

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	29
Stage 7: Review Deconvolution	31
Stage 8: Identify Compounds	39
Identification of compounds using the search method Progenesis MetaScope	40
Incorporating Theoretical Fragmentation	42
Creating and using databases	45
Additional Identification Methods	50
Stage 9: Review Compounds	61
Compounds separated by Drift Time	67
Exporting Compound Data to Pathways Tool(s)	69
Stage 10: Compound Statistics	72
Appendix 1: Stage 1 Data Import and QC review of data set	80
Appendix 2: Stage 2 Processing warnings and failures	85
Appendix 3: Licensing runs (Stage 3)	88
Appendix 4: Manual assistance of Alignment	89
Appendix 5: Within-subject Design	94
Appendix 6: Using LipidBlast for Compound Identification	96
Appendix 7: Editing an adduct to add a missing isotope	98
Appendix 8: Exporting compound identities to IPA	103
Appendix 9: Power Analysis	106
Appendix 10: Using Clip Gallery to Save and Export Pictures and Data	107
Appendix 11: Waters Machine Specifications and approx. timings	109

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 80) 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 60 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

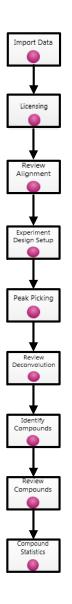
Progenesis QI adopts an intuitive **Workflow** approach to performing comparative LC-MS data analysis. The following user guide describes the various stages of this workflow (see below) focusing mainly on the stages from Alignment to Compound Stats.



Stage

Description

Page



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
Review Normalisation: examine data normalisation methods	29
Review Deconvolution : review and edit the various adduct, forms of a compound	31
Identify Compounds: search identity of compounds using Progenesis MetaScope and or other search engines	39
Review Compounds : managing possible compound identities exploring identity and expression between conditions	61
Compound Statistics : performing multivariate statistical analysis on tagged and selected groups of compounds	72

Restoring or starting a new Metabolomics Tutorial

If working with your original data files then refer to Appendix 1 page 80.

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

Q Progenesis QI		
File		
Experiments		nonlinear
		A Waters Company
Recent experiments	Search P	Getting started Here are some resources to help you get started with Progenesis QI: • Fundamental concepts of the Progenesis QI workflow • User quide and tutorial data
Q	Open Experiment	 Frequently-asked questions
•	Image: Cust Image: Progenesis QI Tutorial HD Image: Cust	EZinfo version information Version 3.0.3.0 of EZinfo is installed.
The	Desktop Name Libraries Progenesis QI Tutorial HDMSe ProgenesisQIArchive Music Pittures Videos Videos Andy Borthwick Computer Cacal Disk (C:) VD RW Drive (Do) Removable Disk (F:) Immediate	Metabolomics Meeting, 20-21 Oct 2015 Waters MS HQ, Wilmslow, UK Are you looking to apply mass spectrometry techniques for routine metabolic profiling and advanced characterisation of metabolomes? Find out more and register at: www.waters.com/metabolomics2015. Latest blog posts • Progenesis QI and the IKnife: the
	File name: Progenesis QI Tutorial HDMSe.Pro Experiments and Archives (*.Prr Open Cancel	cutting edge of food monitoring Progenesis QT v2.1 - available to download now MetSoc2015 - Bridging the gap between community and industry
Other experiments	1	Progenesis SDF Studio v1.0: using MOL files just got easier ASMS 2015: Mass specs, plugs and rock 'n' roll
new 📴 Open	2.1.5674.29829	

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.

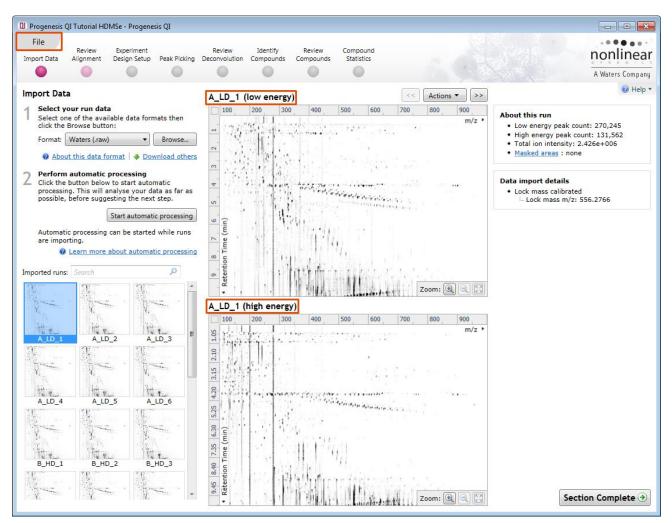
Select the **Create a new experiment** option and select the folder in which you placed the archive, using **Browse**.

U Import Experiment from Archive		
Import experiment from archive After importing the experiment from this archive, any changes to the experiment will be saved to the location below, not back to the archive.		
Replace an existing experiment Experiment to replace:		
Experiment name:	Progenesis QI Tutorial HDMSe	
Save to folder:	s\Customer Data\Progenesis QI Tutorial HDMSe_2.1 Browse	
	Import Cancel	

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.

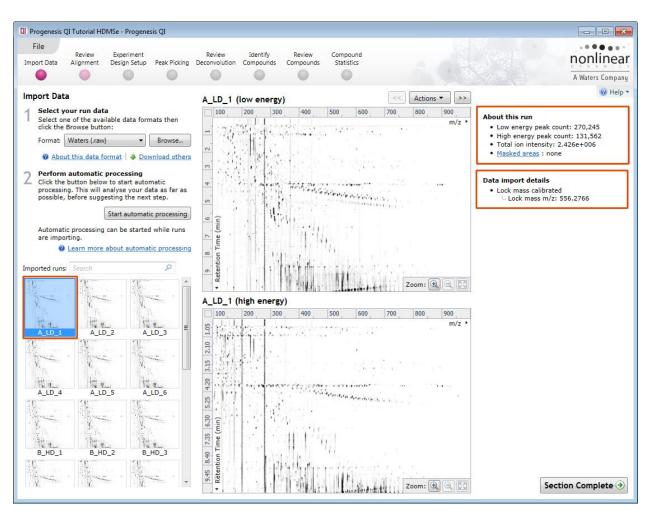
As this example is using HDMSe data it contains both low and high energy exact mass data.

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

Runs in this experiment: 22 Feature detection: High resolution Peak processing: Profile data Adduct definitions used in the experiment:				
Name	Adduct Mass		Abundance	
M+H	1.007276432	1		*
M+2H	2.014552864	2		_
M+NH ₄	18.03382553	1		
M+Na	22.98922068	1		E
M+H+Na	23.99649711	2		
M+CH ₃ OH+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.

Q Start automatic processing	
Select an alignment reference To compensate for drifts in retention time, all runs in the e to a single reference run.	xperiment must be aligned
How do you want to choose your alignment reference?	
Assess all runs in the experiment for suitability	
Output the most suitable run from candidates that I select	
🔘 Use this run:	
A_LD_1 -	
For information on choosing the alignment reference, and 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 select the run with the greatest similarity to all other runs as the alignment reference.

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

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

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

This method is appropriate when you have some prior knowledge of your runs suitability as references: i.e when all the candidate runs are pooled samples or 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**.

0 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 Use the most suitable run from candidates that I select	
🔘 Use this run:	
A LD 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> .	
< Back Next > Cancel	

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**.

Q Start automatic processing		Q Start automatic processing
Select your alignment reference candidates To mark a run for assessment as an alignment reference candidate, double-click it or use the button below to mark multiple runs at once.		Select your alignment reference candidates To mark a run for assessment as an alignment reference candidate, double-click it or use the button below to mark multiple runs at once.
Mark selected runs Clear all Search P		Mark selected runs Clear all Search P
D_QC_1 D_QC_2 D_QC_3	N	Image: Constraint of the second sec
		D.QC.4
0 of 22 runs marked as candidates		4 of 22 runs marked as candidates
< Back Next > Cancel		< Back Next > Cancel

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

QI Start automatic processing	
Automatic alignment After selecting the experiment's alignment reference, the software car automatically align all runs.	n also
After the alignment reference is chosen, do you want to start automatic a Ves, automatically align my runs	lignment?
< Back Next >	Cancel



Progenesis QI User Guide

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.

Q Start automatic processing	QI Start automatic processing
Experiment design Experiment designs allow you to group and compare your samples according to their experimental conditions.	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.	By defining an experiment design in advance, statistical measures such as ANOVA can be calculated automatically.
Set up an experiment design	Set up an experiment design
Enter a name for the experiment design:	Enter a name for the experiment design:
Conditions Load the criteria for grouping runs from this file: Browse Group runs by: <pre><no groups="" valid=""></no></pre>	Conditions Load the criteria for grouping runs from this file: I Tutorial_HDMSe\QI HDMSe Raw data\Progenesis QI_HDMSe Tutorial.spl Browse Group runs by: VERSION VERSION Condition Index
What file formats are supported?	What file formats are supported?
< Back Next > Cancel	< Back Next > 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.

QI 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

QI Automatic Processing (27%) Automatic processing Current step: Choosing an alignment reference	
Importing runs: 22 of 22 processed Selecting reference: Choosing an alignment reference Aligning runs: Pending Creating design: Pending Peak picking: Pending Cancel Cancel	QI Automatic Processing (42%) Automatic processing Current step: Aligning 'A_LD_2' ✓ Importing runs: 22 of 22 processed ✓ Selecting reference: D_QC_4 ✓ Aligning runs: 1 of 21 processed Creating design: Pending Peak picking: Pending
	Cancel

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

Q Automatic Processing (47%)	
Automatic processing Current step: Aligning 'B_HD_2'	
 ✓ Importing runs: 22 of 22 processed ✓ Selecting reference: D QC 4 ✓ Aligning runs: 7 of 21 processed 	Automatic Processing 33 Automatic processing Current step: Analysing
Creating design: Pending Peak picking: Pending	 ✓ Importing runs: 22 of 22 processed ✓ Selecting reference: D_QC_4 ✓ Aligning runs: 21 of 21 processed ✓ Creating design: Created
Cancel	Peak picking: Picking
	Cancel

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

Q Processing Complete			×
Automatic process			
Time taken: 10 minutes :	14 seconds		
 Importing runs: 	22 of 22 processed		
 Selecting reference: 	D_QC_4		
 Aligning runs: 	21 of 21 processed		
 Creating design: 	Created		
 Peak picking: 	6137 compound ions found		
		Clo	ose

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 warnings and managing failure messages are available in Appendix 2 (page 85)

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

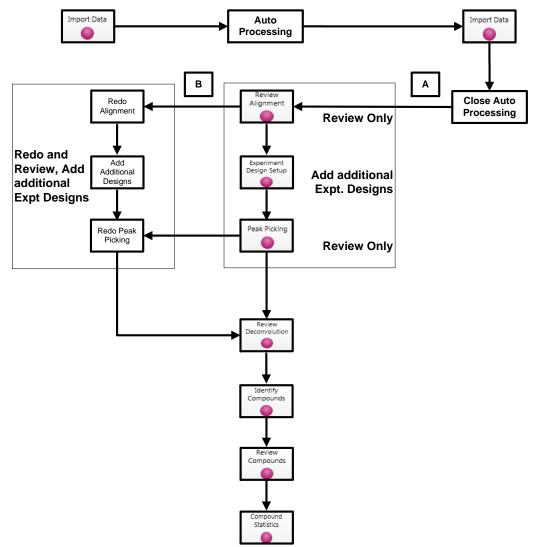
When the Automatic Processing completes click **Close**, the software remains at the Import Data stage of the workflow where you can review the chromatography of the imported runs and remove/replace any runs that have failed to import.

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.

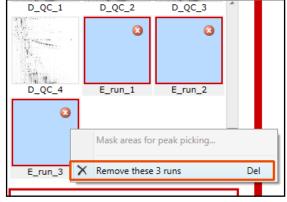
Note: for this data set do not remove any files

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 88)



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 mport Data	Licensing	Review Alignment	Experiment Design Setup	Peak Picking	Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics				nonlin	nea
								•				A Waters Co	ompai
ongle Li	icense Run:	5											
This installa licensed ru	ation is currer ns only.	ntly restricted	to analyse	Run name							Licence state	License this run	^
Toliconso	vour runs, vo	u need an ev	aluation	C:\Users\andy	.borthwick\Docu	uments\Custome	r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	enesis QI	Unlicensed	\checkmark	
	de which can			C:\Users\andy	.borthwick\Docu	uments\Custome	r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	V	
representat	tive.			C:\Users\andy	.borthwick\Docu	uments\Custome	r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	V	
Once licens	sed, your runs	s can be ana	lvsed on	C:\Users\and	.borthwick\Docu	uments\Custome	r Data \Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	\checkmark	
any installa	tion of the so	ftware. The li	cence is	C:\Users\and	.borthwick\Docu	uments\Custome	r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	\checkmark	
automatica experiment	Ily included w	hen archivin	g an	C:\Users\and	.borthwick\Docu	uments\Custome	r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	\checkmark	
	have been lic		dh a s	C:\Users\andy	.borthwick\Docu	uments\Custome	r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	V	E
	click here to m						r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	\checkmark	
available or	n this compute	er.		C:\Users	Unable to save	experiment			— ×	nesis QI	Unlicensed	\checkmark	_
If you have one, you can <u>open a licence file</u> to			<u>e file</u> to	C:\Users					Unlicensed	\checkmark	_		
install.				C:\Users	C:\Users You cannot save the analysis without a valid license. If you close now your analysis will be lost.				Unlicensed	V	_		
lf you have	just installed	a dongle, <u>clic</u>	k here.	C:\Users	· · ·					nesis QI	Unlicensed	v	_
				C:\Users						nesis QI	Unlicensed	v	_
				C:\Users				ОК	Cancel	nesis QI	Unlicensed	v	
				C:\Users					Cuncer	iesis QI	Unlicensed	v	-
									is QI v2 Tutorial\Proge	-		v	_
									is QI v2 Tutorial\Proge			7	_
				C:\Users\andy	.borthwick\Docu	uments\Custome	r Data\Progenes	is QI v2\Progenes	is QI v2 Tutorial\Proge	nesis QI	Unlicensed	V	-
							Run licence of	code:	·		Use Lie	cence Code	•
												ion Comple	_

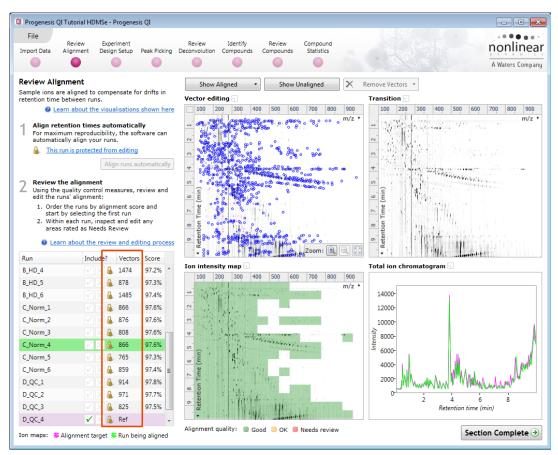
Note: although the analysis will be lost (Alignment Peak Picking etc) the experiment file and all the 'loaded pre-processed runs are retained'. This allows you to reopen the experiment, when you have a licenced code and/or dongle available, and redo the analysis steps without having to reload the data.

For more details on managing the licencing of your data refer to Appendix 3 (page 88)

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.

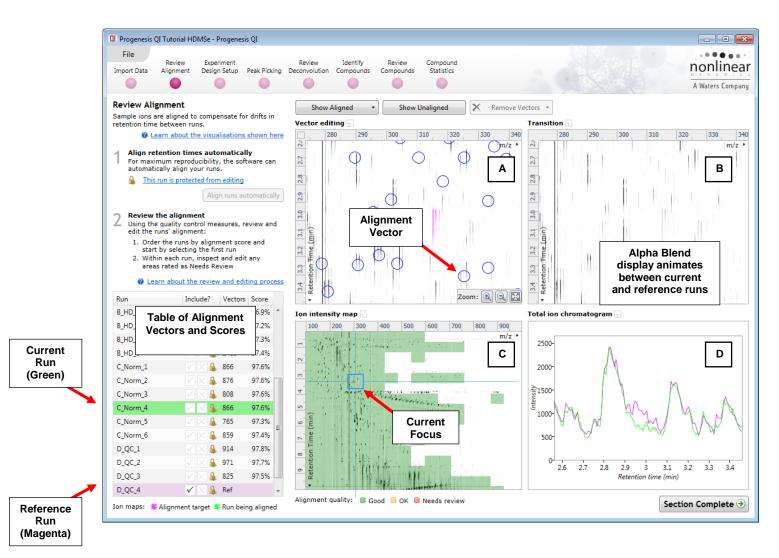
Sar	view Alignment nple ions are aligned to compensate for drifts in		
ret	ention time between runs.	🛛 Protected from editing	
1	 Learn about the visualisations shown here Align retention times automatically For maximum reproducibility, the software can automatically align your runs. This run is protected from editing Align runs automatically 	Delete existing analysis? If you change the alignment, it will invalidate the current analysis including 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.	
2	 Review the alignment Using the quality control measures, review and edit the runs' alignment: 1. Order the runs by alignment score and start by selecting the first run 2. Within each run, inspect and edit any areas rated as Needs Review ② 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 89).

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.

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



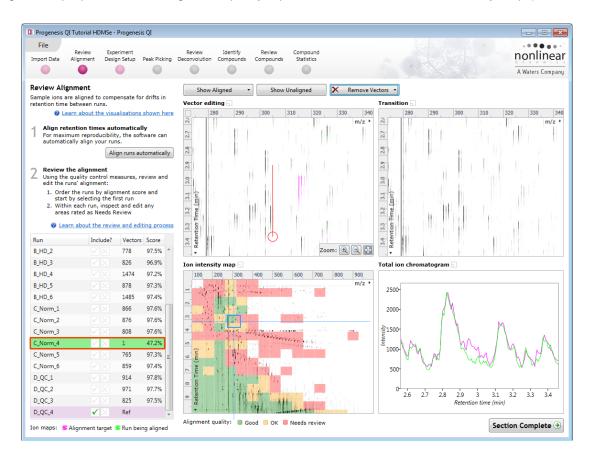
Note: the Retention Time alignment of this data is good between the runs as indicated by a high percentage score (in the absence of any vectors) and the Alignment quality showing as 'all' green.

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'.

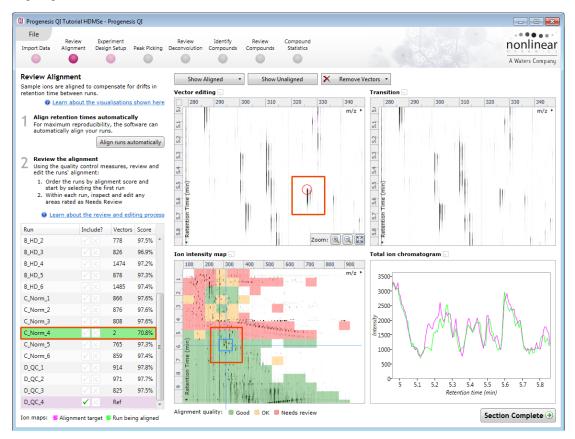
Note: the manual vector is red, to distinguish it from the automatic vectors (blue).

Run	Include?	Vectors	Score	
B_HD_3	X X 4	826	96.9%	*
B_HD_4	< X 🌡	1474	97.2%	
B_HD_5	🗹 🗙 🔒	878	97.3%	
B_HD_6	< X 🌡	1485	97.4%	
C_Norm_1	🗹 🗙 🌲	866	97.6%	
C_Norm_2	X 4	876	97.6%	
C_Norm_3	🗹 🗙 🔒	808	97.6%	
C_Norm_4	🗸 🗙 🌡	866	97.6%	
C_Norm_5	🗹 🗙 🌲	765	97.3%	
C_Norm_6	< X 🌡	859	97.4%	_
D_QC_1	🗹 🗙 🔒	914	97.8%	
D_QC_2	X 4	971	97.7%	
D_QC_3	🗹 🗙 🔒	825	97.5%	
D_QC_4	🖌 🖂 🔒	Ref		-

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 click and release **without** deliberately misaligning the vector.



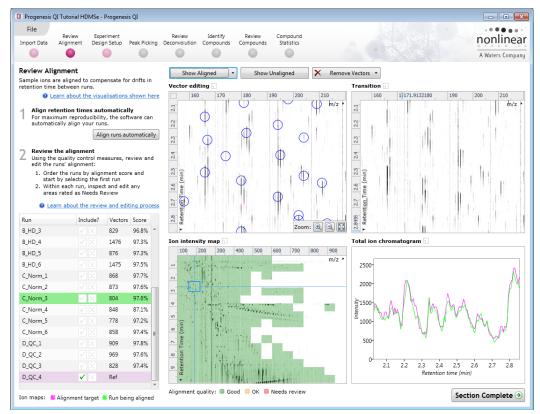
Notice how the correctly placed vector markedly improves the Score this is also reflected in the increased proportion of the Ion Intensity Map showing green indicating a good quality alignment.

Waters THE SCIENCE OF WHAT'S POSSIBLE."

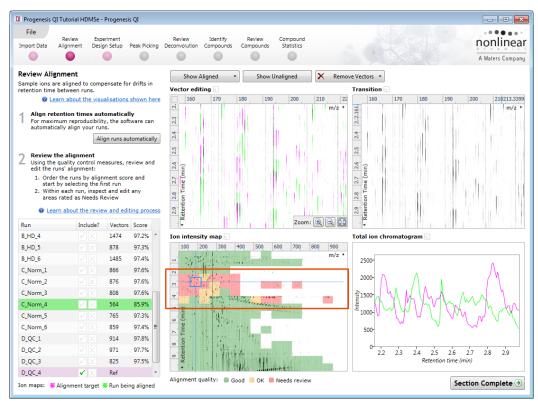
Reviewing quality of alignment vectors

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 lon Intensity Map showing green indicating good quality alignment and the Transition view showing features pulsing slightly but not moving up and down.



