import	
X Selecting reference:	All reference candidates failed to import	
Aligning runs:	Unable to start.	
Creating design:	Unable to start.	
Peak picking:	Unable to start.	
		Close

In this case, the processing dialog has halted indicating that all of the alignment reference candidates failed to import. To proceed restart the process with other runs.

Finally if you have chosen to select the Alignment Reference from a number of QC runs and one of them fails to import then processing will continue as there are alternative runs that did load that can be automatically selected as the reference.

Automatic process	ing complete (with warnings). 43 seconds	
A Importing runs:	23 of 23 processed	
	A 1 failed to import	
A Selecting reference:	D_QC_4	
	🔺 1 candidate failed to import	
 Aligning runs: 	21 of 21 processed	
Creating design:	Created	
Peak picking:	6137 compound ions found	

An example of a failed import is shown below



In this example the run has the wrong polarity compared with the other runs. You can right click on the run and remove it or just click on Remove run to the right of the main display.

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

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.

Appendix 3: Licensing runs (Stage 3)

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



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

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

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

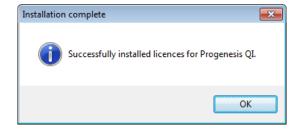
The runs in your experiment will be listed as shown below.

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

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

File Review Experiment nport Data Licensing Alignment Design Setup	Review Identify Review Compound Peak Picking Deconvolution Compounds Compounds Statistics		nonline
Nongle License Runs			A Waters Comp
This installation is currently restricted to analyse icensed runs only.		Licence state	License this run
To license your runs, you need an evaluation or	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	
ease licence code which can be obtained from	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
a sales representative.	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
Once licensed, your runs can be analysed on	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	v
any installation of the software. The licence is	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
automatically included when archiving an experiment.	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	v
	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	
f your runs have been licensed on another computer, click here to make the licences	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	
vailable on this computer.	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
f you have one, you can <u>open a licence file</u> to	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	v
nstall.	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
f you have just installed a dongle, click here.	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
, <u>-</u>	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	V
	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2.2_Demo Suites_Tutorial\Progenesis U	Inlicensed	
	Run licence code: x00:x000x1000x1000x	Use Lice	ence Code

A message confirming successful installation of your licences will appear.



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

Appendix 4: Manual assistance of Alignment

Approach to alignment (using tutorial data)

To place manual alignment vectors on a run (A_LD_1 in this example):

Note: this data is already well aligned so a single vector has already been placed to simulate misalignment so as to allow the demonstration of the use of manual vectors.



- 1. Click on Run A_LD_1 in the **Runs** panel, this will be highlighted in green and the reference run (D_QC_4) will be highlighted in magenta.
- 2. You will need approximately 5 alignment vectors evenly distributed from top to bottom of the whole run.
- 3. First drag out an area on the **Ion Intensity Map** (C), this will reset the other 3 windows to display the same 'zoomed' area.

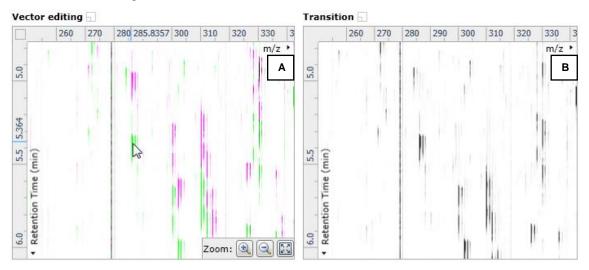
Note: when you click anywhere in the Ion Intensity Map this will **reposition the focus** and update the other views accordingly.

Note: the features moving back and forwards between the 2 runs in the **Transition** window (B) indicating the misalignment of the two runs

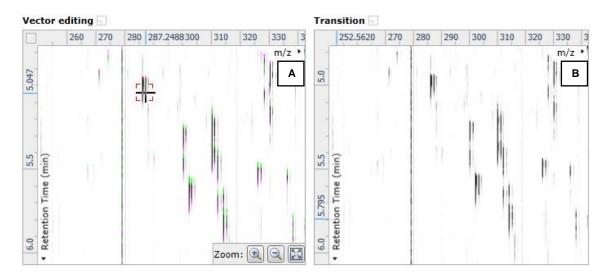
Note: the Ion Intensity Map gives you a colour metric, visually scoring of the current alignment and an overall score is placed next to the Vectors column in the table. With each additional vector this score will update to reflect the overall quality of the alignment. The colour coding on the Ion intensity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