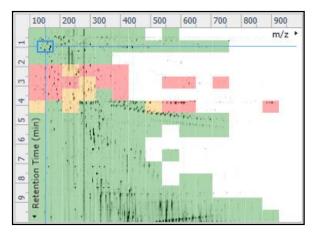
Using the **Simulated** miss-aligned example from the previous section to explain the review process for alignment, the alignment looks as below with a region of poor alignment (highlighted in red).

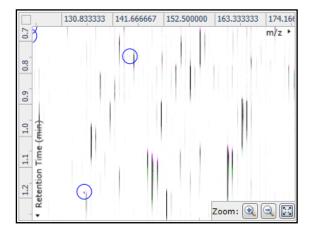


Reviewing Quality of Alignment

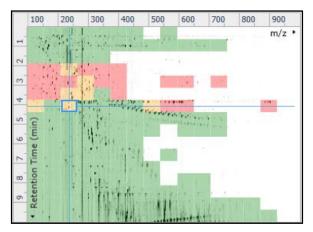
At this point the quality metric, overlaid on the Ion Intensity Map as coloured squares, acts as a guide drawing your attention to areas of the alignment. These range from Good (Green) through OK (Yellow) to Needs review (Red). Drag out a 'Focus' area that corresponds to one of the coloured squares.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

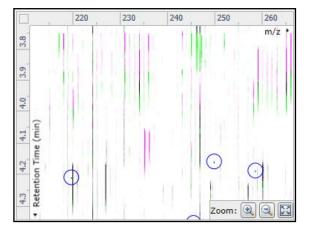
For a 'green' square the majority of the data appears overlapped (black) indicating good alignment. When viewed in the Transition view the data appears to pulse.



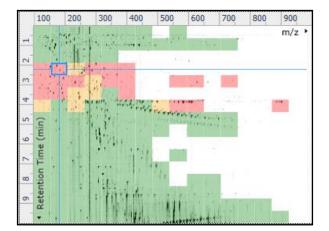


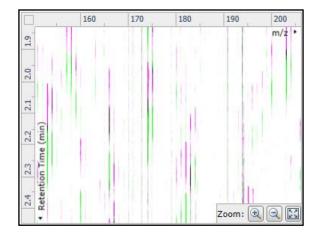
For a 'yellow' square some of the data appears overlapped (black) indicating OK alignment. When viewed in the Transition view some of the data appears to pulse.





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.

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 89).



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.

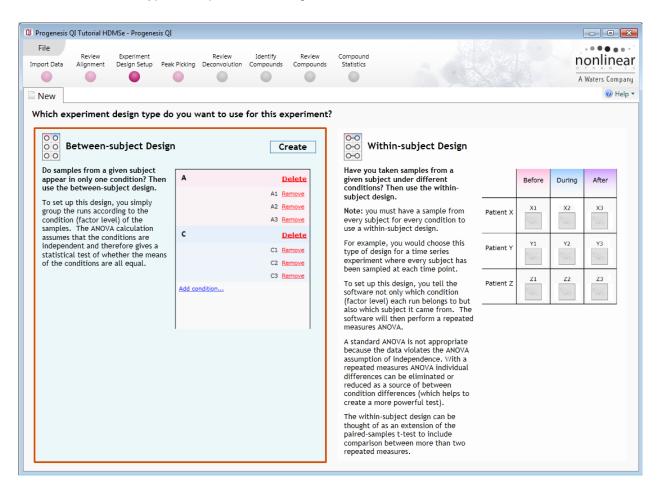
Progenesis QI User Guide

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 94

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.

🛛 Create New Experiment Design 🗾 💽
Enter a name for the experiment design:
All Conditions
How do you want to group the runs?
Group the runs manually
Copy an existing design:
Import criteria from a file: Browse
Group runs by: <pre></pre> <pre></pre>
What file formats are supported? Create design Cancel

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

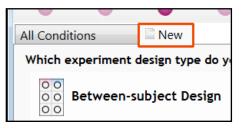
File mport Data	Review Alignment	Experiment Design Setup	Peak Dicking	Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics			nonlinea
										A Waters Compan
Il Conditio	ns I ×	New						/ moultains		🔞 Help
etup condi	itions			Runs Ac	ld Selected Rur	ns to Condition	▼ Search	Q		
etup the cond e.g., control, our samples t	drug A, etc)	ou want to co , and then ass t condition.	mpare below ign each of		Add to new c	ondition	В_	HD_2	B_HD_3	*
C_Norm	ı		<u>Delete</u>		C_Norm A_LD		10	342	Sept.	
		_	m_1 <u>Remove</u> m_2 <u>Remove</u>		_					
			m_3 <u>Remove</u>		B_HD_4		B_	HD_5	B_HD_6	
		-	m_4 <u>Remove</u> m_5 <u>Remove</u>		940.		14	346.	24	
			m_6 <u>Remove</u>		1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1					=
A_LD			Delete							
		-	D_1 <u>Remove</u>		D_QC_1			_QC_2	D_QC_3	
		-	D_2 <u>Remove</u> D_3 <u>Remove</u>				11.1	346 - L		
		-	D_4 <u>Remove</u>							
		-	D_5 <u>Remove</u>		D_QC_4					
Add condition	<u>h</u>	A_L	D_6 <u>Remove</u>							
				L					Sectio	on Complete 🏵

To create a new condition

- 1. Select the runs for the condition by clicking on the required icon in the **Runs** panel, as shown.
- 2. Press the 'black triangle' next to the Add Selected Runs to Condition button on the main toolbar.
- 3. Select Add to new condition... from the drop down menu.
- 4. A new condition will appear in the **Conditions** panel on the left.

- 5. Rename the condition (e.g. 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.



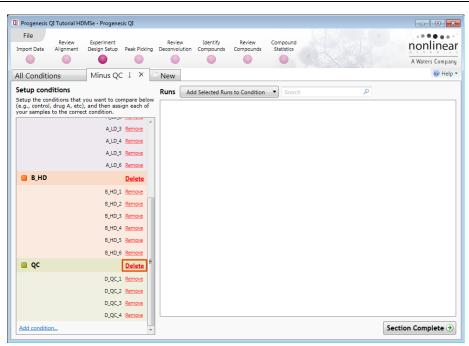
For this example there is a **Progenesis QI_HDMSe Tutorial.SPL** file available in the Experiment Archive you restored at the beginning of this tutorial exercise.

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.

Create New Experiment Desig	n 💌
Enter a name for the experime	ent design:
Minus QC	(-
How do you want to group the	runs?
Group the runs manually	
Copy an existing design:	All Conditions 💌
Import criteria from a file:	igenesis QI_HDMSe Tutorial.spl Browse
Group runs by:	Condition
	VERSION
What file formats are suppor	Condition
	Index

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**.



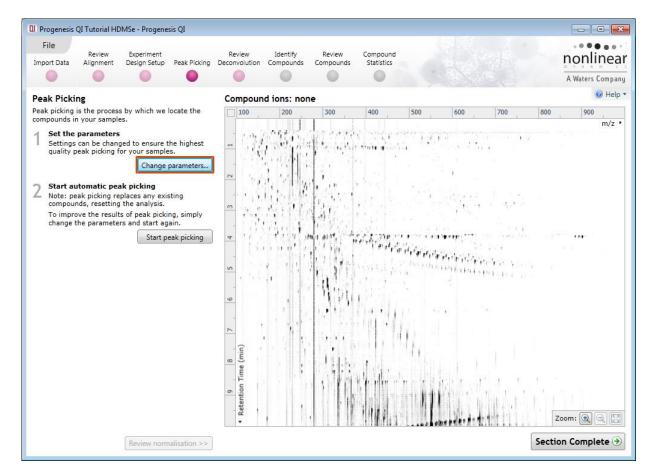
Note: On deleting a condition the runs will reappear in the Runs window.

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.

The ion intensity map displayed before Peak picking is for the Alignment Reference run. As the Peak Picking takes place this is replaced by the Aggregate map generated from all the aligned runs in your experiment.

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.

QI Peak Picking Parameters		Q Peak Picking Parameters	×
Runs for peak picking Peak picking lin Choose runs for peak picking You can tick or un-tick each run to control which will be used by the peak picking algorithm. Although any run which is left un-ticked will not	nits Retention time limits	Runs for peak picking Peak picking Sensitivity You can adjust the sensitivity of the peak picking algorithm using these different methods. Each sensitivity method examines the	ng limits Retention time limits
affect the compound ion outlines, it will still have outlines added to it and will be available in the experiment design setup. Learn more about why you might not want to select all runs.	 ✓ A_LD_3 ✓ A_LD_4 ✓ A_LD_5 	intensities of groups of MS peak to judge whether they are likely form part of an ion or whether they represent noise and so should be ignored. Peaks that ar rejected as noise will not be use to build ion outlines.	to The automatic sensitivity method uses a noise estimation algorithm to determine
		Chromatographic peak wi The chromatographic peak widt gives the length of time over which an ion has eluted. If you a minimum peak width, any ion that has eluted over a shorter period will be rejected.	Apply a minimum peak width
	OK Cancel		OK Cancel

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

Note: 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.

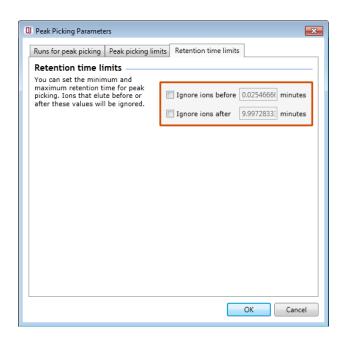
The sensitivity of the peak picking can be controlled by adjusting settings under the **Peak picking limits** tab.

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	n peak widt	th
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.



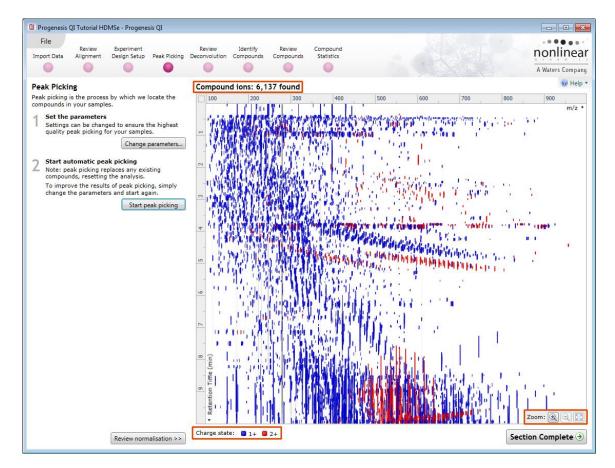
Click **OK** to close the Peak Picking Parameters dialog.

Press Start peak picking to start the detection process.

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

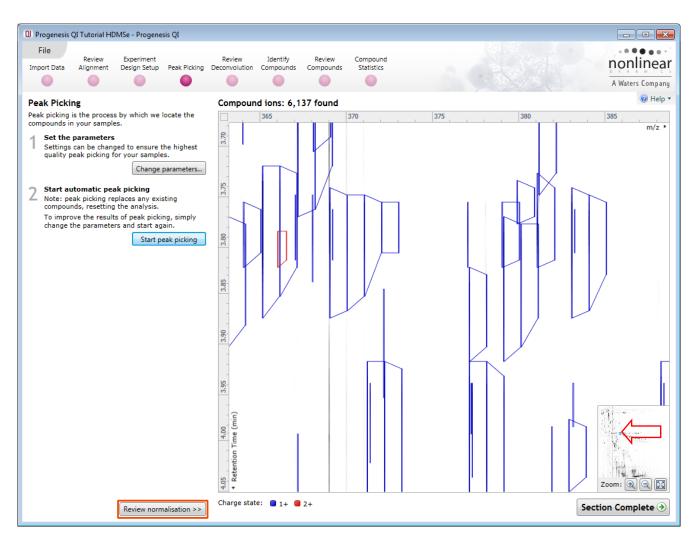
Analysing		

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 at the top of the ion intensity map.

To inspect the detection more closely drag out an area of interest with the zoom tool.



The 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.

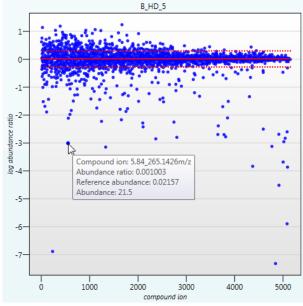


Calculation of Normalisation Factor:

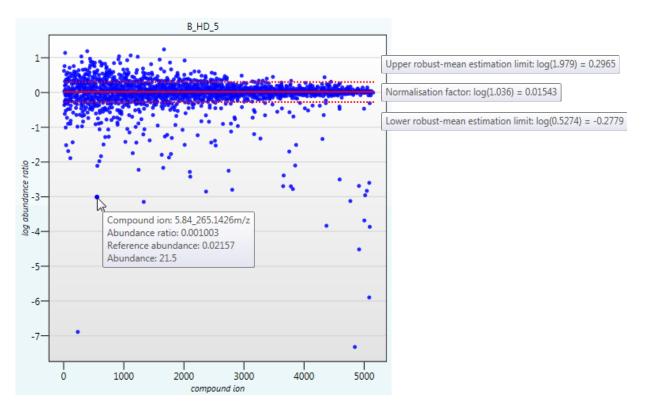
Progenesis QI will automatically select one of the runs that is 'least different' from all the other runs in the data set to be the 'Normalising reference'. The run used is shown above the table of Normalisation factors. (in this example it is A_LD_4)

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 Me	ethod	
Normalise to	all compound	ls 🔻		
Normalise to	all compound	ls		
Normalise to	a set of house	ekeeping comp	ounds	
Normalise u	sing total ion i	ntensity		
Normalise to	external stand	dard		
Don't use an	y normalisatio	n		

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_458.2749n**, 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.

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.

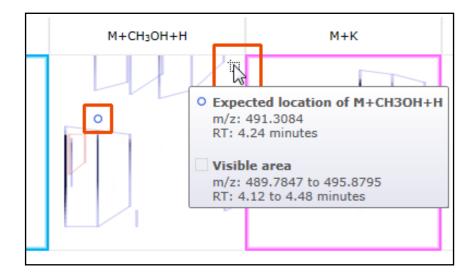
Using **Compound 4.24_458.2749n** we can look at how the information is displayed on the multiple views following deconvolution.

Grid size: 🔹 🗆 🔲 🛈 About this	elp 🔻				
About this compound					
Compound 4.24_458.2749n has the following properties:					
Retention time: 4.237 mins					
 Neutral mass: 458.2749 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.					

Note: for clarity the 'About this' pane can be hidden by clicking on O About this

QI Progenesis QI Tutorial HDMSe - Progenesis QI				
File				
Review Experiment	Review Identify Review	Compound		nonlinear
Import Data Alignment Design Setup Peak Picking	Deconvolution Compounds Compounds	Statistics		
		•		A Waters Company
Review Deconvolution	Compound 4.24_458.2749n		Grid size: 🔹 🗖 🔲	About this
After peak picking, ions are grou a process called deconvolution. Montage of	F MINH, A			
W How doe detected	M+NH4 A	M+Na M+Ci	H3OH+H M+K	M+ACN+H
Review the compounds peaks		E		
Select compounds below to view:				
 areas of the ion map showing the different adduct forms 			•	L / L
 the mass and retention time profiles of their 				
ions				
These views can be used to confirm the validity of your interesting compounds.				\mathcal{V}
How do I use this screen?				
Optimise for your samples				
If any compounds have an ion whose pr				
can be removed. Likewise, missing com	ss Spectra	Chromatograms for Adducts		
ions can be added.	r Adducts	for Adducts		•
No filter applied	Mass Spe	ectra	Chromatograms	Peak
Create				
Compound Accepted ID C - 4	100-	100-		
0 4.39 516.2805n		D	E	
0 4.35 556.2736n	-08 20-			F
0 4.34 502.3022n	tiative Intensity -00 -00 -00	elative Intensity -09 -09		
○ 4.30 542.2944n	tive	tive		
○ 4.28 472.2537n	Pag 40-	9a 40-		
• 4.24 458.2749n	20-	20-		
○ 9.73_402.2998n	20	20		
○ 4.22_225.9611n N				
○ 4.20_442.2051n N	457.5 458 458.5 459 459.5 Neutral I	460 460.5 461 461.5 462 Mass (Da)	4.22 4.24 4.26 4.28 4.3 4.32 4.34 Retention time (min)	4.36 4.38
○ 4.19_498.2686n N				
0 4.10 414.2488n			S	ection Complete 🏵
< >			2	complete (

- (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.

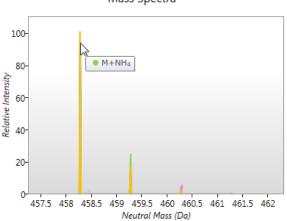


Note: hover the cursor over the icon on the top right of the panel and you will get information about the expected location and the area displayed.

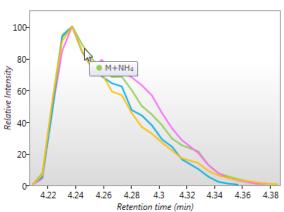
(C) The table displays the information known about the current compound with regards to the number of adducts, identity of the adducts and where deconvolution has been successful the Compound's Neutral mass is displayed. Note: when you hover the cursor over the Neutral mass for the current compound it 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_458.2749n**.

Compound	Accepted ID	Tag 🔹	Adducts		Adduct count	Neutral mass	m/z	Retention time	Fragmented?	
O 4.30_542.2944n			M+H, M+	+NH4, M+Na, M+K	4	542.2944	.3283	4.302	Yes	
O 4.28_472.2537n			M+H, M+	+NH4, M+Na, M+K	4	472.2537	4 2875	4.28	Yes	
• 4.24_458.2749n			M+H, M	+NH₄, M+Na, M+K	4	458.2749	4 7087	4.237	Yes	
O 9.73_402.2998n			M+H, M+	Based on these adducts:		402.2998	4 m/z			
O 4.22_225.9611n			M+H, M+		4	225.9611	a *	z of the most al	oundant compound	
0 4.20_442.2051n			M+H, M+	 M+NH₄ at m/z=476.3087 M+Na at m/z=481.2619 	4	442.2051			ach ion's abundance of un-normalized	
O 4.19_498.2686n			M+H, M+	• M+K at m/z=497.2361	4	498.2686			n across all samples.	
O 4.10_414.2488n			M+H, M+	+NH4, M+Na, M+K	4	414.2488	432.2827	4.104	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.







(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.





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

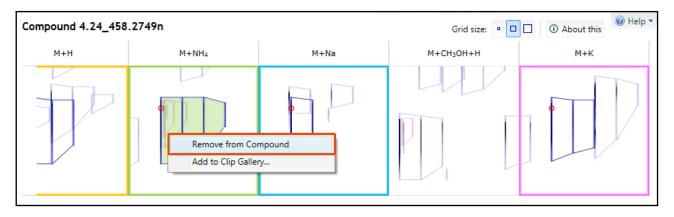
Mass Spectra

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_458.2749n**, 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.

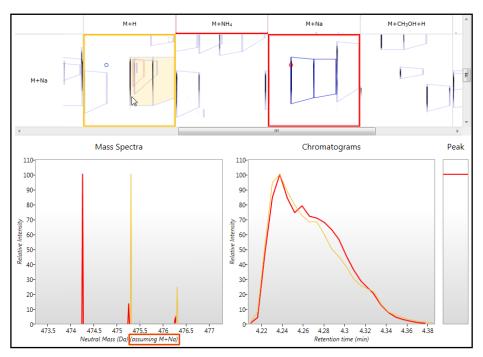
Try removing all the adducts from Compound 4.24_458.2749n to leave just the M+K.

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**

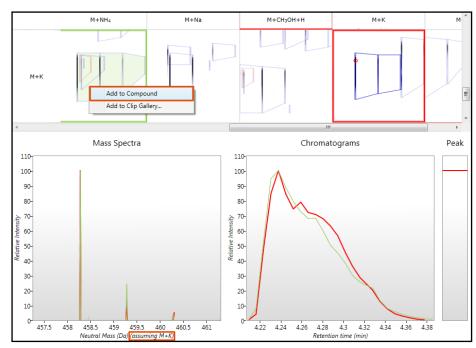
Note: because the peak is a charge state +1 this excludes M+2H and M+H+Na as being the compound's sole adduct on the basis of charge.



As an example to explore the posibility 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.

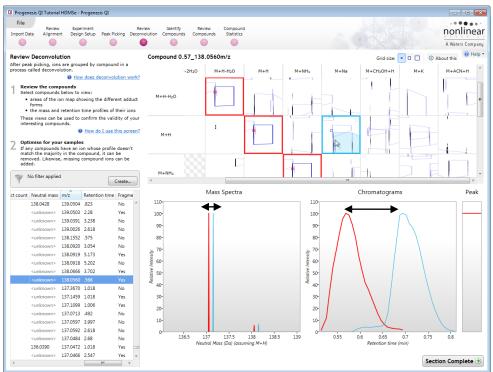


To add the $M+NH_4$ adduct to the compound right click then add this second adduct to the compound. As you add the second adduct the table will update to indicate that the compound has 2 adducts, a **Neutral mass** is calculated (n appears after the compound in the table) and the matrix is replaced by 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



Waters

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).

🗙 Delete tag 🔓 Merge tags
OK Cancel

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...**

No filter a				Create
Compound	Accepted ID	Tag 💌	Adducts	Adduct co
0 4.05_454.2413	n		M+H, M+NH ₄ , M+Na, M+K	4
0 3.78_607.0605	n		M+Na, M+H+Na, M+K	3
0 0.50, 479 6654	n	_	M+H-H ₂ O, M+Na, M+K	3
0 9.47. No	tags to assign		M+H, M+Na, M+K	3
0 0.77 Ne	w tag		M+H-2H2O, M+H-H2O, M+NH4	3
0.46 Qu	ick Tags	 And 	ova p-value	3
0.46 駍 Edi	t tags	Ma	x fold change	3
0 4.45 Ad	d to Clip Gallery	Mir	Minimum CV	
4.49_630.3470	n	Not	identified	3
0.46_700.6089	n	Not	fragmented	3
0 9.57_582.4144	n	Sep	Separated by drift time	
0.79_244.0703	n	Ide	Identified and separated by drift time	
0.46 560.6784	n		M+NH4, M+Na, M+K	

🛛 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

Commented	Assessment ID	Tee	-	Andrewster	م باط	
Compound	Accepted ID	Tag	Tag 💌 Adducts		Adduc	tc
 4.05_454.2413n 				M+H, M+NH4, M+Na, M+K	4	
 3.78_607.0605n 				M+Na, M+H+Na, M+K	3	
0 0.50 479 6654n		_		M+H-H ₂ O, M+Na, M+K	3	
0 9.47 No t	ags to assign			M+H, M+Na, M+K	3	
0 0.77 New	tag		M+H-2H2O, M+H-H2O, M+NH4		3	
 0.46 Quid 	k Tags	•	Anova p-value			
○ 0.46 🔗 Edit	ags		Max	fold change	3	
0 4.45 Add	to Clip Gallery		Minimum CV		3	
o 4.49_630.3470n		_	Not identified		3	
0.46_700.6089n			Not fragmented		3	
○ 9.57_582.4144n			Separated by drift time		3	
0.79_244.0703n			Identified and separated by drift time		3	
0.46 560.6784n			M+NH4, M+Na, M+K		3	

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

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 (≥ 2).

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

On pressing **Create tag** a second tag appears in the table against all the compounds with a Max fold change ≥ 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 appli	ed		Creat	te
Compound	Accepted ID	Tag 💌	Adducts	
O 3.78_824.0656n			M+Na, M+H+Na, M+K	4
○ 5.00_440.1694n			M+NH4, M+Na, M+K	
O 4.06_303.1229n		•	M+H, M+Na, M+K	
○ 10.00_930.6859n			M+NH4, M+Na, M+K	
○ 6.22_310.0805n		•	M+H, M+Na, M+K	
○ 8.20_322.2747n			M+H, M+Na, M+K	
○ 1.65_184.0849n			M+H-2H2O, M+H-H2O, M+H	
○ 1.32_153.0789n			M+H-2H ₂ O, M+H-H ₂ O, M+ACN+H	
8.56_398.2443n			M+H, M+Na, M+K	
○ 6.22_308.0856n			M+H, M+Na, M+K	
○ 9.57_582.4144n			M+NH4, M+Na, M+K	
O 3.97_604.2587n			M+NH₄, M+Na, M+K	

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.

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

Tag filter appl compounds ma			Edit
Compound	Accepted ID	Tag 💌	Adducts
• 7.13_276.1366n		۵	M+H-2H ₂ O, M+H-H
O 6.50_270.1040n		۵	M+H, M+Na, M+K 🗉
O 8.21_322.1218n		۲	M+H, M+Na, M+K
O 6.22_310.0805n		۲	M+H, M+Na, M+K
0 8.01_322.1212n			M+H, M+Na, M+K
○ 7.13_308.1991n		۲	M+H, M+Na, M+K
0 8.20_322.2747n		۵	M+H, M+Na, M+K
O 6.22_308.0856n		۲	M+H, M+Na, M+K
○ 7.11_308.1069n		۲	M+H, M+Na, M+K
0 4.45_278.1274n		۲	M+Na, M+K
0 8.02_322.2790n		۲	M+Na, M+K

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.

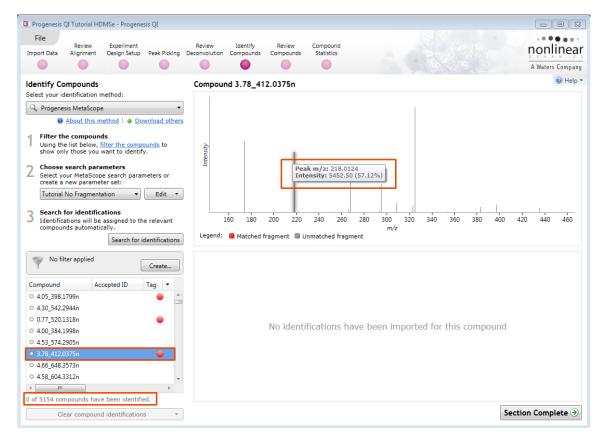
To display all the compounds in the table click **Edit** and then click on **Clear the filter** and then **OK**.

To Find **Compound 3.78_412.0375n**, which has a neutral mass of 412.0375, 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).

Tag filter applied compounds may be hidden	Edit
Search for identifications	1
Create	legend:
pted ID Tag 💌 1	

1								Create
Compound	Accepted ID	Tag	 Identifications 	Fragmented?	Neutral mass	m/z	Drift time	Adduct count
○ 4.05_398.1799n			0	Yes	398.1799	415.2137	3.83	4
○ 4.30_542.2944n			0	Yes	54; m/z	6 H		
○ 0.77_520.1318n			0	Yes	52 ion. To ca	alculate this	s, each ion	compound 's abundance
○ 4.00_384.1998n			0	Yes		l as the sur		rmalized all samples.
O 4.53_574.2905n			0	Yes	574.2905	592.3243		4
• 3.78_412.0375n		- 0		Yes	412.0375	413.0448	3.73	
○ 4.66_648.3573n			0	Yes	648.3573	666.3911	6.21	4
○ 4.58_604.3312n			0	Yes	604.3312	622.3650	5.78	4
○ 3.80_582.1334n			0	Yes	582.1334	583.1522	5.40	4
○ 4.19_498.2686n			0	Yes	498.2686	516.3024	4.75	4
○ 4.40_586.3214n			0	Yes	586.3214	604.3553	5.56	4
○ 4.53_590.3537n			0	Yes	590.3537	608.3875	5.67	4

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 5154 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.

Q Progenesis QI Tutorial HDMSe - Progenesis QI			- • •
File Review Experiment Import Data Alignment Design Setup Peak Picking	Review Identify g Deconvolution Compour		nonlinear
		QI Edit Search Parameters	A Waters Company
Identify Compounds Select your identification method: Progenesis MetaScope	Compound 3.78		
		Fragment tolerance: 12 ppm Perform fragment database search <no database="" selected=""> Fragment tolerance: 12 ppm Save search parameters Cancel</no>	pound
0 of 5154 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\

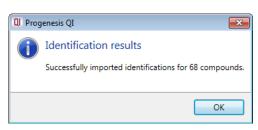
Select: Tutorial Fragmentation.sdf (if not already selected)

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.



No filter applied Create... Tag 💌 Ider Compound Accepted ID O 3.95_370.2213n 0 • 3 🧉 Anova p-value ≤ 0.05 0 0 0 Max fold change ≥ 2 o 4 💽 0 Not identified 0 0 New tag.. 0 Quick Tags Anova p-value.. 0 Max fold change.. 04 🚰 Edit tags Minimum CV... Add to Clip Gallery.. 04 Not identified o 9.73_402.2998n Not fragmented 0 4.22_225.9611n Separated by drift time Identified and separated by drift time 58 of 5154 compounds have been ide Clear all compound identifications

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.

Y	No filter applied	Create

Click **OK** to apply the filter.

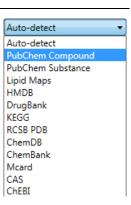
Having imported the search results you can see there are a number of possible identifications for the identified compounds.

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.

Q Filter the compounds Create a filter Show or hide compounds based on a selection boxes to create the filter. For more guidance,	
Available tags: Anova p-value ≤ 0.05 (1243 compounds) Max fold change ≥ 2 (444 compounds)	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 (5086 compounds)
Clear the filter	OK Cancel

Find **Compound 2.91_151.0638n.** 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 applied			Possi	ible identific	ations: 4						
compounds may be hidden	Edit		쓝	Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation score	Mass error (ppm)
Compound Accepted ID	Tag 💌		\star	46506142	Paracetamol	M+H-H	C ₈ H ₉ NO ₂	2.90	51.3	0	3.30
 7.11_308.1069n 	i log i i i		☆	HMDB02210	2-Phenylglycine	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.30
 6.22_308.0856n 	- I		☆	HMDB12219	Dopamine quinone	M+H-H	$C_8H_9NO_2$		38.5	0	3.30
 6.50_270.1040n 			숡	HMDB29703	Methyl 2-aminobenzoate	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.30
8.54_382.2719n											
2.96_314.1436n											
• 2.91_151.0638n	٠										
3.63_179.0946n											
5.18_179.0949n											
< III	÷										
All 68 filtered compounds have beer	n identified.		-								•



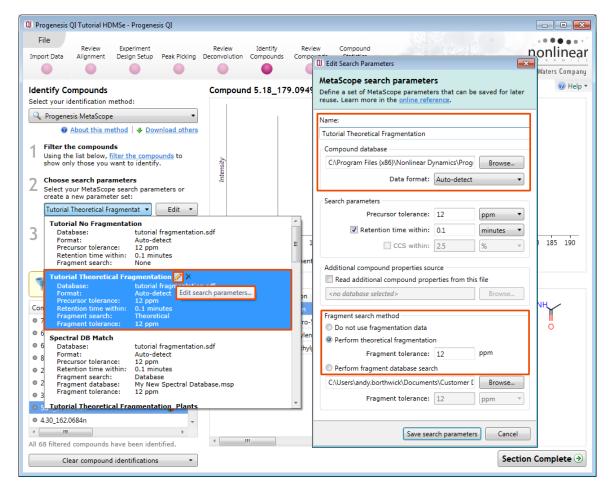
Note: when the Retention time data is **not** included as part of the search scoring then there is no difference in the scoring for the 4 identification candidates in the table to the right.

🖘 Taq filter applied				Possible identifications: 4								
compounds may be l	hidden	Edit	1	Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation score	Mass error (ppm)	
Compound Acce	pted ID	Tag 💌	í	* 46506142	Paracetamol	M+H-H	C ₈ H ₉ NO ₂	2.90	38.5	0	3.30	
• 7.11 308.1069n	piculo	Tug A		HMDB02210	2-Phenylglycine	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.30	
 6.22 308.0856n 		- 👗 🛙	3	HMDB12219	Dopamine quinone	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.30	HN
 6.50_270.1040n 				HMDB29703	Methyl 2-aminobenzoate	M+H-H	C ₈ H ₉ NO ₂		38.5	0	3.30	
 8.54 382.2719n 		•										
 2.96 314.1436n 		•										Γ Ì
• 2.91_151.0638n		4										
3.63_179.0946n												\sim
• 5.18_179.0949n		🌒 ,										
< III.		÷.										ОН
All 68 filtered compounds h	nave been id	dentified.	-	<							÷.	011

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.



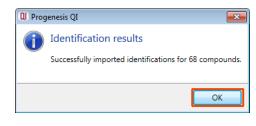
To perform theoretical fragmentation using the **tutorial fragmentation .sdf** as the source of the candidate compounds select the second Search Parameter Set (Tutorial Theoretical Fragmentation) from the drop down

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.

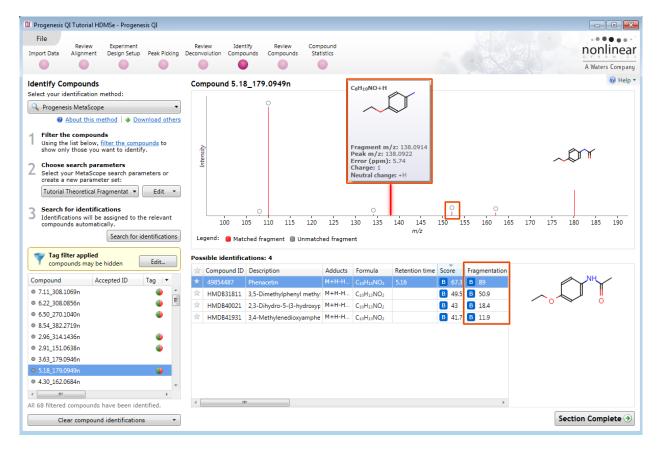
To perform the Theoretical Fragmentation search click on Search for identifications

As before search results will be returned

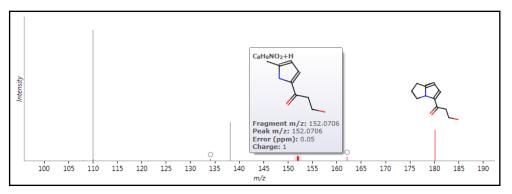


0	Progenesis MetaScope
	🛿 About this method 🖊 Download other
1	Filter the compounds
1	Using the list below, <u>filter the compounds</u> to show only those you want to identify.
2	Choose search parameters Select your MetaScope search parameters or create a new parameter set:
	Tutorial Theoretical Fragmentat 🔹 🛛 Edit 🔹
2	Search for identifications
С	Identifications will be assigned to the relevant compounds automatically.
	Search for identifications

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_179.0949n**).



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 **3rd ranked compound** Fragmentation matches (see below).



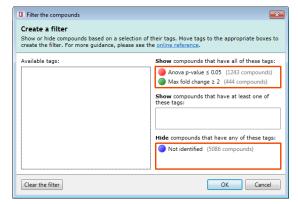
Note: a fragmentation score of 0 indicates either that **no** match was achieved between the theoretical fragments (of a candidate compound) and the measured fragments or in fact the compound has no fragmentation data. For example, in **Compound 2.58_169.0965m/z**:

Q Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment Import Data Alignment Design Setup Peak Picking	Review Identify Review Compound Deconvolution Compounds Statistics	nonlinear A Waters Company
Identify Compounds	Compound 2.58_169.0965m/z	Ø Help ▼
Select your identification method:		
Progenesis MetaScope		
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.		
2 Choose search parameters Select your MetaScope search parameters or create a new parameter set: Tutorial Theoretical Fragment.	This compound has no fragmentation data	
3 Search for identifications Identifications will be assigned to the relevant compounds automatically. Search for identifications		
Search for Identifications	Possible identifications: 3	
Tag filter applied	☆ Compound ID Description Adducts Formula Retention time Score Fragmentation score	он
compounds may be hidden Edit	* HMDB12219 Dopamine quinone M+NH4 C&H4NO2 A 37.1 A 0 * HMDB29703 Methol 2-aminohenzoate M+NH4 C&H4NO2 A 37.1 A 0	
Compound Accepted ID Tag 💌	☆ HMDB29703 Methyl 2-aminobenzoate M+NH4 C ₆ H ₉ NO ₂ A 37.1 A 0 ★ HMDB02210 2-Phenylalycine M+NH4 C ₆ H ₉ NO ₂ A 37.1 A 0	H,N
• 7.91_422.1445m/z		γ
 7.61_346.1201m/z 2.58 169.0965m/z 		
• 2.58_169.0965m/2		
© 2.44_144.0803m/z		
• 2.46_315.1496m/z		
4 <u> </u>		\checkmark
All 68 filtered compounds have been identified.	< H	
Clear compound identifications •		Section Complete →

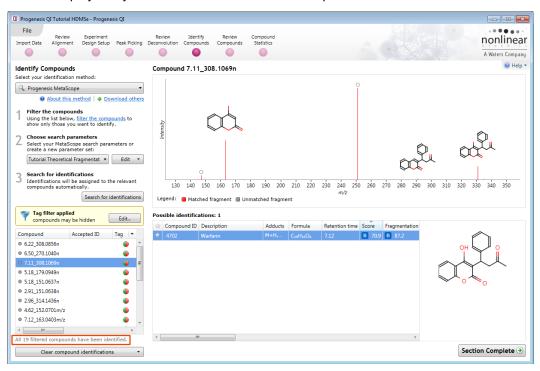
Note: where multiple searches have been performed the search yielding the highest overall score for each candidate is displayed. Also the search parameters giving rise to these scores are displayed as tooltips when the cursor is held over the Search Parameter indicator.

QI Progenesis QI Tutorial HDMSe - Progenesis QI			
File Review Experiment Import Data Alignment Design Setup Peak Picking	Review Identify Review Comp Deconvolution Compounds Compounds Statis		nlinear ters Company
Identify Compounds	Compound 2.91_151.0638n		Help •
Select your identification method:		0	
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.	Intensity		
2 Choose search parameters Select your MetaScope search parameters or create a new parameter set:	4		
Tutorial Theoretical Fragmentat 🔹 Edit 🔹			
3 Search for identifications Identifications will be assigned to the relevant compounds automatically. Search for identifications	105.5 106 106.5 107 107.5 Legend: Matched fragment Unmatche Possible identifications: 4	m/z	114.5 115
Tag filter applied	☆ Compound ID Description	Adducts Formula Retention Score Fragmentation	
Compounds may be hidden Edit	* 46506142 Paracetamol	M+H-H C ₈ H ₉ NO ₂ 2.90 B 69.7 B 92	
Compound Accepted ID Tag 💌	* HMDB12219 Dopamine quinone	M+H-H C ₀ H ₉ NO ₂ B 56.3 B 89	
© 2.96_314.1436n	* HMDB02210 2-Phenylglycine	M+H-H C ₆ H ₉ NO ₂ A 38.5 A Search configuration: B	
• 2.91_151.0638n	🖄 HMDB29703 Methyl 2-aminobenzoate	M+H-H C ₈ H ₉ NO ₂ A 38.5 A Method: Progenesis MetaScope	
• 3.63_179.0946n		Database: tutorial fragmentation.sdf	
• 5.18_179.0949n		Precursor tolerance: 12 ppm Retention time within: 0.1 minutes	
 4.30_162.0684n 4.68_179.0954n 		Theoretical fragment tolerance: 12 ppm	
• 5.18_151.0637n			
< m >			
All 68 filtered compounds have been identified.	<	, OH	
Clear compound identifications •		Section Cor	nplete 🏵

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.

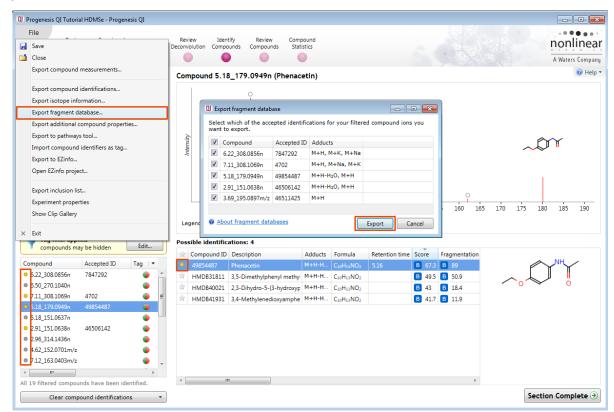


The table now displays only information for 19 identified compounds.



Creating and using databases

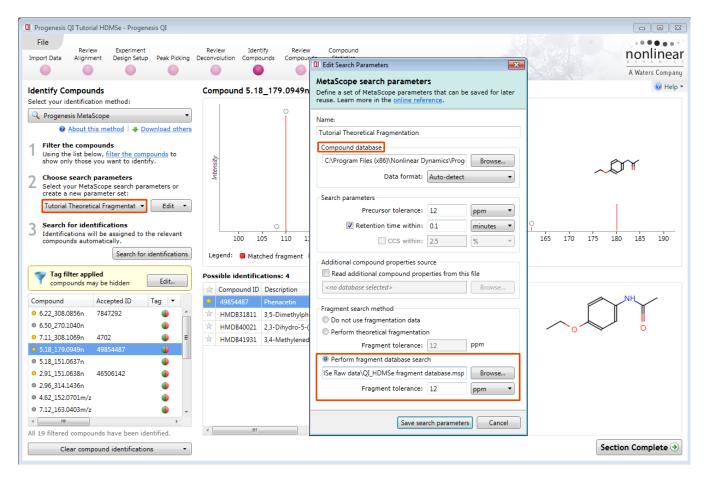
The identification data can be stored and used to perform fragment database searches. This allows you to build up your own (local) databases of fragment information. These can then be used to perform future identifications based on fragment matches when using the same compound database. To create and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



Using the Filter set created for the 19 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.

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.