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

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

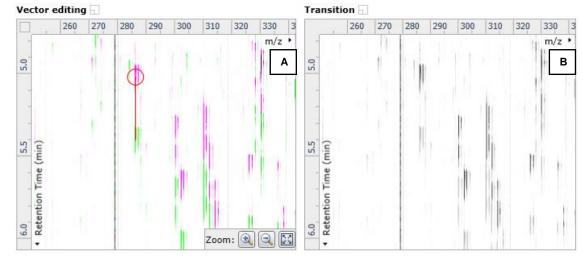


5. As you are holding down the left mouse button (depending on the severity of the misalignment), the alignment vector will automatically find the correct lock. If not, drag the green feature over the corresponding magenta feature of the reference run. The red box will appear as shown below indicating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



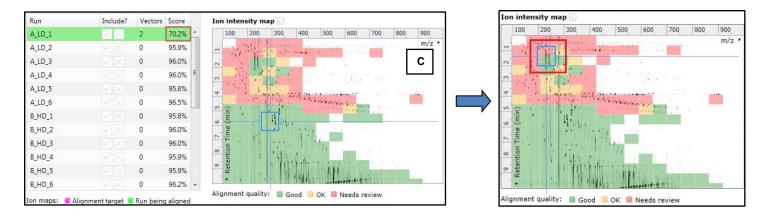
Tip: while holding down the mouse button hold down the **Alt** button. This will allow smooth movement of the cursor as the **Alt** key allows you to override the 'automatic alignment' performed as you depress the mouse button.

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



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

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



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



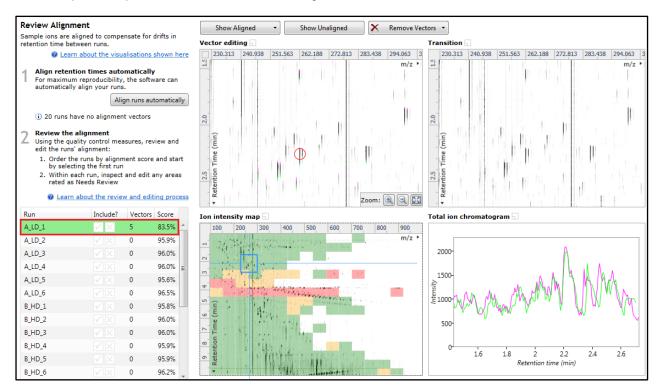
9. Repeat this process moving the focus from top to bottom on the **Ion Intensity Map** view the number you add is dependent on the length of gradient used and the severity of the misalignment. In many cases a single vector per alignment is all that is required to improve the performance of the Automatic Alignment

Note: the manual vectors are red to distinguish them from the automatic vectors which blue

Note: the number of vectors you add is recorded in the Runs table

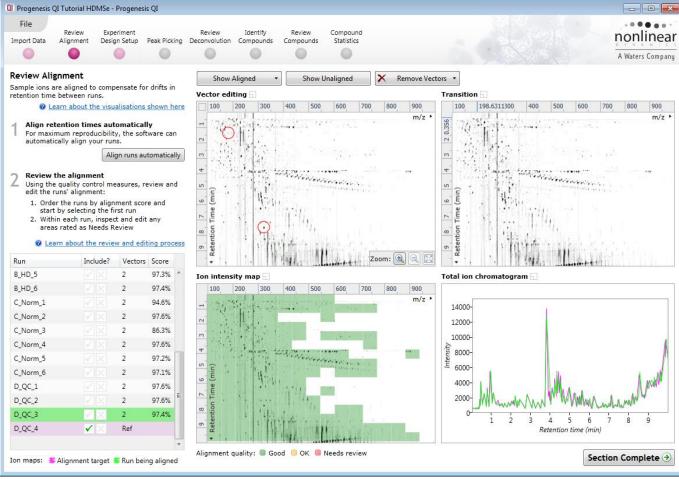
The number of manual vectors that you add at this stage is dependent on the misalignment between the current run and the Reference run. The effect of adding the manual vectors can be seen when you press show aligned

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



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 runs being aligned.



Waters

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.