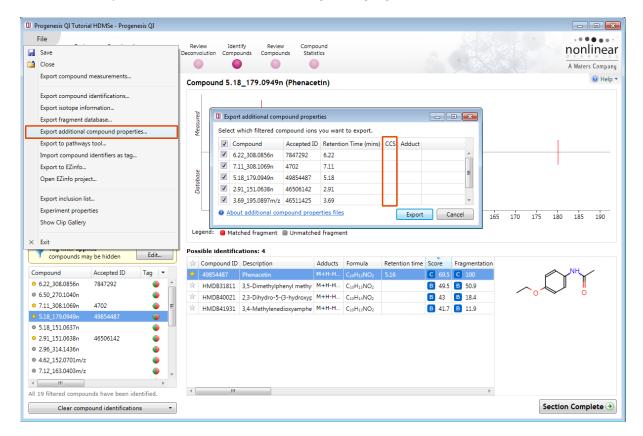
On performing the search you will get a 'mirror plot' where the upper half represents the measured/observed fragmentation and the lower half shows the currently stored fragmentation in the database. This gives rise to a **Fragmentation score** describing how good the match is between the observed fragmentation 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_179.0949n).

	ign Setup Peak Pic	Revie cing Deconvo			npound atistics							A Waters Compar
dentify Compounds		Com	pound 5.18	_179.0949n (Phena	cetin)							🕜 Help
elect your identification metho	d:	_										
🔍 Progenesis MetaScope		•										
About this method	🛛 🦊 Download ot	hers 2										
Filter the compounds Using the list below, filter to show only those you want		hers										
Select your MetaScope sea create a new parameter se Tutorial Theoretical Fragme	t:	abase										
Search for identification Identifications will be assig compounds automatically. Se		t	100 10 end: 🛑 Matc	15 110 115 120 hed fragment Unmate		130 135 nt	140 145 <i>m/z</i>	150 155	160 1	65 170	175 180) 185 190
Identifications will be assig compounds automatically. Se	arch for identificati	t ons Leg		hed fragment 🔳 Unmato				150 155	160 1	65 170	175 180) 185 190
Identifications will be assig compounds automatically. Tag filter applied compounds may be hiddle	arch for identificati	t Dns Leg	end: Matc Matc Model Mo	hed fragment Unmatc ations: 4 Description	ched fragme Adducts	nt Formula	m/z	Score Fr	agmentation	•	175 180	0 185 190
Identifications will be assigned to the set of the set	arch for identificati en Edit d ID Tag I	t Leg	end: Matc Matc Compound ID 49854487	hed fragment Unmatc ations: 4 Description Phenacetin	Adducts	nt Formula C ₁₀ H ₁₃ NO ₂	m/z	Score Fr	agmentation	•	175 180	0 185 190
Identifications will be assig compounds automatically. See Tag filter applied compounds may be hiddi compound Accepte 6.22_308.0856n 7847293	arch for identificati en Edit d ID Tag I	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy	Adducts M+H-H M+H-H	nt Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	m/z	Score Fr C 69.5 C B 49.5 B	agmentation 100 50.9	•	175 180	0 185 190
Identifications will be assigned to a significant of the set of th	arch for identificati en Edit d ID Tag V	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811	hed fragment Unmatc ations: 4 Description Phenacetin	Adducts M+H-H M+H-H	nt Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	175 180) 185 190
Identifications will be assigned to the solution of the solut	ned to the relevar arch for identificati en Edit d ID Tag V 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B	agmentation 100 50.9 18.4	•	0	
Identifications will be assigned in the solution of the solut	en Edit d ID Tag v 2 v 37 v	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy 2,3-Dihydro-5-(3-hydroxy)	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	0	
Identifications will be assigned to a support of the section of	ned to the relevar arch for identificati en Edit d ID Tag V 2 67 67	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy 2,3-Dihydro-5-(3-hydroxy)	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	0	
Identifications will be assigned to a support of the second submitted by the second submitted by the support of the su	ned to the relevar arch for identificati en Edit d ID Tag V 2 67 67	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy 2,3-Dihydro-5-(3-hydroxy)	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	175 180	
Identifications will be assigned in the series compounds automatically. Tag filter applied compounds may be hildl. Compound Accepte 6.50_270.1040n 7.11_308.1069n 4702 7.11_308.1069n 4702 5.18_179.0949n 4985443 5.18_151.0637n 2.96_314.1436n 450614	ned to the relevar arch for identificati en Edit d ID Tag V 2 67 67	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy 2,3-Dihydro-5-(3-hydroxy)	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	175 180	0 185 190
Identifications will be assigned in the solution of the	ned to the relevar arch for identificati en Edit d ID Tag V 2 67 67	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy 2,3-Dihydro-5-(3-hydroxy)	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	175 180	
Identifications will be assigned in the series of the compounds automatically. Tag filter applied compounds may be hidd Compound Accepte 6.22_30.80856n 7847292 6.50_270.1040n 97.11_308.1069n 4702 5.18_179.0949n 4995443 5.18_151.0637n 4650614 2.96_314.1436n 4650614 9.65_13.1436n 4650614 9.63_14.1436n 462_152.0701m/z	en Edit arch for identificati en Edit d ID Tag • 2 • • • • • • • • • • • •	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy 2,3-Dihydro-5-(3-hydroxy)	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	175 180	
Identifications will be assigned in the series of the compounds automatically. Tag filter applied compounds may be hidd Compound Accepte 6.50_270.1040n 6.50_270.1040n 7.11_308.1069n 4702 5.18_179.0949n 4995443 5.18_151.0637n 4650614 2.96_131.4386n 4650614 9.66_2.1.4386n 4650614	arch for identificati en Edit d ID Tag • 2 • 42 • 42 • 42 • 42 • 42 • 42 • 42 • 44 • 45 • 46 • 47 • 47 • 48	t Leg	end: Matc ble identifica Compound ID 49854487 HMDB31811 HMDB40021	hed fragment Unmatc ations: 4 Description Phenacetin 3,5-Dimethylphenyl methy 2,3-Dihydro-5-(3-hydroxy)	Adducts M+H-H y M+H-H p M+H-H	nt Formula C10H11NO2 C10H13NO2 C10H13NO2	m/z	Score Fr C 69.5 C B 49.5 B B 43 B	agmentation 100 50.9 18.4	•	0	

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
$\bullet \bullet \bullet \bullet$		A Waters Company
Identify Compounds Select your identification method:	Compound 5.18_179.0949n	
Progenesis MetaScope		
About this method	ree d	
1 Filter the compounds Using the list below, filter the compounds to show only those you want to identify.	Weaturd	
2 Choose search parameters Select your MetaScope search parameters or create a new parameter set: Tutorial Theoretical Fragmentat V Edit V	Database	
3 Search for identifications Identifications will be assigned to the relevant compounds automatically. Search for identifications	100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 Legend: Matched fragment Unmatched fragment Possible identifications: 4	175 180 185 190
compounds may be hidden Edit	Compound ID Description Adducts Formula Retention time Score Fragmentation	
Compound Accepted ID Tag 🔻	★ 49854487 Phenacetin M+H-H C ₁₀ H ₁₃ NO ₂ 5.16 C 69.5 C 100	
• 6.22_308.0856n	☆ HMDB31811 3,5-Dimethylphenyl methyl M+H-H C ₁₀ H ₁₃ NO ₂	
• 6.50_270.1040n	☆ HMDB40021 2,3-Dihydro-5-(3-hydroxyp M+H-H C ₁₀ H ₁₃ NO ₂	0. 🔨 0
● 7.11_308.1069n 🕘 🗉	☆ HMDB41931 3,4-Methylenedioxyamphe M+H-H C ₁₀ H ₁₃ NO ₂ B 41.7 B 11.9	
• 5.18_179.0949n		
• 5.18_151.0637n		
• 2.91_151.0638n		
Clear identifications from search A (Progenesis Me		
Clear identifications from search B (Progenesis Me		
Clear identifications from search C (Progenesis Me	ascope)	
× Clear all identifications	n >	
Clear compound identifications -		Section Complete 🏵

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

Q Prog	enesis QI
X	Clear identifications for given search
	Clearing all compounds' identifications from search B (Method: Progenesis MetaScope. Database: tutorial fragmentation.sdf. Precursor tolerance: 12 ppm. Retention time within: 0.1 minutes. Theoretical fragment tolerance: 12 ppm.) will clear all identifications that came from that search, both possible and accepted.
	Clear identifications Cancel

Following deleting the identifications based on Search configuration B there will only be one possible remaining identification for **Compound 5.18_179.0949n**

	1ay be hidden			Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation	
ompound	Accepted ID	Tag 💌		49854487	Phenacetin	M+H-H	C10H13NO2	5.16	C 69.5	C 100	
6.22_308.0856n	7847292	4	*								
6.50_270.1040n		4									0 0
7.11_308.1069n	4702	4	Ξ								
5.18_179.0949n	49854487	4									
5.18_151.0637n		4									
2.91_151.0638n	46506142	4									
2.96_314.1436n		4									
Clear identifica Clear identifica	tions from search tions from search	-									
Clear all identif	ications										

Now clear all the compound identifications and re-perform the **Tutorial No Fragmentation** search as described on page 40.

To make use of the **Additional compound properties file**, you created on page 47, 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.

Q Progenesis QI Tutorial HDMSe - Progenesis QI		
File Import Data Review Experiment Alignment Design Setup Peak Picking Der	Re U Edit Search Parameters MetaScope search parameters Define a set of MetaScope parameters that can be saved for later reuse. Learn more in the online reference.	A Waters Company
Identify Compounds Select your identification method: Select your identification method: Progenesis MetaScope About this method Pilter the compounds Using the list below, filter the compounds to show only those you want to identify. Select your MetaScope search parameters or create a new parameter set: Tutorial No Fragmentation Search for identifications Identifications will be assumed to the relevant	Name: Tutoal No Fragmentation Corpound database CAProgram Files (x86)Nonlinear Dynamics/Prog Browse Data format: Auto-detect Search parameters Precursor tolerance: I2 Retention time within: 0.1 minutes ▼ CCS within: 2.5	W Help *
compounds automatically. Search for identifications Tag filter applied compounds may be hidden Edit.	Additional compound properties source Image: Compound properties from this file Image: Compound Properties from this file	145 150 155 160 165 170 175 180 185 190 m/z
Compound Accepted ID Tag	Pragment search method O not use fragmentation data Perform theoretical fragmentation Fragment tolerance: 12 ppm Perform fragment database search ro database selected> Fragment tolerance: 12 ppm Save search parameters Cancel	$ \begin{array}{c c c c c c c c c c c c c c c c c c c $
Clear compound identifications -		Section Complete 🕥

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

QI Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment Design Setup Peak Picking I	Review Identify Review Compound econvolution Compounds Compounds Statistics	A Waters Company
Identify Compounds	Compound 5.18_179.0949n	@ Help
Select your identification method:		
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.	htensity	
2 Choose search parameters Select your MetaScope search parameters or create a new parameter set:	104	
Tutorial No Fragmentation 🔹 Edit 👻		
3 Search for identifications Identifications will be assigned to the relevant compounds automatically. Search for identifications Tag filter applied compounds may be hidden	100 105 110 115 120 125 130 135 140 145 150 155 160 Legend: ■ Matched fragment ■ Unmatched fragment Possible identifications: 4	165 170 175 180 185 190
	The Compound ID Description Adducts Formula Retention time Score Fragment	itation
Compound Accepted ID Tag 💌	★ 49854487 Phenacetin M+H-H C ₁₀ H ₁₃ NO ₂ 5.18 B 59.2 B 0	
© 6.22_308.0856n	Image: marked black with the second	
• 6.50_270.1040n	Image: Microsoft Micros	
	Image: Microsoft Action Microsoft	
5.18_179.0949n 5.18_151.0637n		
2.91 151.0638n		
• 2.96 314.1436n		
• 4.62 152.0701m/z		
• 7.12_163.0403m/z		
< >		
All 19 filtered compounds have been identified.	<	Þ
Clear compound identifications		Section Complete $ i$

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 v2.1 has identification methods for Elemental Composition, LipidBlast, ChemSpider and NIST. These methods are installed by default and are available at the Identify Compounds stage from the identification method drop down. QI Edit Search Parameters

As an example of using Elemental Composition: first Clear all compound identifications that you carried out in the previous section, and leave the current Tag Filter applied.

Then with MetaScope still selected as the current identification method return the Tutorial Theoretical Fragmentation search parameters to the default settings as shown to the right by ticking Perform theoretical fragmentation. Save the search parameters.

Finally re-perform the search for identifications using the Tutorial The toti

heoretical Fragmentation pa	arameter set.							Mass within: 12	ppm *
								Save search parar	meters Cancel
QI Progenesis QI Tutorial HDMSe - Progenesis QI									
File Review Experiment	Review Identify	Review Comp	ound						
		Compounds Stati							nonlinear
									A Waters Company
Identify Compounds	Compound 5.18_179.	0949n							🕢 Help 🔻
Select your identification method:		0							
Progenesis MetaScope		1							
ChemSpider Version: 1.0.5673.29215									
Elemental composition									
Version: 1.0.5673.29222	2				\sim				
LipidBlast Version: 1.0.5673.29212	Intensity			\sim					
Progenesis MetaScope	ri i								~ ~
Version: 1.0.5673.29121 METLIN batch metabolite search									
Version: 1.0.5673.39971									
NIST MS/MS Library Version: 1.0.5673,40082								_	
Identifications will be assigned to the relevant	0			•	I				
compounds automatically.	100 105 1	10 115 120	125	130 135	140 145 <i>m/z</i>	150	155 160	165 170 175	180 185 190
Search for identifications	Legend: 🧧 Matched frag	gment 🔲 Unmatch	ed fragmer	nt	m/z				
Tag filter applied compounds may be hidden	Possible identifications:	4							
Compound Accepted ID Tag 🔻	🚖 Compound ID Descrip	tion	Adducts	Formula	Retention time	Score	Fragmentation		
6.22_308.0856n	* 49854487 Phenace	etin	M+H-H	C ₁₀ H ₁₃ NO ₂	5.16	67.3	89	F	
• 5.18_151.0637n	🖈 HMDB31811 3,5-Dim	ethylphenyl methyl	M+H-H	C10H13NO2		49.5	50.9		
• 2.96_314.1436n	* HMDB40021 2,3-Dih	ydro-5-(3-hydroxyp	M+H-H	C10H13NO2		43	18.4	~	✓ 0
• 5.18_179.0949n	🖄 HMDB41931 3,4-Met	thylenedioxyamphe	M+H-H	$C_{10}H_{13}NO_2$		41.7	11.9		
• 2.91_151.0638n									
• 3.69_195.0897m/z									
• 7.12_121.0290m/z									
• 7.12_163.0403m/z									
• 4.21_135.0443m/z									
• 4.42_136.0758m/z 🖕 👻									
< +									
All 19 filtered compounds have been identified.	< III						•	_	
Clear compound identifications •									Section Complete $ ightarrow$

Then select Elemental composition as the Identification methods.

QJ

×

•

ppm 🔹

minutes 🔹

Browse...

ppm

%

MetaScope search parameters

Tutorial Theoretical Fragmentation

Compound database

Search parameters

Name:

Define a set of MetaScope parameters that can be saved for later reuse. Learn more in the <u>online reference</u>.

C:\Program Files (x86)\Nonlinear Dynamics\Prog

Mass within: 12

CCS within: 2.5

C:\Users\andy.borthwick\Documents\Customer [Browse...

Read additional compound properties from this file

Retention time within: 0.1

Additional compound properties source

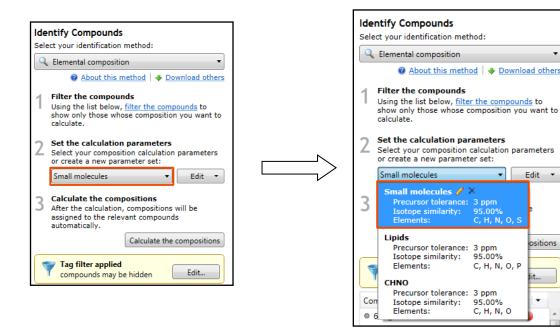
Perform fragment database search

<no database selected>

Fragment search method -Do not use fragmentation data Perform theoretical fragmentation Relative mass error: 12

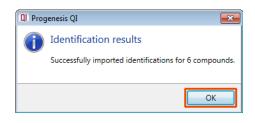
Data format: Auto-detect

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 (%)**.



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

Click OK to import the 6 identifications.



The **Possible identifications** table for the current compound updates to include any compositions that meet the current calculation parameters.

File Review Experiment mport Data Alignment Design Setup Peak Picking	Review Deconvolution C	Identify Review Compounds Compoun				nonlir	near
		•				A Waters Co	mpany
lentify Compounds	Compound	5.18 179.0949n				0	Help 👻
elect your identification method:							
Elemental composition		1					
Ø About this method I Download other	5						
Filter the compounds Using the list below, <u>filter the compounds</u> to show only those whose composition you want to calculate.	Intensity						
Set the calculation parameters Select your composition calculation parameters or create a new parameter set:							
Small molecules Edit							
Calculate the compositions After the calculation, compositions will be	100			135 140 145 15 <i>m/z</i>	0 155 160 165	170 175 180 185 1	90
assigned to the relevant compounds automatically.		Matched fragment	Unmatched fragment				
assigned to the relevant compounds automatically. Calculate the compositions		ntifications: 5	Unmatched fragment				
automatically. Calculate the compositions	Possible ider	-	Unmatched fragment	Adducts	Formula		
automatically. Calculate the compositions Tag filter applied	Possible ider	ntifications: 5 nd ID Description	Unmatched fragment	M+H-H	C ₁₀ H ₁₃ NO ₂		
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit	Possible ider Compour 4985448 HMDB31	ntifications: 5 nd ID Description	-	M+H-H M+H-H	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂		
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit compound Accepted ID Tag	Possible ider ☆ Compour ☆ 4985448 ☆ HMDB31 ☆ HMDB40	ntifications: 5 nd ID Description 7 Phenacetin 1811 3,5-Dimethylphe	-	M+H-H M+H-H yrrolizine M+H-H	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂		
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit onpound Accepted ID Tag	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD840 ☆ HMD840	ntifications: 5 nd ID Description 7 Phenacetin 13,5-Dimethylphe 1021 2,3-Dihydro-5-(3 1931 3,4-Methylenedi	enyl methylcarbamate 3-hydroxypropanoyl)-1H-p ioxyamphetamine	M+H-H M+H-H yyrrolizine M+H-H M+H-H	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂		
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit ompound Accepted ID Tag Calculate the compositions Calculate t	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD840 ☆ HMD840	ntifications: 5 nd ID Description 7 Phenacetin 13,5-Dimethylphe 1021 2,3-Dihydro-5-(3 1931 3,4-Methylenedi	enyl methylcarbamate 3-hydroxypropanoyl)-1H-p ioxyamphetamine	M+H-H M+H-H yrrolizine M+H-H	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	No structure data	
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit ompound Accepted ID Tag	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD840 ☆ HMD840	ntifications: 5 nd ID Description 7 Phenacetin 13,5-Dimethylphe 1021 2,3-Dihydro-5-(3 1931 3,4-Methylenedi	enyl methylcarbamate 3-hydroxypropanoyl)-1H-p ioxyamphetamine	M+H-H M+H-H yyrrolizine M+H-H M+H-H	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	No structure data	
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit ompound Accepted ID Tag	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD840 ☆ HMD840	ntifications: 5 nd ID Description 7 Phenacetin 13,5-Dimethylphe 1021 2,3-Dihydro-5-(3 1931 3,4-Methylenedi	anyl methylcarbamate 3-hydroxypropanoyl)-1H-p ioxyamphetamine 1NO2 determined by elem	M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	No structure data	
automatically. Calculate the compositions Tag filter applied compounds may be hidden 6.22 308.0856n 6.50.270.1040n 7.11.308.1069n 5.518.179.0949n 5.518.151.0637n	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD840 ☆ HMD840	ntifications: 5 nd ID Description 7 Phenacetin 13,5-Dimethylphe 1021 2,3-Dihydro-5-(3 1931 3,4-Methylenedi	anyl methylcarbamate 3-hydroxypropanoyl)-1H-p ioxyamphetamine 1NO2 determined by elem	M+H-H M+H-H M+H-H M+H-H M+H-H ental composition M+H-H identifications: 5	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂		
automatically. Calculate the compositions Tag fitter applied compounds may be hidden Edit ompound Accepted ID Tag 6.50,270.1040n 6.50,270.1040n	Possible ider ☆ Compour ☆ 4985448 ☆ HMDB31 ☆ HMDB40 ☆ HMDB41	ntifications: 5 nd ID Description 7 Phenacetin 13,5-Dimethylphe 1021 2,3-Dihydro-5-(3 1931 3,4-Methylenedi	anyl methylcarbamate 3-hydroxypropanoyl)-1H-p ioxyamphetamine 1NO2 determined by elem	M+H-H M+H-H M+H-H M+H-H M+H-H ental composition M+H-H identifications: 5	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	No structure data Retention time error (mins) Isotope simila
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit ompound Accepted ID Tag Calculate the compositions Calculate the com	Possible ider ☆ Compour ☆ 4985448 ☆ HMDB31 ☆ HMDB40 ☆ HMDB41	ntifications: 5 nd ID Description 7 Phenacetin 13,5-Dimethylphe 1021 2,3-Dihydro-5-(3 1931 3,4-Methylenedi	enyl methylcarbamate 3-hydroxypropanoyl)-1H-p ioxyamphetamine INO2 determined by elem	M+H-H M+H-H M+H-H M+H-H M+H-H ental composition M+H-H identifications: 5	C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂) Isotope simil 98.29
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit ompound Accepted ID Tag Calculate the compositions Calculate the com	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD842 ☆ HMD844 ☆ HMD844 ★ C10H181	httifications: 5 hd ID Description 7 Phenacetin 13,5-Dimethylph 1021 2,3-Dihydro-5-(2) 23,3-Al-Methylenedi 1021 2,3-Dihydro-5-(2) 231 3,4-Methylenedi 1022 Formula C10H13	anyl methylcarbamate 3-hydroxypropanoyl)-1H-rg ioxyamphetamine INO2 determined by elem Score	M+H-H yrrolizine M+H-H M+H-H m+H-H identifications: 5 Fragmentation score	C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	Retention time error (mins	
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit ompound Accepted ID Tag Calculate the compositions Calculate th	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD842 ☆ HMD844 ☆ HMD844 ★ C10H181	httifications: 5 hd ID Description 7 Phenacetin 13,5-Dimethylph 1021 2,3-Dihydro-5-(2) 23,3-Al-Methylenedi 1021 2,3-Dihydro-5-(2) 231 3,4-Methylenedi 1022 Formula C10H13	anyl methylcarbamate 3-hydroxypropanoyl)-1H-rg ioxyamphetamine INO2 determined by elem Score	M+H-H M+H M+	CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2	Retention time error (mins	98.29
automatically. Calculate the compositions Tag filter applied compounds may be hidden Edit Compound Accepted ID Tag Calculate the compositions Accepted ID Tag Calculate the compositions Calculate the composit	Possible ider ☆ Compour ☆ 4985448 ☆ HMD831 ☆ HMD842 ☆ HMD844 ☆ HMD844 ★ C10H181	httifications: 5 hd ID Description 7 Phenacetin 13,5-Dimethylph 1021 2,3-Dihydro-5-(2) 23,3-Al-Methylenedi 1021 2,3-Dihydro-5-(2) 231 3,4-Methylenedi 1022 Formula C10H13	anyl methylcarbamate 3-hydroxypropanoyl)-1H-rg ioxyamphetamine INO2 determined by elem Score	M+H-H M+H-H yrrolizine M+H-H h+H-H ental composition M+H+H identifications: 5 Fragmentation score A 89 A 50.9	CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2 CasHaNO2	Retention time error (mins	98.29 98.29 98.29

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:	paramete	rs
	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.

Q Edit ele	mental co	mpositio	n calculation	parame	ters																×
			n calculati				s that	can be	saved	for lat	er reu	se. Lea	rn moi	re in th	ne <u>onlir</u>	ne refe	rence.				
Name: Small mo	olecules (2))		н																	He
Composit	ion:			Li	Be											В	С	N	0	F	Ne
Element		То	Remove?	Na	Mg											AI	Si	D	s	СІ	Ar
н	0	50	×	ING	ivig											~	31	ſ	3	G	Ai
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	T	Xe
0	0	20	×	KD	or	T	21	IND	IVIO	IC.	ĸu	- NN	Pu	Ag	Ca	IU	on	30	Te	1	ve
S	0	2	×	Cs	Ba		Hf	Та	w	Re	Os	Ir	Pt	Au	Hg	TI	Pb	Bi	Ро	At	Rn
				Fr	Ra		Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	FI	Uup	Lv	Uus	Uuo
				1		La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Но	Er	Tm	Yb	Lu	
Tolerance Precursor	tolerance:	3	ppm 🔻			Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	
Isotope si	imilarity:	95	%																		
																		Save		Car	cel

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.

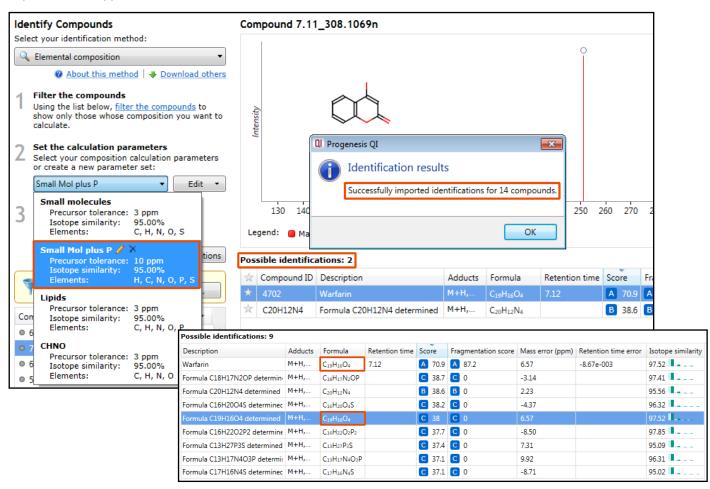
Q Edit eler	mental c	ompositio	on calculation	parame	ters																×
Elemen	tal con	npositio	on calculati	on pa	rame	ters															
Define a s	set of ele	emental o	composition ca	lculatio	on para	meter	s that	can be	saved	for lat	er reu	se. Lea	irn moi	re in th	e <u>onlir</u>	ne refe	rence.				
Name:				н																	He
Small Mo	ol plus P			"																	ne
Composit	ion:			Li	Be											В	С	N	0	F	Ne
Element	From	То	Remove?	Na	Ma											AI	Si	Р	s	CI	Ar
н	0	50	×	ina	Mg											AI	21	P	3	G	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	v	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	т	Xe
0	0	20	×	KD	JI JI	T	21	IND	IVIO	IC.	ĸu	ΝΠ.	Pu	Ag	Ca	m	on	30	Te	1	ve
Ρ	0	4	×	Cs	Ba		Hf	Та	w	Re	Os	Ir	Pt	Au	Hg	TI	Pb	Bi	Ро	At	Rn
S	0	2	×	Fr	Ra		Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	FI	Uup	Lv	Uus	Uuo
					Thu -		- Ni	00	59	bii	113	inte	03	ing	Cir	out		oup		003	000
Tolerance						La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Но	Er	Tm	Yb	Lu	
Precursor		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_308.1069n current in the table and click **Calculate the composition**.

When the compositions are calculated using the adjusted composition and tolerances an additional 7 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.

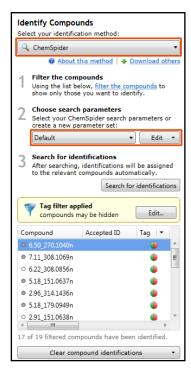
Running a ChemSpider Search

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 608 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.

🛛 Edit ChemSpider Search	Parameters	×
ChemSpider search Define a set of ChemSpir reuse. Learn more in the	der parameters tha	at can be saved for later
Name: Default		
Required search parame Precursor tolerance:	5	ppm •
Data source:	KEGG.	Select data sources
Optional search parame		
Fragment tolerance:	5	ppm
Filter by isotope similar	ilarity score	_
Isotope similarity:	95	%
Filter by elemental control control control control composition		50 N: 0-10 O: 0-30
	Select el	emental composition
		Save Cancel

on the left.	n from	a wide variety of o	ata sources. These are listed in the panel urces most appropriate to your analysis.
	Q]	
Data sources to choose from:			Data sources to use for searching:
4C Pharma Scientific A&J Pharmtech A1 BioChem Labs A2Z Chemical Abacipharm Abblis Chemicals Abbam ABI Chemicals Abmole Bioscience ACB Blocks Accela ChemBio ACD Labs ACD Labs Accemol Acemtex Scientific	*	Add > Add all >> < Remove < Remove all	KEGG 1 of 608 data sources chosen for searching

I 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_179.0949n	🛞 Help 🔻
Select your identification method:		
🔍 ChemSpider 🔹		
About this method + Download others		
Filter the compounds Using the list below, filter the compounds to show only those you want to identify.	Intensity	
2 Choose search parameters Select your ChemSpider search parameters or create a new parameter set:	<u>u</u>	
Default		
3 Search for identifications After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications	Searching for identifications (77/103 masses)	Cancel
Tag filter applied compounds may be hidden Edit		Identification results
Compound Accepted ID Tag 💌		Successfully imported identifications for 14 compounds.
o 6.50_270.1040n		
• 7.11_308.1069n		ОК
○ 6.22_308.0856n ○ 5.18 151.0637n	No identifications have been	n imported for this compound
o 2.96 314.1436n		
• 5.18_179.0949n		
○ 2.91_151.0638n		
○ 3.69_195.0897m/z		
0 of 19 filtered compounds have been identified.		
Clear compound identifications 👻		Section Complete 🏵

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_179.0949**, based on a mass error tolerance of 5ppm there are 9 other possibilities with the same mass error.

File Review Experimen nport Data Alignment Design Setu		Revi Deconv											nonli	nea
	-	-											A Waters C	ompai
entify Compounds		Com	npound 5.18	179.0949n									Q) Help
elect your identification method:				-										
🔍 ChemSpider	•			1										
🛛 About this method 🛛 🐥 🛛	Download others													
Filter the compounds Using the list below, <u>filter the cor</u> show only those you want to ider		Intensity	forei a											
Choose search parameters Select your ChemSpider search p create a new parameter set:		Inte										I		
Default • Search for identifications After searching, identifications wi			100	105 110 115	120	125 120	125	140 145	150 155	160 165	170 175	180	105 1	1
Search for identifications After searching, identifications wi to the relevant compounds auton Search fo	ill be assigned			105 110 115 hed fragment Unmatche		125 130 t	135	140 145 m/z	150 155	160 165	170 175	180	185 1	90
Search for identifications After searching, identifications wi to the relevant compounds autom	ill be assigned natically.	Poss	gend: 🛛 Matcl	hed fragment Unmatche				m/z				180	185 1	90
Search for identifications After searching, identifications wi to the relevant compounds auton Search for Tag filter applied compounds may be hidden	ill be assigned natically. rr identifications	Poss	gend: 🛑 Match	hed fragment Unmatche tions: 10 Description	ed fragmen Adducts	t	Rete Score	m/z	150 155 Mass error (ppm) 1 1.37			180	185 1	90
Search for identifications After searching, identifications wi to the relevant compounds autor Search for Tag filter applied compounds may be hidden mpound Accepted ID	ill be assigned natically. or identifications	Poss	gend: 🔲 Match sible identifica Compound ID	hed fragment Unmatche tions: 10 Description	Adducts	t Formula	Rete Score 39.3	m/z Fragmentatior	Mass error (ppm)	Ret Isotope similarity		180	185 1 NH	90
Search for identifications After searching, identifications wit to the relevant compounds autor Search for Carg filter applied compounds may be hidden mpound Accepted ID 6.50_270.1040n	ill be assigned natically. rr identifications	Poss	gend: Matcl sible identifica Compound ID CSID16606	tions: 10 Description 3,5-Dimethylphenyl methyl	Adducts	t Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	Rete Score 39.3 39.3	m/z Fragmentatior 0	Mass error (ppm) 4 1.37	Ret Isotope similarity 98.29		180	185 1	90
Search for identifications After searching, identifications wi to the relevant compounds autor Search for Tag filter applied compounds may be hidden mpound Accepted ID 6.50,270.1040n 7.11_308.1069n	ill be assigned natically. rr identifications	Poss	gend: Matcl sible identifica Compound ID CSID16606 CSID16141	tions: 10 Description 3,5-Dimethylphenyl methy 3,4-Dimethylphenyl methyl	Adducts M+H-H M+H-H	t Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	Rete Score 39.3 39.3 39.3	m/z Fragmentation 0	Mass error (ppm) 1 1.37 1.37	Ret Isotope similarity 98.29 98.29		180	185 1	90
Search for identifications After searching, identifications wi to the relevant compounds autor Search for Compounds may be hidden mpound Accepted ID 6.50,270.1040n 7.11_308.1069n 6.22_308.0856n	ill be assigned natically. rr identifications Edit Tag V	Poss ☆ ☆ ☆	gend: Matcl sible identifica Compound ID CSID16606 CSID16141 CSID23083	tions: 10 Description 3,5-Dimethylphenyl methy 3,4-Dimethylphenyl methy IPC	Adducts M+H-H M+H-H M+H-H	t Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	Rete Score 39.3 39.3 39.3	m/z Fragmentatior 0 0 0	Mass error (ppm) 1 1.37 1.37 1.37	Set Isotope similarity 98.29 98.29 98.29 98.29		180	185 1	90
Search for identifications After searching, identifications wit the relevant compounds autor Search for Compounds may be hidden mpound Accepted ID 650,270.1040n 7.11_308.1069n 6.22_308.0856n 5.18_151.0637n	ill be assigned natically. rr identifications Edit Tag V		gend: Match sible identifica Compound ID CSID16606 CSID16141 CSID23083 CSID2006639	tions: 10 Description 3,5-Dimethylphenyl methy 3,4-Dimethylphenyl methy IPC (5)-2-amino-4-phenylbutai	Adducts M+H-H M+H-H M+H-H M+H-H	t Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	Rete Score 39.3 39.3 39.3 39.3	m/z Fragmentatior 0 0 0 0 0 0 0 0	Mass error (ppm) 1 1.37 1.37 1.37 1.37 1.37	Ret Isotope similarity 98.29 98.29 98.29 98.29 98.29 98.29		180		90
Search for identifications After searching, identifications wi to the relevant compounds autor Search for Campounds may be hidden mpound Accepted ID 6.50,270.1040n 7.11,308.1069n 6.22,308.0856n 5.28,151.0637n 2.96,314.1436n	ill be assigned natically. rr identifications Edit Tag V		gend: Matcl sible identifica Compound ID CSID16606 CSID16141 CSID23083 CSID2006639 CSID82699	tions: 10 Description 3,5-Dimethylphenyl methyl 3,4-Dimethylphenyl methyl IPC (S)-2-amino-4-phenylbutar Salsolinol	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H	t Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	Rete Score 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3	m/z Fragmentatior 0 0 0 0 0 0 0 0 0	Mass error (ppm) 1 1.37 1.37 1.37 1.37 1.37 1.37	Set Isotope similarity 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29		180		90
Search for identifications After searching, identifications wi to the relevant compounds auton Search for Tag filter applied compounds may be hidden	ill be assigned natically. rr identifications Edit Tag • E E		gend: Matcl sible identifica Compound ID CSID16606 CSID16141 CSID23083 CSID2006639 CSID82699 CSID4590	tions: 10 Description 3.5-Dimethylphenyl methy 3.4-Dimethylphenyl methy JPC (S)-2-amino-4-phenylbutai Sakolinol Phenacetin	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H	Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	Rete Score 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3	m/z Fragmentatior 0 0 0 0 0 0 0 0 0	Mass error (ppm) 8 1.37 1.37 1.37 1.37 1.37 1.37 1.37	Ret Isotope similarity 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29		180		90
Search for identifications After searching, identifications with the relevant compounds autori Search for Tag filter applied compounds may be hidden mpound Accepted ID 6.50, 270.1040n 7.11,308.1069n 6.22,308.0856n 5.18,151.0637n 2.26,314.1436n 5.18,179.0049n	ill be assigned natically. vr identifications Edit Tag V Egit Edit		gend: Match sible identifica Compound ID CSID16606 CSID16141 CSID23083 CSID2006639 CSID82699 CSID4590 CSID324	tions: 10 Description 3.5-Dimethylphenyl methy JPC (5)-2-amino-4-phenylbutar Salsolinol Phenacetin Fusaria caid Phenprobamate	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H	Formula C10H13NO2	Rete Score 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3	m/z Fragmentatior 0 0 0 0 0 0 0 0 0	Mass error (ppm) 4 1.37 1.37 1.37 1.37 1.37 1.37 1.37 1.37	Ret Isotope similarity 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29		180		90
Search for identifications After searching, identifications wi to the relevant compounds autor Search for Tag filter applied compounds may be hidden mpound Accepted ID 6.50,270.1040 7.11_308.1069n 6.22_308.0856n 5.18_151.0637n 2.96_314.1436n 5.18_179.049n 2.91_151.0638n	ill be assigned natically. vr identifications Edit Tag V Egit Edit		Matcl sible identifica Compound ID CSID16606 CSID16141 CSID23083 CSID2609 CSID2699 CSID2590 CSID324 CSID324	tions: 10 Description 3.5-Dimethylphenyl methy JPC (5)-2-amino-4-phenylbutar Salsolinol Phenacetin Fusaria caid Phenprobamate	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H	Formula C ₁₀ H ₁₃ NO ₂ C ₁₀ H ₁₃ NO ₂	Rete Score 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3 39.3	m/z Fragmentation 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Mass error (ppm) 1 1.37 1.37 1.37 1.37 1.37 1.37 1.37 1.3	Ret Isotope similarity 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29		180		90

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 cutoff. **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.

I Edit ChemSpider Search	Parameters	×						
ChemSpider search Define a set of ChemSpid reuse. Learn more in the	• er parameters that (can be saved for later						
Name:								
Default								
Required search paramet	ers							
Precursor tolerance:	5	ppm •						
Data source:	KEGG.							
	S	elect data sources						
Optional search paramet	0.00							
Perform theoretical fr								
Fragment tolerance:	5	ppm						
Filter by isotope simil	arity score							
Isotope similarity:	95	%						
Filter by elemental co	mposition	,						
Elemental composition:	C: 0-100 H: 0-150	N: 0-10 O: 0-30						
	P: 0-2 S: 0-2							
	Select elemental composition							
	Sa	ave Cancel						

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_179.0949** making Phenacetin to be the more likely compound identification.

File Review Experiment port Data Alignment Design Setup Peak Pickin	Rev ng Deconv			itics					ex.	nonline A Waters Com
entify Compounds lect your identification method:	Con	npound 5.18	_179.0949n							() H
ChemSpider search parameters Chefault Chefault ChemSpider search parameters Chefault Chefault ChemSpider Search parameters ChemSpider Se	Intensity	freesawin	0			~Ĉ	Ţ	0		
Search for identifications After searching, identifications will be assigned to the relevant compounds automatically. Search for identification	_	100 gend: 🛑 Matcl	105 110 115 hed fragment 🔲 Unmatche		25 130 t	0 135 14	0 145 <i>m/z</i>	150 155	160 165 17	70 175 180 185 190
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification	ns Le		105 110 115 hed fragment 🔲 Unmatche			- ī			160 165 17	70 175 180 185 190
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification	ns Le Pose	gend: 📕 Matcl	105 110 115 hed fragment Unmatche ations: 10			- ī	m/z	150 155	160 165 17	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification	ns Le Pose	gend: 😑 Matcl	105 110 115 hed fragment Unmatche ations: 10	ed fragmen Adducts	t	135 14	m/z	150 155	Ret Isotope similar 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag V	ns Le Pose	gend: 🔲 Matcl sible identifica Compound ID	105 110 115 hed fragment Cunmatche ations: 10 Description	Adducts	Formula	135 14 Rete Score	<i>m/z</i> Fragmentati	150 155	Ret Isotope similar	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag • 550,270.1040n •	ns Le Pos: ☆	gend: Matcl sible identifica Compound ID CSID4590	105 110 115 hed fragment C Unmatche ations: 10 Description Phenacetin	Adducts M+H-H	Formula C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50	m/z Fragmentati	150 155 ior Mass error (ppm) 1.37	Ret Isotope similar 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden mpound Accepted ID Tag 550_270.1040n 7.11_308.1069n	ns Le Poss ☆	gend: Matcl sible identifica Compound ID CSID4590 CSID82699	105 110 115 hed fragment II Unmatche stions: 10 Description Phenacetin Salsolinol	Adducts M+H-H M+H-H	Formula C10H13NO2 C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50 B 49.9	m/z Fragmentati B 87 B 53.4	150 155 ior Mass error (ppm) 1.37 1.37	Ret Isotope similar 98.29 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag • 50,270.1040n • 5.22,308.0856n •	ns Lee Poss ☆ ★ ★	gend: Matcl sible identifica Compound ID CSID4590 CSID82699 CSID2006639	105 110 115 hed fragment Unmatche ations: 10 Description Phenacetin Salsolinol (S)-2-amino-4-phenylbutar	Adducts M+H-H M+H-H M+H-H	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50 B 49.9 B 49.5	m/z Fragmentati B 87 B 53.4 B 52.6	ior Mass error (ppm) 1.37 1.37 1.37	Ret Isotope similar 98.29 98.29 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden mpound Accepted ID Tag • 550 270.1040n 0.11.308.1069n 5.18_151.0637n 0.11.1000 1.12.2000.0056n 1.12.2000.0	ns Le Poss ☆ ☆ ☆ ☆	gend: Match sible identifica Compound ID CSID4590 CSID82699 CSID2006639 CSID16141	105 110 115 hed fragment Unmatche tions: 10 Description Phenacetin Salsolinol (S)-2-amino-4-phenylbutar 3.4-Dimethylphenyl methy	Adducts M+H-H M+H-H M+H-H M+H-H	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50 B 49.9 B 49.5 B 49.5	m/z Fragmentati B 87 B 53.4 B 52.6 B 50.9	150 155 ior Mass error (ppm) 1.37 1.37 1.37 1.37	Ret Isotope similar 98.29 98.29 98.29 98.29 98.29 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden mpound Accepted ID Tag • 650,270.1040n Accepted ID Tag • 622,308.065n 622,308.065n 218,151.0637n 296,314.1436n 0	ns Le Poss ☆ ☆ ☆ ☆ ☆	gend: Match sible identifica Compound ID CSID4590 CSID82699 CSID2006639 CSID16141 CSID16606	105 110 115 hed fragment Unmatche stions: 10 Description Phenacetin Salsolinol (5)-2-amino-4-phenylbutar 3,4-Dimethylphenyl methyl 3,5-Dimethylphenyl methyl	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50 B 49.9 B 49.5 B 49.5	m/z Fragmentati B 87 B 53.4 B 52.6 B 50.9 B 50.9	ior Mass error (ppm) 157 137 137 137 137 137 137	Ret Isotope similar 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag 6.50,270.1040n 7.11_308.1069n 6.22_308.0856n 5.18_151.0637n 5.18_151.0637n 5.18_151.0637n 5.18_179.0249n U	ns Le Poss ☆ ☆ ☆ ☆	gend: Match Sible identifica Compound ID CSID4590 CSID2006639 CSID2006639 CSID16141 CSID16606 CSID13491	105 110 115 hed fragment I Unmatche tions: 10 Description Phenacetin Salsolinol (S)-22-amino-4-phenylbutar 3.4-Dimethylphenyl methyl 4-Amino-3-phenyl butyric	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50 B 49.9 B 49.5 B 49.5 B 49.3	m/z Fragmentati B 87 B 53.4 B 52.6 B 50.9 B 50.9 B 50.9 B 49.7	Iso 155 Iso 155 Iso 137 137 137 137 137 137 137 137 137 137 137	Ret Isotope similar 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Tag filter applied compounds may be hidden Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag 5.50,270.1040n 0 7 7.11,308.1069n 0 20 5.18_151.0637n 0 29 2.88_170.0049h 0 29 2.91_151.0638n 0 0	le Poss ☆ ☆ ☆ ☆ ☆	gend: Match sible identifica Compound ID CSID4590 CSID2609 CSID2006639 CSID16141 CSID16006 CSID13491 CSID4606	105 110 115 hed fragment I Unmatche tions: 10 Description Phenacetin Satsolinol (S)-2-amino-4-phenylbutar 3.4-Dimethylphenyl methyl 3.5-Dimethylphenyl methyl 4-Amino-3-phenyl butyric Phenprobamate	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50 B 49.9 B 49.5 B 49.3 B 48.7	m/z Fragmentati B 87 B 53.4 B 52.6 B 50.9 B 50.9 B 50.9 B 49.7 B 46.9	ior Mass error (ppm) 137 137 137 137 137 137 137 137	Ret Isotope similar 96.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29	
After searching, identifications will be assigned to the relevant compounds automatically. Search for identification Tag filter applied compounds may be hidden Edit mpound Accepted ID Tag 6.50,270.1040n 7.11_308.1069n 6.22_308.0856n 5.18_151.0637n 5.18_151.0637n 5.18_151.0649n 5.18_179.0649n	IS Le Poss ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆	gend: Matcl sible identifica Compound ID CSID4590 CSID2006639 CSID16066 CSID16066 CSID13491 CSID4606 CSID23083	105 110 115 hed fragment Unmatche ations: 10 Description Phenacetin Salsolinol (S)-2-amino-4-phenylbutar 3,4-Dimethylphenyl methyl 3,5-Dimethylphenyl methyl 4-Amino-3-phenyl butyric Phenprobamate IPC	Adducts M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H M+H-H	Formula C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2 C10H13NO2	135 14 Rete Score B 56.7 B 50.0 B 49.9 B 49.5 B 49.5 B 49.3 B 49.3 B 49.3 B 49.3 B 48.7 B 48.7	m/z Fragmentati ■ 87 ■ 53.4 ■ 52.6 ■ 50.9 ■ 50.9 ■ 50.9 ■ 49.7 ■ 46.9 ■ 46.9 ■ 46.9 ■ 17	ior Mass error (ppm) 137 137 137 137 137 137 137 137 137 137	Ret Isotope similar 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29 98.29	

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. CHN Filter) and tick **Filter by elemental composition** and untick Perform theoretical fragmentation.