A	Automatic Alig	Inment		×
	Choose which	runs to automatically align:		
	Run	Notes	Vectors	•
	✓ A_LD_1	run has user vectors		2
	A_LD_2	run has user vectors		2 😑
	A_LD_3	run has user vectors		2
	🔽 A_LD_4	run has user vectors		2
	A_LD_5	run has user vectors		2
	🛛 A_LD_6	run has user vectors		2
	B_HD_1	run has user vectors		2
	B_HD_2	run has user vectors		2
	B_HD_3	run has user vectors		2 🛫
			ОК Салов	

Note: the tick box next to the Run name controls whether automatic vectors will be generated for each run.

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

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 Treatment: Day1 and then at 3 times following treatment: Day2, Day3 and Day4.

🔲 Progenesis QI Tutorial HDMSe - Progenesis QI												8
File Review Experiment Import Data Alignment Design Setup Peak P	icking Dec	Review convolution	Identify Compounds	Review Compounds	Compound Statistics					D	A Waters Co	I C S
Conditions Minus QC	New										0	Help 🔻
Which experiment design type do y	vou wan	t to use	for this ex	xperiment?								
Between-subject Design Do samples from a given subject appear in only one condition? Then	A			Delete	0-0 Have you	ithin-subject Des taken samples from a nder different conditi	given		Before	During	After	
use the between-subject design.	Í	QI Create I	New Experime		Jubjeeru	×	Diam.		Defore	During	Antei	
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			ame for the e atment of pat	experiment des ients	ign:		m to	Patient X	X1	X2	X3	
conditions are independent and therefore gives a statistical test of whether the means of the conditions are all equal.	с	Grou	ou want to gr up the runs ma v an existing d	·			nis nas	Patient Y	¥1	Y2	Y3	
	Add con	Cop	y an existing o	iesign:	Create	e design Cancel	but The	Patient Z	Z1	Z2	Z3	
					measures		peated					
					because t assumptic repeated difference reduced a condition	d ANOVA is not approp he data violates the Ah n of independence. Wi measures ANOVA indiv es can be eliminated or s a source of between differences (which hel nore powerful test).	NOVA ith a idual r					
					thought o paired-sa	n-subject design can be f as an extension of th mples t-test to include n between more than measures.	e					

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.

	Deconvolution Cor	patients I ×	nds Statistics		d		A Waters Company Waters Company
Setup conditions and subjects Setup the conditions and subjects for your experiment design on the right, and then assign each of your samples to the correct subject/condition cell n the grid.		Day 1	Day 2	Day 3	Day 4	Add Condition	
 Add a column for each condition. Add a row for each subject. Drag each of your samples to the correct location in the grid. Filter samples: P 	Patient 1	A_LD_1	A_LD_2	A_LD_3	A_LD_4		
C_Norm_2	Patient 2	B_HD_1	B_HD_2	B_HD_3	B_HD_4		
C_Norm_3	Patient 3	C_Norm_1	sere Sample	Select Sample	Select Sample		
C_Norm_5	Add Subject		<u>.</u>			_	
							Section Complete 🤿

You can create additional Experimental Designs using the New tab

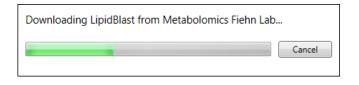
All of these Experimental Designs are available at all the following stages in the Progenesis QI workflow.

Appendix 6: Using LipidBlast for Compound Identification

When using LipidBlast for the **first** time you will require internet access to download the LipidBlast databases. Click on the download link.

I Progenesis QI Tutorial HDMSe - Progenesis QI File Import Data Alignment Design Setup Peak Picking	Review Identify Review Compound Deconvolution Compounds Compounds	nonlinear A Waters Company
Identify Compounds Select your identification method: LipidBlast About this method Download others Before performing any LipidBlast searches, you'll need to download the databases. Download the LipidBlast databases Litter the compounds	Compound 0.77_520.1319n	
 Using the list below, filter the compounds to show only those you want to identify. Set the search parameters Enter the mass error tolerances for matching each compound and its fragments: Precursor tolerance: 5 ppm • Fragment tolerance: 5 ppm • 	I LipidBlast database download This will download the LipidBlast libraries from Metabolomics Fiehn	520 540 560
 Report isomers as a single search hit Bearch for identifications After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications 	LipidBlast project. Download Cancel	

Dialog will open acknowledging the source of the LipidBlast libraries, click Download.



When performing a LipidBlast search the following parameters can be set:

Precursor tolerance

Set the allowable threshold (in ppm or Da) for the difference between the database neutral mass and your measured neutral mass (default is 5ppm).

Fragment tolerance

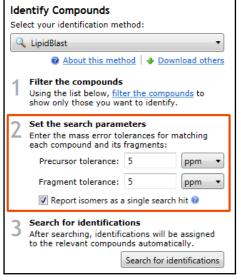
Set threshold (in ppm or Da) that determines whether a database ms/ms peak is considered to match a measured peak. The difference in m/z must fall below the threshold to be considered a match.

Report isomers as a single search hit

The search will merge hits that differ only in double bond positioning and/or geometry thus reporting isomers as a single search hit.

For example, if a search returns two identifications, GPCho(16:0/20:3(8Z,11Z,14Z)) and GPCho(16:0/20:3(5Z,8Z,11Z)), these will be reported as a single hit GPCho(16:0/20:3). When not selected, all lipid forms will be returned as separate search hits.

Having set the parameters click **Search for identifications**



On performing the search, you will get a 'mirror plot' where the upper half represents the measured/observed fragmentation and the lower half shows the currently matched fragmentation spectrum in the database.

	view Experimen ment Design Set		Review Deconvolution	Identify on Compounds	Review Compounds	Compound Statistics				nonlinea
•										A Waters Compa
entify Compo lect your identific			Compo	und 11.76_8	352.7195n					🔞 Help
LipidBlast		•								
 About 	this method 🖊	Download others	ured		Peak m/z: 5					
	n pounds elow, <u>filter the con</u> e you want to iden		Measured		Intensity: 11 Database m, Δ m/z: 0.000					
	h parameters error tolerances f d and its fragment		Database							Ι
Precursor to	lerance: 5	ppm 🔻	a a							
					1					
Fragment to	lerance: 5	ppm 🔻	-				1		- I I	
-			540	560 580	600 620 640	0 660 680 700		740 760 780	800 820	840 860 880 900
Report iso	omers as a single se					0 660 680 700 matched fragment) 720 <i>m/z</i>	740 760 780	800 820	840 860 880 900
Report iso Search for ide After searching	omers as a single se	arch hit 🕐	Legend		fragment 🔲 Unr			740 760 780	800 820	9 840 860 880 900
Report iso Search for ide After searching	omers as a single se entifications , identifications w compounds autor	arch hit 🕐	Legend Possible	: 📕 Matched	fragment 🔲 Unr			740 760 780 Retention time	800 820) 840 860 880 900
Report isc Search for ide After searching to the relevant	omers as a single se entifications d, identifications w compounds autor Search fo	arch hit @ ill be assigned natically.	Legend Possible	: Matched identification pound ID	fragment 🔲 Unr	matched fragment Adducts	m/z		800 820	9 840 860 880 900
Report iso Search for ide After searching	omers as a single se entifications d, identifications w compounds autor Search fo	arch hit @ ill be assigned natically.	Legend Possible ☆ Con ★ TG ☆ TG	Matched identification inpound ID 16:0/18:2/18:3)	fragment Unr ns: 8 Description	Adducts 3) M+H, 5) M+H,	m/z Formula CssH96O6 CssH96O6	Retention time	800 820	9 840 860 880 900
Report iso Search for ide After searching to the relevant	entifications a, identifications w compounds autor Search for blied	arch hit ill be assigned natically. or identifications	Legend Possible (con Con Con Con Con Con Con Con C	Matched identification inpound ID 16:0/18:2/18:3) 16:0/16:0/20:5)	fragment Unr ns: 8 Description TG(16:0/18:2/18:	Adducts 3) M+H, 5) M+H, 4) M+H,	m/z Formula CssH96O6 CssH96O6 CssH96O6	Retention time	800 820	9 840 860 880 900
Report iso Search for idd After searching to the relevant No filter app ompound	emers as a single se entifications a, identifications w compounds autor Search fo Search fo Diled Accepted ID	arch hit @ ill be assigned matically.	Legend Possible ☆ Con ★ TG ☆ TG ☆ TG ☆ TG	Matched identification apound ID 16:0/18:2/18:3) 16:0/16:0/20:5) 16:0/16:1/20:4)	fragment Unr ns: 8 Description TG(16:0/18:2/18: TG(16:0/16:0/20:	Adducts 3) M+H, 5) M+H, 4) M+H, 3) M+H,	m/z Formula CssH96O6 CssH96O6	Retention time	800 820	9 840 860 880 900