Identify Compounds Select your identification method:	Compound 2
Q ChemSpider	-
	wnload others
Filter the compounds Using the list below, <u>filter the comp</u> show only those you want to identif	
2 Choose search parameters Select your ChemSpider search par create a new parameter set:	ameters or
Default 🔹	Edit 👻
2 Search for identifications	Edit
After searching, identifications will to the relevant compounds automate	
Search for it	Create a copy
Tag filter applied compounds may be hidden	Delete ti

CHN Filter		
Required search parame	eters	
Precursor tolerance:	5	ppm •
Data source:	KEGG.	
		Select data sources
Fragment tolerance:	5	ppm
Filter by isotope sim	ilarity score	
Isotope similarity:	95	%
Filter by elemental c	omposition	
Filter by elemental c	omposition	

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

The existing composition is displayed.

lement	From	То	Remove?	н																	н
н	0	150	×		_														-	_	
0	0	100	×	Li	Be											В	C	N	0	F	Ν
N	0	10	×	Na	Mg											AI	Si	Р	s	CI	4
C P	0	30 2	×	к	Ca	Sc	Ti	۷	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	•
5	0	2	×	Rb	Sr	Y	Zr	Nb	Мо	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	Ι	>
				Cs	Ba		Hf	Ta	w	Re	Os	Ir	Pt	Au	Hg	TI	Pb	Bi	Ро	At	F
				Fr	Ra		Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	FI	Uup	Lv	Uus	U
						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.

QI Select C	hemSpid	er Element	al Compositio	n																	×
ChemSp	oider el	emental	l compositi	on																	
Define a s	set of ele	ments and	d their count r	anges t	o use	as an e	elemen	tal con	npositi	on filte	er. Lear	rn mor	e in the	e <u>onlin</u>	e refer	ence.					
Element	From	То	Remove?	н																	He
н	0	150	×		_											_			_	-	
C	0	100	×	Li	Be											В	C	N	0	F	Ne
Ν	0	10	×	Na	Mg											AI	Si	Ρ	S	СІ	Ar
				к	Ca	Sc	Ti	۷	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
				Rb	Sr	Y	Zr	Nb	Мо	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	Ι	Xe
				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	Uuo
						La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	l
						Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	I
	Save Cancel																				

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

QI Edit ChemSpider Search	n Parameters	×
ChemSpider search Define a set of ChemSpi reuse. Learn more in the	der parameters that	can be saved for later
Name:		
CHN Filter		
Required search parame	eters	
Precursor tolerance:	5	ppm 🔹
Data source:	KEGG.	
	5	Select data sources
Optional search parame		
Fragment tolerance:	5	ppm
Filter by isotope sim	ilarity score	
Isotope similarity:	95	%
Filter by elemental c Elemental composition		N: 0-10
	Select elem	nental composition
	S	ave Cancel

The new set now displays the Elemental Composition Filter

	Choose search parameters Select your ChemSpider search p create a new parameter set:	ara	ameter	s or
	CHN Filter 🔻		Edi	t 🔻]
3	Default Precursor tolerance: Data sources: Theoretical fragmentation: Isotope similarity filter: Elemental composition filter:	Ki of of	f	ned tions
Con • 4	CHN Filter / X Precursor tolerance: Data sources: Theoretical fragmentation: Isotope similarity filter: Elemental composition filter:	K of of	f	

As an example using the Default search parameters ChemSpider returns 10 possible identities for **Compound 2.46_315.1496**

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. Search for identifications				This o	ompound	has	no fra	gmentation	data	
Tag filter applied		ible identifica					Ŧ			
Compounds may be hidden	\$	Compound ID		Adducts	Formula	Rete	Score	Fragmentatior	Mass error (p	pp
Compound Accepted ID Tag 💌	*	CSID2274171	(2E)-2,3-Diphenyl-3-(2-pyr	M+CH3	C ₂₀ H ₁₄ N ₂		36.2	0	1.34	
● 4.42_136.0758m/z	☆	CSID10128107	(7beta,17R,19E,21alpha)-21	M+H-2	$C_{21}H_{22}N_2O_3$		36	0	1.03	
● 4.62 152.0701m/z	☆	CSID26332999	(1R,10S,12R,13R,14S,16S,18	M+H-2	$C_{21}H_{22}N_2O_3$		36	0	1.03	T T T
● 4.34 176.0710m/z	☆	CSID391595	Methyl (19alpha)-19-methy	M+H-2	$C_{21}H_{22}N_2O_3$		36	0	1.03	
4.08 235.1822m/z	☆	CSID4445235	Methyl (15alpha,19E)-16-fc	M+H-2	C21H22N2O3		36	0	1.03	E N
3.67 368.1013m/z	☆	CSID391596	Methyl (19beta)-19-methyl	M+H-2	C21H22N2O3		36	0	1.03	*
2.97 190.1357m/z	\Rightarrow	CSID4573616	Methyl (19E)-16-formylsar	M+H-2	C21H22N2O3		36	0	1.03	
2.46_315.1496m/z	☆	CSID2272523	(E)-Ranitidine	M+H	C13H22N4O3S		35.9	0	3.25	
0.63_320.1086m/z	☆	CSID77849	Aphistar	M+H	C13H22N4O3S		35.9	0	3.25	
	☆	CSID10128463	Testosterone, 9-chloro-17-	M+H-2	C20H27CIO3		31.5	0	-4.20	*
of 19 filtered compounds have been identified.	٠ .								Þ	b .
Clear compound identifications •										Section Complete

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

2 Choose search parameters Select your ChemSpider search parameters or create a new parameter set: CHN Filter Edit 3 Search for identifications After searching, identifications will be assigned to the relevant compounds automatically. Search for identifications			This co	ompound	has	no fra	gmentation	data	
Tag filter applied compounds may be hidden	Possible identificati			5	0.1		5	M	
	· · · · · · · · · · · · · · · · · · ·	Description 2E)-2,3-Diphenyl-3-(2-pyr	Adducts M+CH ₃	Formula C ₂₀ H ₁₄ N ₂		Score 36.2	-	Mass error (ppm) 1.34	
Compound Accepted ID Tag -		2t)-2,3-01pHenyi-3-(2-pyi-		C201114142		50.2	V	1.34	
• 4.42_136.0758m/z									
○ 4.62_152.0701m/z									
 4.34_176.0710m/z 4.08 235.1822m/z 									N N
 3.67_368.1013m/z 									\checkmark
0 2.97_190.1357m/z									
• 2.46_315.1496m/z									
○ 0.63_320.1086m/z									
< III >									
3 of 19 filtered compounds have been identified.	*							÷.	
Clear compound identifications -									Section Complete $ e e e e e e e e e e e e e $

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

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 96)

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.

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.

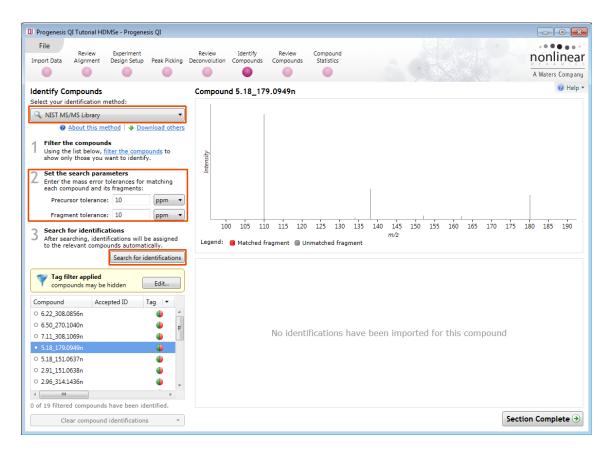
Precursor tolerance

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.

Q Prog	jenesis QI
1	Identification results Successfully imported identifications for 16 compounds.
	ОК

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.



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.

QI Progenesis	; QI Tutorial HI)MSe - Progene	sis QI		
File	Review	Experiment		Review	Identify
Import Data	Alignment	Design Setup	Peak Picking	Deconvolution	Compounds
					•
	ompounds	Compoun	d 5.18_179		
🔍 NIST M	S/MS Library		•		
0	About this m	ethod 🖊 Dov	vnload others]	
Using t		ds filter the compo want to identify		insity	

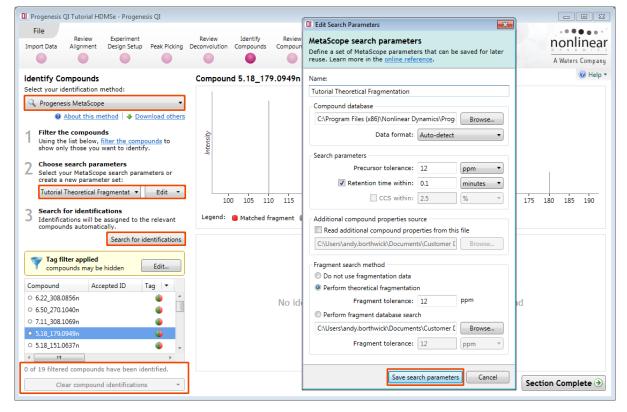
For each identification method the **About this method** link will take you to relevant page(s) for the current method.

Note: there is also a link to Download other methods that are not installed by default.

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), select to clear all identifications, 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.



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 37. The display should look as below.

Q Progenesis QI Tutorial HDMSe - Progenesis QI						
File	Deview Identifie	I Filter the compo	unds			
Import Data Alignment Design Setup Peak Picking	Review Identify Deconvolution Compounds	Create a filter	pounds based			Move tags to the appropriate boxes to eference.
Identify Compounds Select your identification method:	Compound 5.18_17	79. Available tags:			Show o	ompounds that have all of these tags:
Progenesis MetaScope About this method Download others	°				-	ova p-value ≤ 0.05 (1243 compounds) x fold change ≥ 2 (444 compounds)
Filter the compounds Using the list below, filter the compounds to show only those you want to identify.	Intensity				Show c these ta	compounds that have at least one of lgs:
2 Choose search parameters Select your MetaScope search parameters or create a new parameter set: Tutorial Theoretical Fragmentat	0					mpounds that have any of these tags: t identified (5086 compounds)
3 Search for identifications Identifications will be assigned to the relevant compounds automatically. Search for identifications	100 105 11 Legend: Matched f	fra Clear the filter				OK Cancel
	🚖 Compound ID Des	scription	Adducts	Formula	Retention time	Sco
Tag filter applied compounds may be hidden	対 49854487 Phe	enacetin	M+H-H ₂ O	C ₁₀ H ₁₃ NO ₂	5.16	68.4 Fri N
	対 HMDB31811 3,5-	-Dimethylphenyl methyl	M+H-H ₂ O	C10H13NO2		54.1 0
Compound Accepted ID Tag 💌	* HMDB40021 2,3-	-Dihydro-5-(3-hydroxyp	M+H-H ₂ O	C10H13NO2		50.6
6.22_308.0856n 6.50_270.1040n	🖈 HMDB41931 3,4-	-Methylenedioxyamphe	M+H-H ₂ O	C ₁₀ H ₁₃ NO ₂		48.4
7.11_308.1069n						
• 5.18 151.0637n						
All 19 filtered compounds have been identified.	۰ III.					Þ
Clear all compound identifications						Section Complete $ ightarrow$

Finally to accept Phenacetin as **Compound 5.18_179.0949n'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 nport Data Alignment Design Setup Peak Picking D	Review Identify Deconvolution Compound								ņ	online
									A	Waters Compa
eview Compounds	Find a compound: See	arch	۶		Filter cor	mpounds	✓ Filter is	active		Hel
sing this screen, you can find the compounds of terest in your experiment.	Compound	Neutral mass	m/z	z Rete	ntion time	Drift time	Peak Width	Tag 💌	Accepted ID	Identification
Create a shortlist to review	 3.69_195.0897m/z 	<unknown></unknown>	195.0897	1 3.69		1.94	0.16			2
In the table, sort and <u>filter the compounds</u> based	6.50_270.1040n	270.1040	271.1113	1 6.50		2.75	0.14			1
on their measurements, to generate a shortlist for further review.	• 5.18_179.0949n	179.0949	180.1043	1 5.18		2.05	0.18	۵	49854487	4
How are the measurements calculated?	4.42_136.0758	hknown>	136.0758	1 4.42		1.57	0.14	•		1
To sort the table by a given value, simply click	2.96_314.1436	.1436	315.1509	1 2.96		3.02	0.27			1
the relevant column header.	• 4.08_235.1822m/z	<unknown></unknown>	235.1822	1 4.08		2.54	0.18			1
Review the compounds For each compound of interest, inspect the ions'	• 2.91_151.0638n	151.0638	152.0707	1 2.91		1.67	0.18			4
alignment and peak picking:	• 7.12_163.0403m/z	<unknown></unknown>	163.0403	1 7.12		1.62	0.17			1
Review selected compound	• 2.46_315.1496m/z	<unknown></unknown>	315.1496	1 2.46		2.97	0.06	۲		1
B Iso double-click to review a compound.	•									÷
	Compound 5.18_179.	0949n:								
Choose the correct identifications For each compound, select one of its possible	Compound abundan	ce Possible id	lentifications	3D M	lontage [Drift time me	ontage			
identifications as the accepted one.		C Norm				A LD			B HD	
To speed this up, you can <i>automatically</i> accept	a								-	
identifications in compounds where only one of the possible identifications has:	10 – 10 –)• (}	
Score ≥ 50.0 Accept identifications	pung -									
	ArcSinh Normalised Abundance									
Export data for further processing By exporting your data to external tools, there's	sile 5 –									
no limit to your analysis.					г	_				
Export to pathways tool						C				
	- ArcS	18			L					
periment design	-									, r

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_179.0949n as an example the 4 views for Window C show:

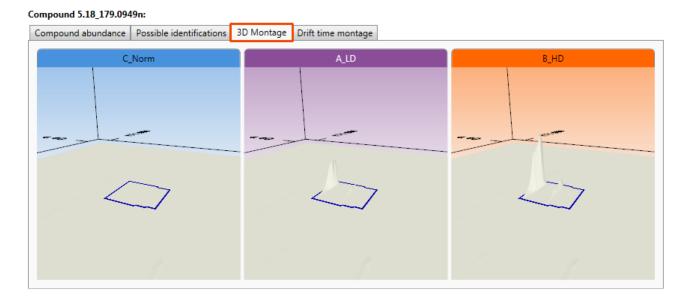
	· .		5.18_179.094						
Co	mpo	und	l abundance	Possible identifications	3D Montage	Drift time montage			
				C_Norm		A_LD			
ArcSinh Normalised Abundance	10					M		I∳I	
A	0		4						•

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

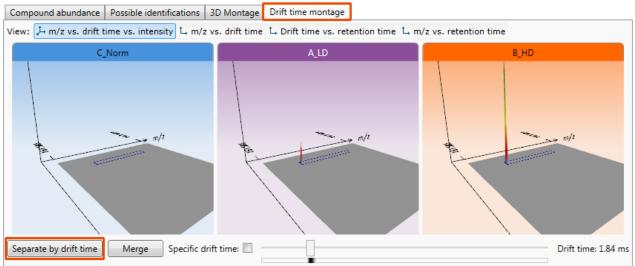
Possible identifications for current compound with Overall Score based on, Mass error, Retention Time (if available), Isotope Similarity based on the comparison of the measured isotope distribution for the compound vs the expected based on the compound formula, Fragmentation Score and CCS if available.

Cor	npound abunda	nce Possible identifications	5 3D Mon	tage Drift	time montage			
os	sible identific	ations: 4						
${\approx}$	Compound ID	Description	Adducts	Formula	Retention time	Score	Fragmentation score	
☆	49854487	Phenacetin	M+H-H	C ₁₀ H ₁₃ NO ₂	5.16	67.3	89	NH
☆	HMDB31811	3,5-Dimethylphenyl methyl	M+H-H	C ₁₀ H ₁₃ NO ₂		49.5	50.9	
☆	HMDB40021	2,3-Dihydro-5-(3-hydroxyp	M+H-H	C ₁₀ H ₁₃ NO ₂		43	18.4	
☆	HMDB41931	3,4-Methylenedioxyamphe	M+H-H	C10H13NO2		41.7	11.9	0 4 -

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



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



Compound 5.18_179.0949n:

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_308.1069n click on the 'star'.

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

File Review Experiment moort Data Alignment Design Setup Peak Picking D	Revi		Identify		ompound						n	online
	•			•							A	Waters Comp
eview Compounds	Find a	a compour	nd: Search		2	Filter compoun	ds 👻 Filter i	s active		1999 A. 1997		🔞 He
sing this screen, you can find the compounds of terest in your experiment.	dth	Tag 💌	Accepted ID	Identifications	Anova (p)	Max fold change	Highest mean	Lowest mean	Isotope distribution	Max Abundance	Min CV%	Description
Create a shortlist to review			49854487	4		6.51E+04	B HD	C Norm		12144.2550	1.82	Phenacetin
In the table, sort and filter the compounds based				1		6.59E+03	B HD	C Norm		273.4031	1.70	
on their measurements, to generate a shortlist for further review.		-		1	< 1.1E-16		B HD	C Norm	-	11753.4811	4.22	
Ø How are the measurements calculated?		-		1	< 1.1E-16		B HD	C Norm		8719.9385	2.42	
To sort the table by a given value, simply click		-		4	< 1.1E-16		B HD	C Norm		1290.5303	2.00	
the relevant column header.		-		1		1.38E+05	B HD	A LD		3722,4198	1.09	
Review the compounds		-		1	1.11E-16		B HD	A LD		95,4432	7.58	
 For each compound of interest, inspect the ions' alignment and peak picking; 			4702	± 1		1.69F+04	B HD	A LD		18493.4165	2.91	Warfarin
Review selected compound			4702	1	1.99E-14		B HD	A LD		43.8736	8.92	wanann
Review selected compound	4	-		1	1.995-14	190	6_HU	A_LU		43.0730	6.92	
You can also double-click to review a compound.												
Choose the correct identifications For each compound, select one of its possible identifications as the accepted one.	Con	npound at	1_308.1069n: oundance Pc ntifications:	ossible identifica	tions 3D I	Nontage Drift tim	ne montage					, *
Choose the correct identifications For each compound, select one of its possible identifications as the accepted one. To speed this up, you can <i>automatically</i> accept	Con	npound at sible ider	oundance Po	ossible identifica 1	tions 3D I Adducts		tion time	e Fragmentati	on score			
Choose the correct identifications For each compound, select one of its possible identifications as the accepted one. To speed this up, you can <i>automatically</i> accept identifications in compounds where only one of the possible identifications has:	Con Pos	npound at sible ider	oundance Po ntifications:	ossible identifica 1 otion	Adducts		tion time Scor	Fragmentati 87.2	on score	f		
Choose the correct identifications For each compound, select one of its possible identifications as the accepted one. To speed this up, you can automatically accept identifications in a select one of the possible identifications have the possible identifications have Score ≥ 50.0 Accept identifications Export data for further processing by exporting your data to external tools, there's no limit to your analysis. Export to pathways tool	Con Pos:	npound ab sible ider Compour	ntifications: nd ID Descrip Warfari	ossible identifica 1 otion	Adducts	Formula Reter	tion time Scor		Ĺ	ОН		0
Choose the correct identifications For each compound, select one of its possible identifications as the accepted one. To speed this up, you can <i>automatically</i> accept identifications in compounds where only one of the possible identifications. Score ≥ 50.0 Accept identifications Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	Con Pos	npound ab sible ider Compour	oundance Po ntifications: nd ID Descrip	ossible identifica 1 otion	Adducts	Formula Reter	tion time Scor		on score	ОН		°

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 Experiment mport Data Alignment Design Setup Peak Picking	Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics							nonline
											A Waters Comp
eview Compounds	Find a com	pound: Search		P	Filter compo	unds 👻 Filte	r is active				🔞 He
sing this screen, you can find the compounds of terest in your experiment.	th Tag	Accepted ID	Identifications	Anova (p)	Max fold change	Highest mean	Lowest mean	Isotope distribution	Max Abundance	Min CV%	Description
Create a shortlist to review		14708992	1	< 1.1E-16	6.59E+03	B_HD	C_Norm		273.4031	1.70	Acetanilide
In the table, sort and filter the compounds based on their measurements, to generate a shortlist	4	7847292	1	< 1.1E-16	1.5E+05	B_HD	A_LD		17223.3822	4.50	Alprazolam
for further review.	4	46511425	2	< 1.1E-16	1.32E+03	B_HD	C_Norm		6171.0708	2.51	Caffeine
How are the measurements calculated?	۷	855682	1	< 1.1E-16	Infinity	B_HD	C_Norm		8719.9385	2.42	Lidocaine
To sort the table by a given value, simply click the relevant column beader.	(49854487	4	< 1.1E-16	6.51E+04	B_HD	C_Norm		12144.2550	1.82	Phenacetin
the relevant column header.		7847488	1	< 1.1E-16	814	B_HD	C_Norm		11753.4811	4.22	Ranitidine
Review the compounds For each compound of interest, inspect the ions'	۷	9357	1	< 1.1E-16	3.33E+03	B_HD	C_Norm		1166.0506	3.91	Tolbutamide
alignment and peak picking:	4	4702	1	2.22E-16	1.69E+04	B_HD	A_LD		18493.4165	2.91	Warfarin
Review selected compound			1	1.99E-14	198	B_HD	A_LD		43.8736	8.92	
You can also double-click to review a compound.	*										÷
	Compound	5.18_179.0949	n:								
Choose the correct identifications For each compound, select one of its possible	Compour	d abundance	Possible identific	ations 3D	Montage Drift	time montage					
identifications as the accepted one.	Possible	identification	s: 4								
To speed this up, you can automatically accept identifications in compounds where only one of	🚖 Com	pound ID Desc	ription	Adducts	Formula F	Retention time	Score Fragmer	ntation scor			
	🗙 498	54487 Phen	acetin	M+H-H	C10H13NO2	5.16 6	57.3 89			<u> </u>	\checkmark
the possible identifications has:	🚖 HMI	DB31811 3,5-D	imethylphenyl m	M+H-H	C10H13NO2	4	9.5 50.9		ſ		
the possible identifications has: Score ≥ 48 Accept identifications	🚖 HMI	0B40021 2,3-D	ihydro-5-(3-hyd	M+H-H	C10H13NO2	4	3 18.4	/	\sim		0
Score ≥ 48 Accept identifications	~ FINIL						1.7 11.9		U I		
Score 2 48 Accept identifications Export data for further processing By exporting your data to external tools, there's)B41931 3,4-N	lethylenedioxyar	M+H-H	C10H13NO2	4					
Score ≥ 48 Accept identifications Export data for further processing)B41931 3,4-N	lethylenedioxyar	M+H-H	C ₁₀ H ₁₃ NO ₂	4					
Score 2 48 Accept identifications Export data for further processing By exporting your data to external tools, there's		DB41931 3,4-N	lethylenedioxyar	r M+H-H	C ₁₀ H ₁₃ NO ₂	2					
Score ≥ 48 Accept identifications Export data for further processing By exporting your data to external tools, there's no limit to your analysis.		DB41931 3,4-N	lethylenedioxyar	r M+H-H	C ₁₀ H ₁₃ NO ₂	2					
Score ≥ 48 Accept identifications Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	☆ HMC		lethylenedioxyar	r M+H-H	C ₁₀ H ₁₃ NO ₂	2					
Score ≥ 48 Accept identifications Export data for further processing By exporting your data to external tools, there's no limit to your analysis.		DB41931 3,4-N	lethylenedioxyar	r M+H-H	C ₁₀ H ₁₃ NO ₂	2		Þ			

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_151.0638n) and Phenacetin (5.18_179.0949) were accepted manually.