Report isc Search for idd After searching to the relevant No filter app ompound 12.11_854.7374n	emers as a single se entifications g, identifications w compounds autor Search fe olied Accepted ID	arch hit @ ill be assigned natically. or identifications	Legend Possible ☆ Con ★ TG ☆ TG ☆ TG ☆ TG	Matched identification apound ID 16:0/18:2/18:3) 16:0/16:0/20:5) 16:0/16:1/20:4) 17:1/17:1/18:3)	fragment Unr ns: 8 Description TG(16:0/18:2/18: TG(16:0/16:0/20: TG(16:0/16:1/20:	Adducts 3) M+H, 5) M+H, 4) M+H, 3) M+H,	m/z Formula CssH96O6 CssH96O6 CssH96O6	Retention time		9 840 860 880 900
Report isc Search for idd After searching to the relevant No filter app mopound 12.11_854.7374n 11.76_852.7195n	omers as a single se entifications g, identifications w compounds autor Search fe olied Accepted ID	arch hit @ ill be assigned natically. or identifications	Legend Possible ☆ Con ★ TG ☆ TG ☆ TG ☆ TG ☆ TG ☆ TG	Matched identification apound ID 16:0/18:2/18:3) 16:0/16:0/20:5) 16:0/16:1/20:4) 17:1/17:1/18:3) 17:0/17:2/18:3)	fragment Unit TS: 8 Description TG(16:0/18:2/18: TG(16:0/16:0/20: TG(16:0/16:1/20: TG(17:1/17:1/18:	Adducts 3) M+H, 5) M+H, 4) M+H, 3) M+H, 3) M+H,	m/z Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention time		
Report isc Report isc After searching to the relevant No filter app Ompound 12.11_854.7374n 11.76_852.7195n 11.12_876.7191n	omers as a single se entifications g, identifications w compounds autor Search fo slied Accepted ID	arch hit @ ill be assigned natically. or identifications	Legend Possible ☆ Con ★ TG(☆ TG(Matched identification apound ID 16:0/18:2/18:3) 16:0/16:0/20:5) 16:0/16:1/20:4) 17:1/17:1/18:3) 17:0/17:2/18:3) 16:1/18:1/18:3)	fragment Unr 15: 8 Description TG(16:0/18:2/18: TG(16:0/16:2/20: TG(16:0/16:1/20: TG(17:1/17:1/18: TG(17:0/17:2/18:	Adducts 3) M+H, 5) M+H, 4) M+H, 3) M+H, 3) M+H, 3) M+H,	m/z Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention time		
Report isc Search for idd After searching to the relevant No filter app Ompound 12.11_854.7374n 11.76_852.7195n 11.12_876.7191n 12.72_856.7518n	omers as a single se entifications on identifications w compounds autor Search for solied Accepted ID	arch hit @ ill be assigned natically. or identifications	Legend Possible ☆ Con ★ TG ☆ TG	Matched identification apound ID 16:0/18:2/18:3) 16:0/16:0/20:5) 16:0/16:1/20:4) 17:1/17:1/18:3) 17:0/17:2/18:3) 16:1/18:1/18:3) 16:1/18:2/18:2)	fragment Unr 15: 8 Description TG(16:0/18:2/18: TG(16:0/18:2/18: TG(16:0/16:1/20: TG(16:0/16:1/20: TG(17:1/17:1/18: TG(17:1/17:2/18: TG(16:1/18:1/18:	Adducts 3) M+H, 5) M+H, 4) M+H, 3) M+H, 3) M+H, 3) M+H, 3) M+H, 3) M+H, 2) M+H,	m/z Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention time		
Report isc Report isc After searching to the relevant No filter app Ompound 12.11_854.7374n 11.76_852.7195n 11.12_876.7191n	omers as a single se entifications on identifications w compounds autor Search for solied Accepted ID	arch hit @ ill be assigned natically. or identifications	Legend Possible ☆ Con ★ TG ☆ TG	Matched identification apound ID 16:0/18:2/18:3) 16:0/16:0/20:5) 16:0/16:1/20:4) 17:1/17:1/18:3) 17:0/17:2/18:3) 16:1/18:1/18:3) 16:1/18:2/18:2)	fragment Unr TG(16:0/18:2/18: TG(16:0/18:2/18: TG(16:0/16:1/20: TG(16:0/16:1/20: TG(17:1/17:1/18: TG(17:1/17:1/18: TG(16:1/18:1/18: TG(16:1/18:2/18:	Adducts 3) M+H, 5) M+H, 4) M+H, 3) M+H, 3) M+H, 3) M+H, 3) M+H, 3) M+H, 2) M+H,	m/2 Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention time		
Report isc Search for idd After searching to the relevant No filter app No filter app 12.11_854.7374n 11.76_852.7195n 11.12_876.7191n 12.72_856.7518n 12.83_880.7504n ""	omers as a single se entifications on identifications w compounds autor Search for solied Accepted ID	arch hit ill be assigned natically. or identifications Create	Legend Possible ☆ Con ★ TG ☆ TG	Matched identification apound ID 16:0/18:2/18:3) 16:0/16:0/20:5) 16:0/16:1/20:4) 17:1/17:1/18:3) 17:0/17:2/18:3) 16:1/18:1/18:3) 16:1/18:2/18:2)	fragment Unr TG(16:0/18:2/18: TG(16:0/18:2/18: TG(16:0/16:1/20: TG(16:0/16:1/20: TG(17:1/17:1/18: TG(17:1/17:1/18: TG(16:1/18:1/18: TG(16:1/18:2/18:	Adducts 3) M+H, 5) M+H, 4) M+H, 3) M+H, 3) M+H, 3) M+H, 3) M+H, 3) M+H, 2) M+H,	m/2 Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention time		

Note: a matched fragment is indicated in red, spanning the x axis. If you hold the cursor over a matched fragment details of the measured and database m/z are reported as well as the mass difference.

Where fragmentation data is available then the order will be based on the fragmentation score. In the absence of fragmentation data the score will be based on mass error and Isotope similarity.

Note: LipidBlast does not contain structural data. Clicking on the URL in the in the **Link** field, of the possible identifications table, takes you to Lipid Maps where a structural diagram is available.

Appendix 7: Exporting compound identities to IPA

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

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

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

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

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

🔲 Progenesis QI T	utorial HD	MSe - Progene	esis QI										×
	Review lignment	Experiment Design Setup	Peak Picking	Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics					A Waters Com	C S
-	-											A waters com	
Review Comp Using this screen		6		Find a comp	ound: Search	h	<i>P</i>	Filter compound	Is 🔻 Filter i	s active		U II	eih .
interest in your e				Compound	N	eutral mass m	/z z Ret	ten Filter comp	ounds	ig 💌	Accepted ID	Identifications	
1 Create a sh			QI Export Pati	nways Informati	on			_	×	-	7847292	1	<u></u>
In the table, on their mea		filter the com s, to generat		thways tool						🕘 🖉	49854487	4	=
for further r				thways tool fro the links belo		ow. You can fi	nd out more or	download new	0.18	8	4702	1	
		measureme	plugino usinj	g the links belo					0.14	•	9357	1	
To sort the t the relevant	table by a : column h	given value, eader.		vays tool do yo		?			0.14	٠	14708992	1	
a built with a		- 4-	IPA: Ingenui	ty Pathway Anal	ysis				• 0.27		7847488	1	Ŧ
2 Review the compounds For each compound of interest, ins alignment and peak picking:					About this	plugin 🛛 🕹 Do	ownload other plugi	ns			,		
		Review select							ntage				
You can also	o double-c	lick to review									B_HC)	
5 For each cor	mpound, s	identificatio elect one of i accepted one											
identification	To speed this up, you can <i>automat</i> identifications in compounds where the possible identifications has:												
Score ≥	50.0	Accept											
	g your dat	her processi a to external sis.				< B	ack Nex	kt > Cancel					
		Export to p	oathways tool	ArcSink 0.200		,	Ĩ						
Experiment des Review your data		ifferent nerco	ective		4							•	
Current design:			ecuve:								Sectio	on Complete	•

Export of the Accepted IDs only is

recommended then select the appropriate identifier type used for compounds in the experiment.

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

Choose which identifications to export and the type of analysis you w	ant to perform.
Select which identifications to export:	
Accepted IDs only (recommended) All possible IDs	
Select the identifier type used for compounds in the experiment:	
Human Metabolome Database (HMDB)	,
CAS Registry Number	
KEGG	
Human Metabolome Database (HMDB) PubChem CID	

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

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

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

Q Export Pathways Information								
Configure your export Choose which identifications to export and the type of analysis you want to perform.								
Select which identifications to export:								
Accepted IDs only (recommended) All possible IDs All po								
Select the identifier type used for compounds in the experiment:								
Human Metabolome Database (HMDB)								
Select the way you want to export the compounds to IPA:								
Upload expression dataset 🔹								
For expression data, choose two experimental conditions that you would like to compare.								
Baseline: Comparison:								
C_Norm								
If greater, gives a negative fold change.								
ALD								
B_HD								
K Sack Export compounds to IPA Close								

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

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

Click Export compounds to IPA .