Compound Validation View

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

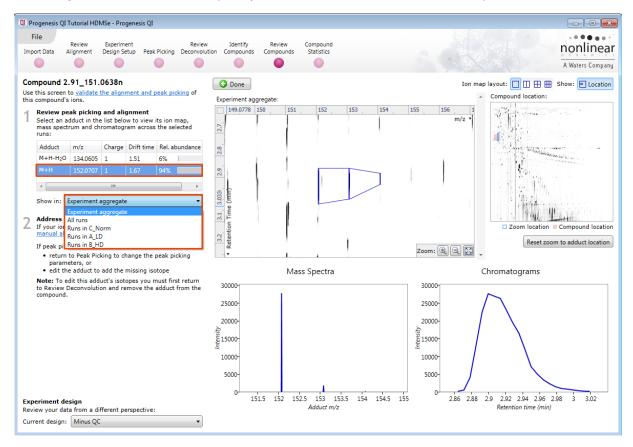
In this view:

- 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_151.0638n (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.



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.

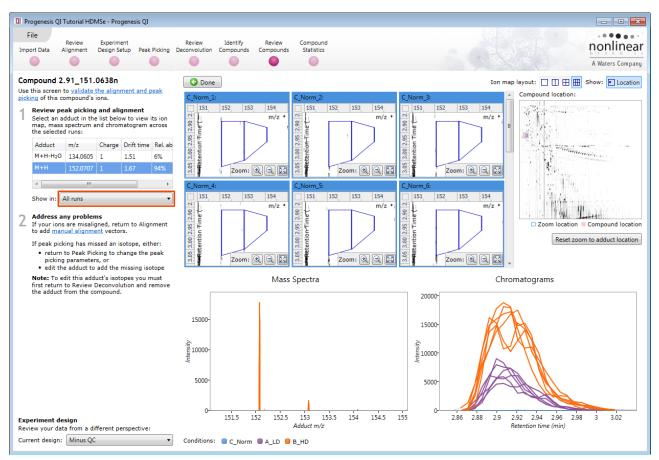
Below the list of adducts, a drop-down list lets you select whether you want to see the selected adduct:

- 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. The aggregate pattern of detection can be edited to add a missing isotope using the **Edit adduct** facility.

Note: details on Editing adducts is available in Appendix 7 (page 98)

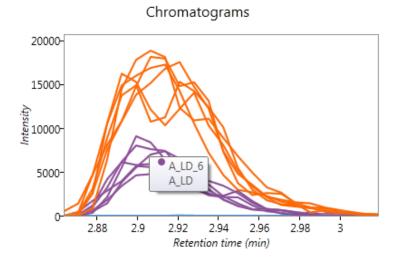
To validate the alignment you will find it best to select to select the **All runs** option and then adjust the ion map layout to display multiple runs



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.

	Neutral mass 179.0949 Anova p-value ≤ Max fold change Not identified	180.1043 0.05	z 1 1 1	Retention time 5.18 3.69 2.91 4.42	2.05 1.94 1.67	Peak Width 0.18 0.16 0.18	Tag 🔍	Accepted ID 49854487 46511425 46506142	Identifications 4 2	< 1.1E-16	Max fold change 6.51E+04 1.32E+03	*				
4 0 0 0 0	Anova p-value ≤ Max fold change Not identified	0.05	1 1 1	3.69 2.91	1.94 1.67	0.16	4	46511425	4			Ê				
	Max fold change Not identified		1 1 1	2.91	1.67		-		2	< 1.1E-16						
	Max fold change Not identified		1 1			0.18	۱	46506142								
	Not identified		1	4.42	1 57				4	< 1.1E-16	64.9					
N					1.57	0.14	٠	14708992	1	< 1.1E-16	6.59E+03					
	New tag		1	2.96	3.02	0.27	۹	7847488	1	< 1.1E-16	814					
_	Quick Tags		r-	Anova p-val	242	0.00		7847292	1	< 1.1E-16	1.5E+05					
~	<u> </u>			Max fold cha			•	855682	1	< 1.1E-16	Infinity					
	Edit tags		Minimum CV				476106	1	< 1.1E-16	1.38E+05						
6.50_270.1 Add to Clip Gallery	• 6.50_270.1 Add to Clip Gallery	Add to Clip Gallery			Add to Clip Gallery						•	9357	1	< 1.1E-16	3.33E+03	
069n	308.1069	347.0700		Not fragmen	ted		•	4702	1	2.22E-16	1.69E+04					
272m/z	<unknown></unknown>	511.3272	Г	Separated by	drift time				0	0.261	1.29					
133m/z	<unknown></unknown>	502.8133		Identified an	d separated	by drift time			0	0.144	1.7	Ŧ				
2	069n 272m/:		069n 308.1069 347.0700 272m/z <unknown> 511.3272</unknown>	069n 308.1069 347.0700 272m/z <unknown> 511.3272</unknown>	Add to Cip Salety Not identifie 069n 308.1069 347.0700 Not identifie 272m/z <unknown> 511.3272 Separated by</unknown>	Add to Cip Galley Not identified Not fragmented 272m/z < unknown > 5113272 Separated by drift time	Add to Cip Galley Not identified 069n 308.1069 347.0700 272m/z sunknown> 511.3272 Separated by drift time Separated by drift time	Add to Cip Galley Not identified 069n 308.1069 347.0700 Vot identified Not identified 272m/z cunknown> 511.3272 Separated by drift time Separated by drift time	Add to Cip Galley 9357 069n 308.1069 347.0700 Not identified 4702 272m/z separated by drift time	Not identified 9357 1 069n 308.1069 347.0700 Not identified 4702 1 272m/z support Separated by drift time 0 0	Add to Cip Galety 9357 1 < 1.1E-16 069n 308.1069 347.0700 Not identified 4702 1 2.22E-16 272m/z sunknown> 511.3272 Separated by drift time 0 0.261	Not identified 9357 1 < 1.1E-16 3.33E+03 069n 308.1069 347.0700 Not identified 4702 1 2.22E-16 1.69E+04 272m/z cunknown> 511.3272 Separated by drift time 0 0.261 1.29				

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 to create the filter. For more guidance, please s	of their tags. Move tags to the appropriate boxes see the <u>online reference</u> .
Available tags:	Show compounds that have all of these tags:
Not identified (5086 compounds) Anova p-value \$ 0.05 (1243 compounds) Max fold change 2 2 (444 compounds)	Show compounds that have at least one 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.

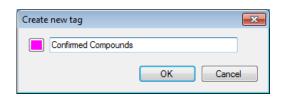
	le Review rt Data Alignment	Experiment Design Setup Peak Picking De	Review Identify convolution Compounds	Review Compounds	Compound Statistics							non	line
-												A Water	s Comp
ev	iew Compounds		Find a compound: Seal	rch	2	Filter com	pounds •	 Filter is a 	tive				🕜 Hel
	g this screen, you can f est in your experiment		Compound	Neutral mass	m/7 7	Retention time	Drift time	Peak Width	Tag	Accented ID	Identifications	Anova (n)	Max fr
	Create a shortlist to		 3.78_400.0568n 	400.0568	401.0641 1		4.05	0.06		Accepted to	0	0.489	1.07
	In the table, sort and	filter the compounds based	 3.82 867.1741n 	867.1741	885.2079 1		8.26	0.12			0	0.882	1.07
	on their measurement: further review.	s, to generate a shortlist for	 3.82_007.1741m 3.81 436.0958m/z 	<unknown></unknown>	436.0958 1		4.27	0.05			0	0.0561	1.72
		e measurements calculated?	 3.81_436.0356m/2 3.82_885.2068m/z 	<unknown></unknown>	885.2068 2		5.13	0.09			0	0.0279	1.16
	To sort the table by a	given value, simply click the	 3.82_003.2008m/z 3.82_901.1809m/z 	<unknown></unknown>	901.1809 2		5.13	0.09			0	0.132	1.10
	relevant column heade		 3.82_901.1809m/2 3.80_302.0462m/z 	<unknown></unknown>	302.0462 2		2.00	0.13			0	0.152	1.18
	Review the compour	nds	 3.80_302.0462m/2 3.80 525.0562m/z 	<unknown></unknown>	525.0562 2		3.19	0.13			0	0.0298	1.10
		interest, inspect the ions'							-		0		
	alignment and peak pi	Review selected compound	• 3.81_311.0515m/z	<unknown></unknown>	311.0515 2	3.81	2.05	0.16	-		0	0.00537	1.06
	identifications as the a To speed this up, you identifications in comp the possible identificat Score ≥ 55.0 Export data for furth	an automatically accept ounds where only one of ions has: Accept identifications her processing to external tools, there's	View: 上 m/z vs. drif	t time vs. inter	nsity L₁ m/z	vs. drift time	Drift time A_LD		time 🔓 r	n/z vs. retenti	B_HD		
	riment design w your data from a di	Export to pathways tool	Separate by drift time	Merge	Specific dr	ift time:	k t=					Drift tim	

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.

Progenesis QI Tutor	ial HDMSe - Progenes	is Qi												
File Revie mport Data Alignn		Peak Picking	Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics						~	nonline	ą
•												-	A Waters Comp	an
eview Compour	nds		Find a com	pound: Sear	ch	P	Filter com	pounds 👻					🕑 He	alp
sing this screen, yo terest in your expe	u can find the compo riment.	ounds of	Tag 💌	Accepted ID	Identifications	Anova (p)	Max fold change	Highest mean	Lowest mean	Isotope distribution	Max Abundance	Min CV%	Description	
Create a short	list to review		۵	14708992				B_HD		_ _	273.4031			4
	t and filter the comp ements, to generate		۵	78472	Anova p-value s	< 0.05	.5E+05	B_HD		-	17223.3822	4.50	Alprazolam	
further review.			۵	46511	Max fold change		32E+03				6171.0708			
How	are the measurement	nts calculated?	۵	05560	Not identified		ifinity	B_HD		 _	8719.9385			
To sort the table relevant column	by a given value, si	mply click the	۵	46506	Separated by dri	ift time	4.9	B_HD			1290.5303	2.00	Paracetamol	
relevant column	nedder.		۵	49854	New tag		51E+04	B_HD		 _	12144.2550	1.82	Phenacetin	
Review the con	mpounds und of interest, inspe	et the ione'	۵	78474			14	B_HD			11753.4811	4.22	Ranitidine	
alignment and p		et the lons	۵	9357	Quick Tags	,	33E+03	B_HD	C_Norm		1166.0506	3.91	Tolbutamide	
	Review select	ed compound	٩	4/02	Edit tags		.69E+04	B_HD	A_LD		18493.4165	2.91	Warfarin	
	uble-click to review a		٨		Add to Clip Gall	ery	7E+05	B HD	C Norm		229.3539	4.81		١.

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



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

 [I] Filter the compounds Create a filter Show or hide compounds based on a selection boxes to create the filter. For more guidance, p 	
Available tags: Available tags: Max fold change ≥ 2 (444 compounds) Not identified (5086 compounds) Separated by drift time (130 compounds)	Show compounds that have all of these tags: Confirmed Compounds (9 compounds) Show compounds that have at least one of these tags: Hide compounds that have any of these tags:
Clear the filter	OK Cancel

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:

🔛 Export	Compound Identifications
	Export complete
	Open File Open Folder Close

Export Compound Identification	ns	×
Choose properties to be included in	n exported file	
Compound		
Compound ID		
Accepted?		
Adducts		
Formula		
Score		
Fragmentation Score		
Mass Error (ppm)		
Isotope Similarity		
Theoretical Isotope Distribution	1	
Link		
Description		
Neutral mass (Da)		
m/z		
Charge		
Retention time (min)		
🗹 Anova (p)		
Max Fold Change		
Tags		
	ок с	ancel

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)

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 103)

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:

Q Progenesis QI Tutorial HDMS	- Progenesis QI			
	,	Review Compound mpounds Statistics		A Waters Company
Review Compounds	Find a compound: Search	P Filter compour	nds 💌 Filter is active	🕜 Help 🔻
Using this screen, you can find interest in your experiment.		I mass m/z z Reten Filter con	npounds	Accepted ID Ide
Create a shortlist to re In the table, sort and filte on their measurements,	r the compounds based 0 4.42 136.0758m/z curker	Clear filte	@	49854487 4 * 14708992 1 =
for further review.	QI Filter the compounds		0.14	9357 1 7847292 1
To sort the table by a giv the relevant column head	Create a filter Show or hide compounds based on a selection of their boxes to create the filter. For more guidance, please s		0.27	7847488 1 +
 2 Review the compound of int alignment and peak pickling and peak pick	 Not identified (5086 compounds) Anova p-value ≤ 0.05 (1243 compounds) Max fold change ≥ 2 (444 compounds) Separated by drift time (130 compounds) 	confirmed Compounds (9 compounds)	ne montage	HD
	Clear the filter	OK Cancel		
Experiment design Review your data from a differ	ant perspective:			,
Current design: Minus QC	•		Secti	ion Complete 🏵

Λ	Export data for further processing
t	By exporting your data to external tools, there's no limit to your analysis.

Export to pathways tool

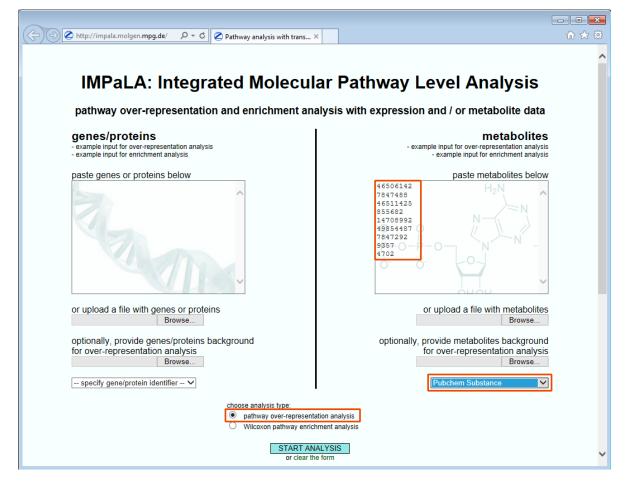
Then click Export to pathways tool, and select IMPaLA from the Pathways tools and click Next.

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	×
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) Select the type of analysis to perform:	
Pathway over-representation analysis	•
Pathway over-representation analysis Wilcoxon pathway enrichment analysis Open IMPaLA in my browser	
< Back Copy compounds to clipboard Close	

When IMPaLA opens, paste the metabolites in to the right hand panel.



Select specific metabolite identifier

Click Start Analysis

Analysis results list is returned:

32 pathways found.				Results per page: 50 V Go to page (previous) 1 of 1 (next)			
pathway name	pathway source	overlapping metabolites	all metabolites	Pmetabolites	Qmetabolites		
Warfarin Pathway, Pharmacodynamics	PharmGKB	1	1 (1)	0.000867	1		
Warfarin Pathway, Pharmacokinetics	PharmGKB	1	1 (1)	0.000867	1		
Warfarin Action Pathway	SMPDB	1	5 (5)	0.00433	1		
Benzodiazepine Pathway, Pharmacokinetics	PharmGKB	1	11 (11)	0.00951	1		
Xenobiotics	Reactome	1	51 (51)	0.0435	1		
Cytochrome P450 - arranged by substrate type	Reactome	1	115 (116)	0.0961	1		
Phase 1 - Functionalization of compounds	Wikipathways	1	137 (137)	0.114	1		
Phase 1 - Functionalization of compounds	Reactome	1	150 (151)	0.124	1		
Defective CYP27A1 causes Cerebrotendinous xanthomatosis (CTX)	Reactome	1	150 (151)	0.124	1		
Defective CYP1B1 causes Glaucoma	Reactome	1	150 (151)	0.124	1		
Defective CYP4F22 causes Ichthyosis, congenital, autosomal recessive 5 (ARCI5)	Reactome	1	150 (151)	0.124	1		
Defective CYP11B2 causes Corticosterone methyloxidase 1 deficiency (CMO-1 deficiency)	Reactome	1	150 (151)	0.124	1		
Defective CYP24A1 causes Hypercalcemia, infantile (HCAI)	Reactome	1	150 (151)	0.124	1		
Defective MAOA causes Brunner syndrome (BRUNS)	Reactome	1	150 (151)	0.124	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

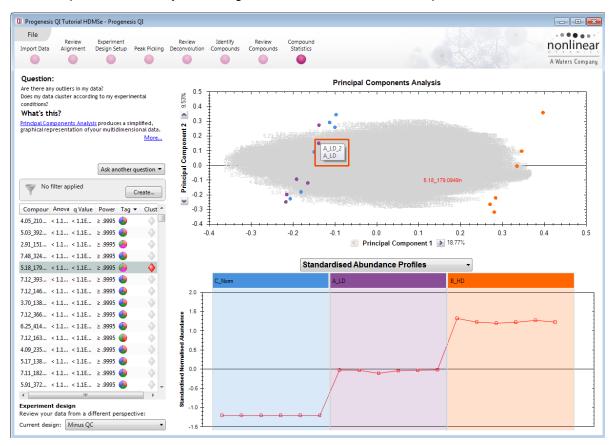
The user guide now describes the functionality of the Multivariate Statistics.

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 of the 5154 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.

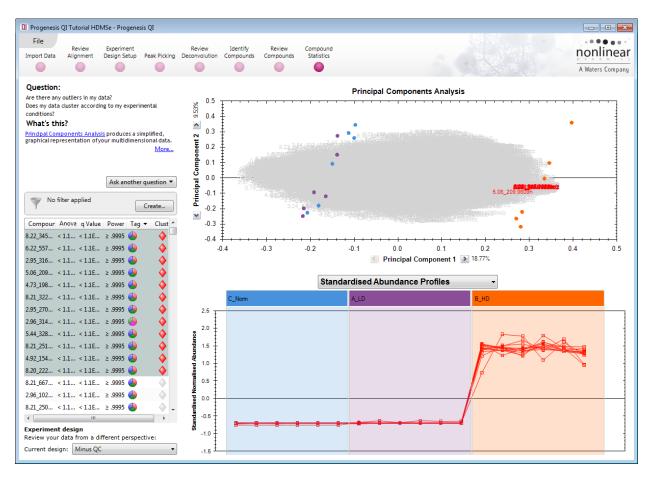
Principal Component Analysis (PCA)

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 9 (page 106)

Correlation Analysis

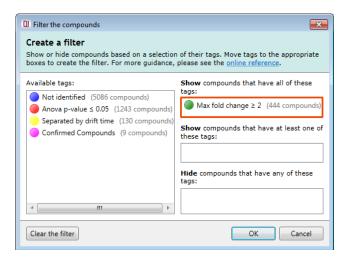
Use the tags created in Review Compounds to filter the compounds displayed in the table. We are going to explore the Correlation Analysis for all the Compounds (with possible identifications) that display a significant 2 fold or greater difference in abundance.