INGENUITY	
Welcome! Please login	Contact Customer Support
Email smith@work.com Password Remember my password LOG IN Sign Up Forgot Password	Customer Support Phone: 650.381.5111 Hours: 6am - 5pm (PST) Monday - Friday (excluding holidays) support@ingenuity.com For Product and Sales related inquiries contact: 650.381.5056 sales@ingenuity.com

Waters

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

IPA will open displaying the imported data from Progenesis.

Edit View Windo	w Help	Provide Feedback Support Janucz Nykiel Close
	тер	
	Genes and Chemicals Diseases and Functions Pathways and Tox Li	
v×	Enter gene names/symbols/IDs or chemical/drug names here	SEARCH Advanced Search
set Upload - New Data	aset 2015-05-11 07:35 AM	Dataset Upload Workflow Instructions
elect File Format:	Ingenuity File Format A or B More Info	Data Unload Workflow
Contains Column Head	er: O Yes No	Data Upload Workflow
elect Identifier Type:	Human Metabolome Database (HMDB) Specify the identifier type found in t	he dataset. Use Dataset Upload to import your dataset file into IPA. Once uploaded, many different analysis options exist including
		Biomarker Filter, Molecular Tox and Core Analyses, Review
Array platform used for	experiments: Not specified/applicable Vertex Select relevant array platform as a relevant array platform array platform as a relevant array platform array platfo	che difference cype of analyses and see which one besches
Jse the dropdown men	us to specify the columns that contain identifiers and observations. For observations, select the app	ropriate expression value type. your needs.
		1. To upload a dataset file, click here.
aw Data (567) Dataset	Summary (182)	
		······································
ID	▼ Observation 1 ▼	Look jn: 🗀 Multiple Rank 👻 🕼 🏠 💕 😂 🗁
	Fold Change	Batch (mult-timepoint)
		Affy_with_p=value_and_fold.txt UL_with_fold and_Normalized.txt
1 HMDB01988	Infinity	
2 HMDB05015 3 HMDB40639	1.72884040538655	
3 HMDB40639 4 HMDB40285	3.98236765102672 2.1144344578431	
5 HMDB39847	1.91509320582337	
6 HMDB05049	1.36771507748685	
7 HMDB30934	-2.37287619375142	File (jame:
8 HMDB37820	1.69294555757503	Files of Type: All Files
9 HMDB15294 10 HMDB39350	-1.4569308616069 -1.46035570178035	Open Cancel
11 HMDB41326	1.89049486451889	
12 HMDB34382	1.75866062252427	Select the dataset file from your computer and click the
13 HMDB35337	3.98452117071048	Open button.
14 HMDB36734	2.38572167550178	3. Select Flexible format for the file format from the
15 HMDB13809 16 HMDB14821	-1.25125092715796 1.08777276460426	dropdown menu.
16 HMDB14821 17 HMDB00782	1.14588290027386	
18 HMDB40806	1.83433974738324	 Select an Identifier Type from the dropdown menu. IPA supports many identifiers and symbols and will attempt to
19 HMDB30926	2.51317063326778	guess at the type of identifier in your dataset file. To
20 HMDB40179	1.42264275983645	override the selection, uncheck the option and simply select
21 HMDB13250	1.17037352363088	the most appropriate one. If more than one type of identifier
22 HMDB37712 23 HMDB38057	1.50281207491669 1.43917459561637	exists in your dataset, select all appropriate ones.
24 HMDB40285	1.84463305905821	GenBank
25 HMDB02725	1.80218245845514	
26 HMDB29978	2.11110785701812	Affymetrix
27 HMDB37272	1.27732135837318	Afymetrix SNP ID
28 HMDB39848 29 HMDB29828	1.14877655967276 -1.35134004957123	Agilent
30 HMDB30932	1.82757219666658	Applied Biosystems
31 HMDB13300	1.38248130204499	CAS Registry Number
32 HMDB32791	1.21282984334921	CodeLink Entrez Gene (LocusLink)
		□ Entrez Gene (Locusuink) GenBank
		GenPept
		CI Number

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

Appendix 8: Power Analysis

Power analysis is a statistical technique, which is used to gauge how many replicates are needed to reliably observe the abundance differences in your data. It is available through the Compound 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 Compound Stats screen. A selection of 4 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 compounds according to how similar their abundance profiles are.
Ľ	Power Analysis How many replicates should I run? What is the power of my experiment?
	Adduct Abundance What adducts do my runs contain?

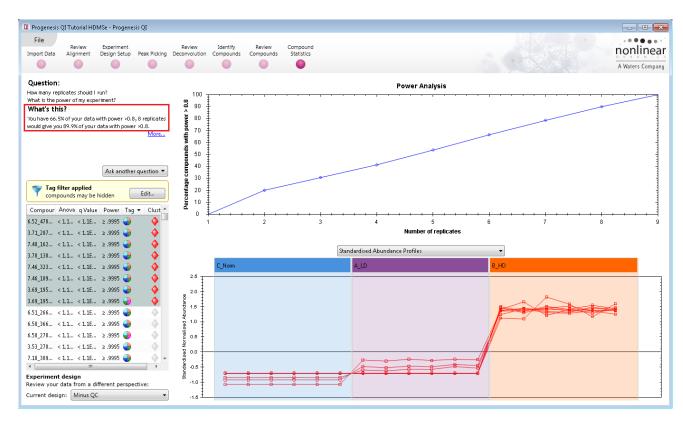
Select the option

'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 compounds with a power >0.8'

Using only the compounds tagged **Anova p<0.05 (i.e. 1342 compounds)**, as an example, view the power analysis.



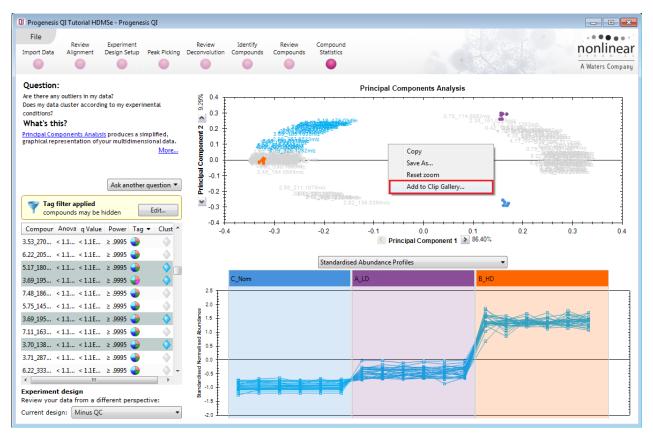
This is displayed graphically showing that 66.5% of the 1342 compounds have a power of 80% or that 8 replicates would give you 89.9% of your data with power > 0.8.

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

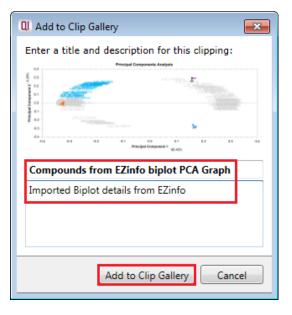
At every stage of the Progenesis QI workflow the images and data tables can be added to the Clip Gallery.

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

As an example of using the Clip Gallery, at the **Compound Statistics** view, displaying the PCA plot. Right click on the **Biplot** View and select Add to clip gallery....



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



Enter details as required and click Add to clip gallery

	QI P	rogenesis QI Tutorial HDMSe - Pro	oge
		File	
To view, edit and/or export from the clip galley the gallery can be		Save	tu
accessed from the File menu.		Close	L
		Export inclusion list Export to EZinfo Open EZinfo project	ф
		Experiment properties	s i
		Show Clip Gallery	im
Selecting an image in the gallery makes available an Actions menu that allows you to manipulate the output of the image.	×	Exit	

Clip Gallery							- • •
Clip Gallery			Ite	em size:	•	Search	Q
Actions Edit Copy Imag Copy Title Copy Desc Export Imag		ription	nesis (וך	Data analysi	s performed using: Progenesis	
Compounds from EZinfe Graph Imported Biplot details fro	-	use in presentat	logo n image that you ca tions and posters to are was used in your	an / r s	A high-reso use in prese	5 QI analysis badg lution image that yo entations and posten oftware was used in y	u can s to
						Export all.	. Close

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

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