To filter the data click Create

	Ask another question 💌
No filter applied	Create

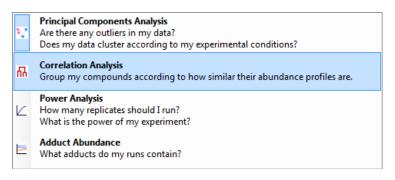
Set up the filter as shown below





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'



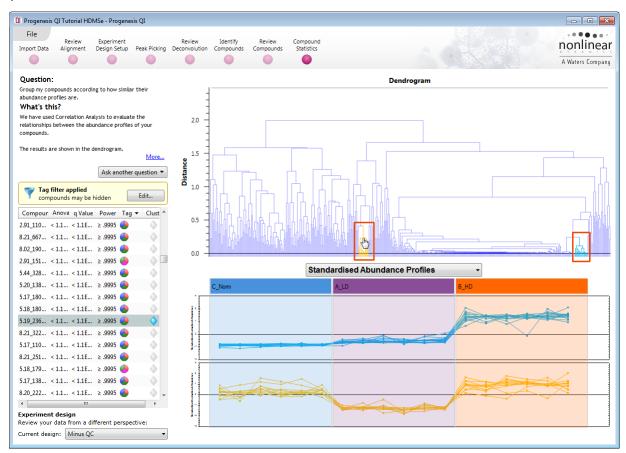
This time the statistically based question(s) being asked is:

'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.

Q Progenesis QI Tutorial HDMSe - Progenesis QI							
File							
	Save	nt tup	Peak Picking				
	Close						
	Export inclusion list						
	Funnation F7infa	imila	r their				
	Export to EZinfo						
	Open EZinfo project	L .					
	Experiment properties		e the of your				
	Show Clip Gallery	Unites	or your				
	·····/						
×	Exit	ľ	More				

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 H	HDMSe - Progenesis QI			
File Review Import Data Alignment	Experiment Design Setup Peak Pick	Review Identify Review ng Deconvolution Compounds Compounds	Compound Statistics	nonlinear
	•		•	A Waters Company
Conditions	Minus QC L	ow vs High I 🗡 🗋 New		🕑 Help 🕶
Setup conditions		Runs Add Selected Runs to Condition	▼ Search	Q
	at you want to compare be tc), and then assign each o rect condition.		C_Norm_2	C_Norm_3
Low	Dele	te Sign	244	
	A_LD_1 Remo	e la	1. 19 A. 19 C. 1	- 39 F 14 1
	A_LD_2 Remo	<u>e</u>		
	A_LD_3 Remo	C_Norm_4	C_Norm_5	C_Norm_6
	A_LD_4 <u>Remo</u>			
	A_LD_5 Remo			1997 - E
🔲 High	A_LD_6 <u>Remo</u>			
la High	B_HD_1 Remo		D_QC_2	D_QC_3
	B_HD_2 Remo			
	B_HD_3 Remo	154 21	2. A. C.	
	B_HD_4 Remo			
	B_HD_5 Remo			
	B_HD_6 Remo	D_QC_4		-
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.

Q Progenesis (QI Tutorial HDMSe - Pro	igenesis QI						• ×
File								
🛃 Save		nt tup Peak Pic	Review Identify king Deconvolution Compound		mpound atistics		nonli	near
Close					•		A Waters	Comnanu
Export inc	clusion list				-			rempang
			8 07		Principal Compor	nents Analysis		
Export to		perimental	QI Save As			×		
Open EZir	nfo project		~~~~	Data 🕨 Progenesis QI		earch Progenesis QI v2	•	1
Experimer	nt properties	s a simplified.		Data 🖡 Progenesis QI	v2 🕨 👻 😽 Se			
Show Clip		imensional da	Organize 🔻 New folder			:= 🔹 🕡		
	,	Mo	☆ Favorites		Documents lib	Arrange by: Folder 🔻	Start Zara	
× Exit			E Desktop	=	Progenesis QI v2	Arrange by: Folder 🔻	• 772	
	Ask ar	nother questio	🐌 Downloads		Name	*	557.1310m/z	
			😌 Dropbox		Progenesis QI v2 To	utorial	101	-
No fil	lter applied	Create.	Recent Places		Jan Högenesis Qr 12 H			
			🔊 ShareFile					
	Anova q Value Power 1.1 < 1.1E ≥ .9995	-	Mesktop	-	•	•	.3 0.4	0.5
-	1.1 < 1.1E ≥ .9995 1.1 < 1.1E ≥ .9995	-						
-	1.1 < 1.1E ≥ .9995	-		nesis QI Tutorial HDMSe		•		
	1.1 < 1.1E ≥ .9995	-	Save as type: UMetri	cs EZinfo project (*.usp)	▼		
	1.1 < 1.1E ≥ .9995	-						
		-			_			
-	1.1 < 1.1E ≥ .9995	-	Hide Folders			Save Cancel		
-	1.1 < 1.1E ≥ .9995	-				ii.		
-	1.1 < 1.1E ≥ .9995	-	2 0.5 -		/	/		-
-	1.1 < 1.1E ≥ .9995	-			/			
5.91_372 <	1.1 < 1.1E ≥ .9995		Po.0					
5.03_392 <	1.1 < 1.1E ≥ .9995		- ²					1
•	III		- 0.5					
Experiment			k -1.0					-
	data from a different po n: Low vs High	erspective:	-					
Current desig	in: 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.

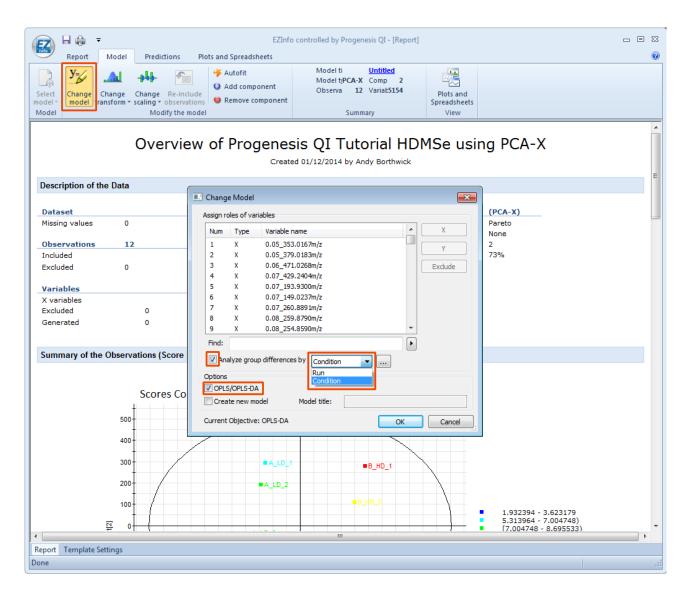
Note: the dialog may be obscured by the EZinfo application as it opens.

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).

Q 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 co for identification or further investigation.	mpounds 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



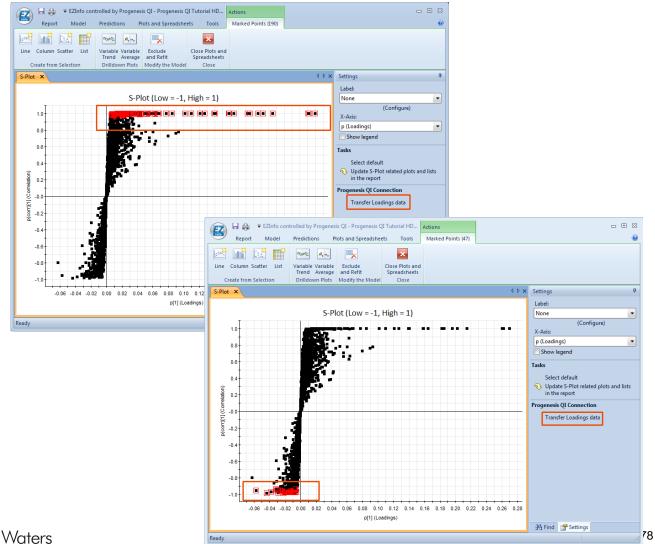
Select analyse group differences by Condition.

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

🛃 H 🖨 🔻 _ 0 X EZInfo con olled by Progenesis QI - Progenesis QI Tutorial HDMSe Report Model Predictions Plots and Spreadsheets Tool ht ht Predictionset * in the second VIP https://www. × Classification list Scores Hotelling's Distance Observed vs. T2 Range to Model Predicted Observation plots Goodness X/Y of Fit Overview Model plots Spreadsheets VIP Loadings Bi Plot Loadings Coeffici 5-Plot Variable Trend Close Plots and Spreadsheets Close ∢ ⊳ × S-Plot × Settings ņ Label: S-Plot (Low = -1, High = 1) None -(Configure) 1.0 .. . - -- - - -- -X-Axis: p (Loadings) • Sho v legen 0. Tasks Select default 0.6 Progenesis QI Conr Transfer Loadings data 0. 0.: 5 р(соп)[1] (Со -0. -0. -0. -0.6 -0.8 -1. -0.06 -0.04 -0.02 0.00 0.02 0.04 0.06 0.08 0.10 0.12 0.14 0.16 0.18 0.20 0.22 0.24 0.26 0.28 p[1] (Loadings) 者 Find 🛛 😤 Settings Displays p(corr) - the y-related loading vector - versus the OPLS-regression coefficients (b) or weights (w), or loadings (p). Important X-variables have large positive or la

The PCA will update the new model. Now Select **Plots and Spreadsheets** on the main tool bar and click on the **S-Plot**

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.

QI Filter the compounds

Clear the filter

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

Create new tag	X
S_Plot_B	
	OK Cancel

QI 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 c for identification or further investigation.	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 💌						
Import	tags Cancel					
Import	tags Cancel					

×

S_Plot_B (47 compounds)

Hide compounds that have any of these tags:

OK

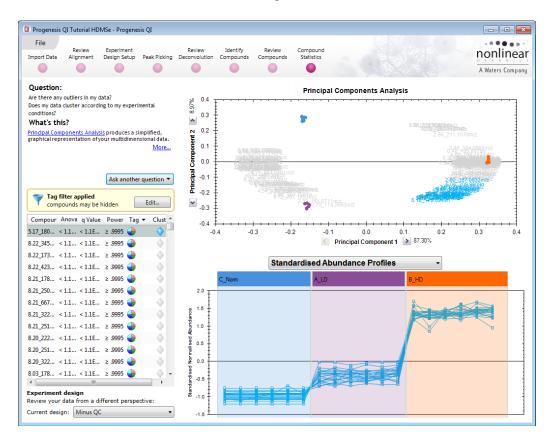
Cancel

Click Import tags

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

	Create a filter Show or hide compounds based on a selection of to create the filter. For more guidance, please s	
Set a new Tag Filter in Compound Statistics to Show	Available tags:	Show compounds that have all of these tags:
ů i	 Not identified (5086 compounds) 	
compounds that have at least one of these tags.	Anova p-value ≤ 0.05 (1243 compounds)	
	Max fold change ≥ 2 (444 compounds)	
	Separated by drift time (130 compounds)	Show compounds that have at least one of these tags:
	Confirmed Compounds (9 compounds)	these tags:
	- · · · · ·	S_Plot_A (190 compounds)

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



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

Appendix 1: 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 experin	nents	Search	٩	Getting started Here are some resources to help you get started with Progenesis QI:
	QI Create New Experiment Create a new ESI experiment Analyse samples run using electrospray injection (ESI) rompounds. Enter a name for the new experiment: Progenesis QI Tutorial HDMSe	to generate the charge on		Eundamental concepts of the Progenesis QI workflow User guide and tutorial data Frequently-asked questions EZinfo version information Version 3.0.3.0 of EZinfo is installed.
There	Select the location to store the experiment files: Select the location to store the same folder as the run data Choose an experiment folder	Browse	nts	Metabolomics Meeting, 20-21 Oct 2015 Waters MS HQ, Wilmslow, UK Are you looking to apply mass spectrometry techniques for routine metabolic profiling and advanced characterisation of metabolomes?
	Back	Next Cancel		Find out more and register at: www.waters.com/metabolomics2015. Latest blog posts • Progenesis OI and the iKnife: the cutting edge of food monitoring • Progenesis OI v2.1 – available to download now • MetSoc2015 – Bridging the gap between community and industry • Progenesis SDF Studio v1.0: using MOL
Other experime	ents 🎯 Open		2.1.5674.29829	 Flogeness SDF State VLO. Using Flot files just got easier ASMS 2015: Mass specs, plugs and rock 'n' roll

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

Q Create New Experiment Set the analysis parameters The type of machine, and the settings used, determin analysed.	e 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 S	r ICIEX TripleTOF, Agilent QTOF, Bruker Maxis, Thermo LTQ Orbitrap
© Centroided data	Thermo LTQ Iontrap in Enhanced n	node.
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	
Positive		
Negative	Direct sample analysis Pre-processed data from Progenes	is 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: Profile data Centroided data Resolution (full width at half maximum): Select the ionisation polarity for this analysis: Positive Negative	
Back Next Can	cel

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

Q Create New Experiment	X	QI Create New Experiment	•••
Select the possible adducts To correctly identify the compounds are formed.	in your samples, we need to know how the ions	Select the possible adducts To correctly identify the compounds in your samp are formed.	les, we need to know how the ions
Adducts in library:	Adducts in this experiment:	Adducts in library:	Adducts in this experiment:
2M+K 2M+ACN+H 2M+ACN+Na M+H-2H20 M+H-H20 M+H+G M+N4 M+N4 M+N4 M+CH30H+H M+K	>> << Move adducts into this list to add them to your experiment.	M-H ₂ O-H //>//////////////////////////////////	Move adducts into this list to add them to your experiment.
Edit adduct library	Back Create experiment Cancel	Edit adduct library Back	Create experiment 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 New Adduct	
The adducts listed to experiments.				when creating new		Enter the name and details of the new adduc	t. harge:
Adduct	Adduct Mass	Masses	Charge	Add		2M V +K 38.9632	-
🔁 M+Na	22.9892	1	1	* Remove		2MI +K 58.9032	
H+CH3OH+H	33.0335	1	1				
🔁 М+К	38.9632	1	1	=		Add Car	cel
M+ACN+H	42.0338	1	1		L		
🗄 M+2Na-H	44.9712	1	1				
🗄 M+ACN+Na	64.0158	1	1				
🔁 M+2H	2.0146	1	2				
- M I N	02.0005		-	v			
			Save	Cancel			

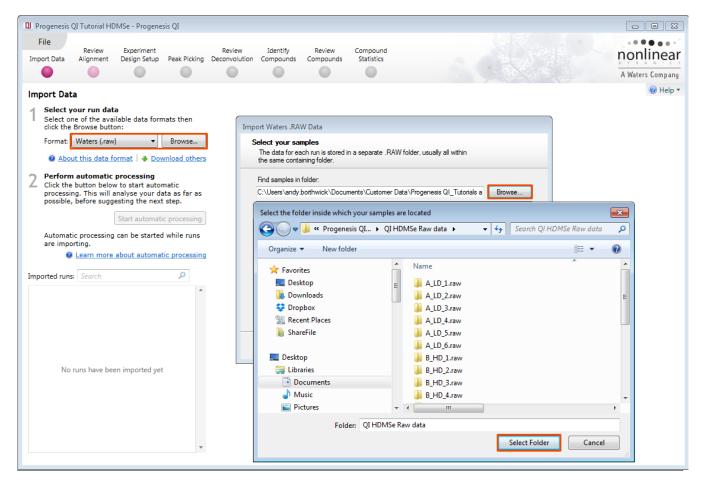
Now select the expected adducts based on you knowledge of your experimental conditions. In this example the polarity was positive and a basic set of Adducts was selected as shown on the right panel below.

dducts in library:			Adducts in this experime	ent:
H+2Na-H		>>	H+H-2H₂O	*
🗄 M+ACN+Na			🚹 M+H-H ₂ O	
🗄 M+2Na		< <	🔁 M+H	
🔁 M+3H	=		🔁 M+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.

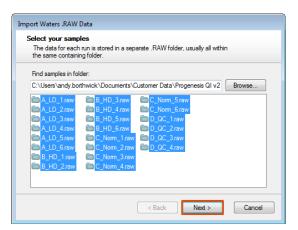
Select the 'Import Data file format', in this example they are Waters (.raw) files

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 samples contain lock mass calibration information. Please provide the calibration m/z.
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 samples are ready to be imported. Please review the options below.
22 samples 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 you will be warned if any of the data files have been 'centroided' during the data acquisition and conversion process.

File				
Review nport Data Alignment	Experiment Design Setup Peak Picking	Review Identify Review Deconvolution Compounds Compounds	Compound Statistics	nonlinea
			0	A Waters Compar
nport Data			<< Actions >>	 Help
Select your run da	ata	A_LD_1 (low energy)		
Select one of the av click the Browse but Format: Waters (.ra	ailable data formats then tton: w) • Browse		n 600 200 300 900 m/z ►	About this run • Low energy peak count: 270,245 • High energy peak count: 131,552 • Total ion intensity: 2.426e+006 • Masked areas: none
About this data	format 🕹 Download other	9 m	경험 수밖에 가지 않는 것이 같다.	- <u>Hakea areas</u> i hone
processing. This will	c processing ow to start automatic analyse your data as far as gesting the next step.		Andreas Marina ang ang ang ang ang ang ang ang ang a	Data import details • Lock mass calibrated └ Lock mass m/z: 556.2766
	Start automatic processing		54 M.	
are importing.	g can be started while runs e about automatic processing	Time 1		
ported runs: Search	٩	- Retention	Zoom: 🔍 🔍 🔀	
	i i i i i i i i i i i i i i i i i i i	A_LD_1 (high energy)		
A LD 1 A L	A LD_3	100 200 300 400 500 S S S S S S S S S S S S S		
		12	1.1	
15. The second s	Pending	8	n an	
A_LD_4 A_L	LD_5 A_LD_6		felm desan desan desan	
Pending Pend	ding Pending	9.45 8.40 7.35 6.30 5.25 Retention Time (min)		
B_HD_1 B_H	HD_2 B_HD_3	45 8.4 etention		

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 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.

Processing Complete Automatic processing complete (with warnings). Time taken: 9 minutes 58 seconds									
 Importing runs: Selecting reference: 	22 of 22 processed D QC 3								
Aligning runs:	21 of 21 processed A 1 run failed to align - continuing without it								
 Creating design: 	Created								
 Peak picking: 	6069 compound ions found								
		Close							

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	ing complete (with 21 seconds	warnings).	
A Importing runs:	22 of 22 processed A 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).

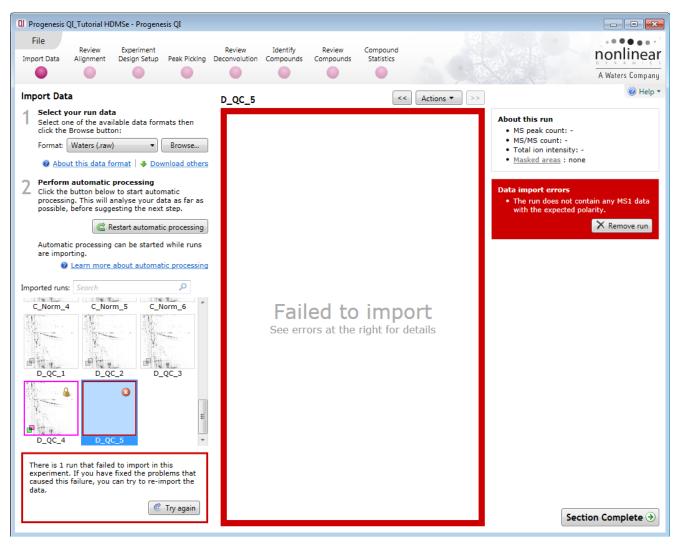
Q Processing Complete		×
Automatic process Time taken: 28 seconds	ing failed.	
A Importing runs:	19 of 19 processed 1 failed to import	
× Selecting reference:	All reference candidates failed to import	
Aligning runs:	Unable to start.	
Creating design:	Unable to start.	
Peak picking:	Unable to start.	
		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 Time taken: 10 minutes 4	ing complete (with warnings). 43 seconds	
A Importing runs:	23 of 23 processed	
	🔺 1 failed to import	
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.

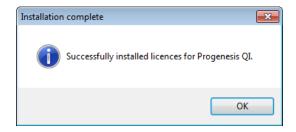
The runs in your experiment will be listed as shown below.

To activate license(s) for the selected runs enter the code in the space provided and click $\mbox{Use Licence code}.$

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

File Review Experiment	Review Identify Review Compound Peak Picking Deconvolution Compounds Compounds Statistics		nonline
	0 0 0 0 0		A Waters Comp
ongle License Runs			
This installation is currently restricted to analyse icensed runs only.	Run name	Licence state	License this run
To license your runs, you need an evaluation	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	
icence code which can be obtained from a sales	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	V
epresentative.	C:\Users\andy.botthwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	V
Once licensed, your runs can be analysed on	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	v
any installation of the software. The licence is	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	
automatically included when archiving an	C:\Users\andy.botthwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	V
experiment.	C:\Users\andy.botthwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	
f your runs have been licensed on another computer, click here to make the licences	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	
available on this computer.	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	V
f vou have one, vou can open a licence file to	C:\Users\andy.botthwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	V
nstall.	C:\Users\andy.borthwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	
f you have just installed a dongle, click here.	C:\Users\andy.bothwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	V
· , · · · , - · · · · · · · · · · ·	C:\Users\andy.botthwick\Documents\Customer Data\Progenesis QI v2\Progenesis QI v2 Tutorial\Progenesi	Unlicensed	V
	$\label{eq:c:Users} C: \label{eq:c:Users} Or v2 \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	Unlicensed	
	Run licence code: xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	Use Lic	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

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.

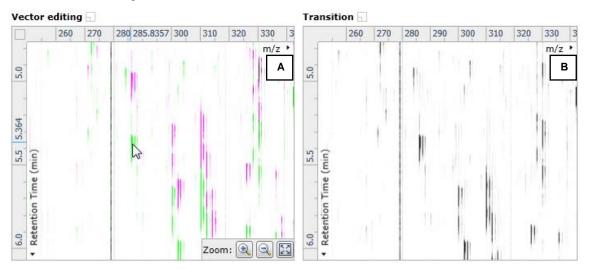
Note: when you click anywhere in the Ion Intensity Map this will **reposition the focus** and update the other views accordingly.

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

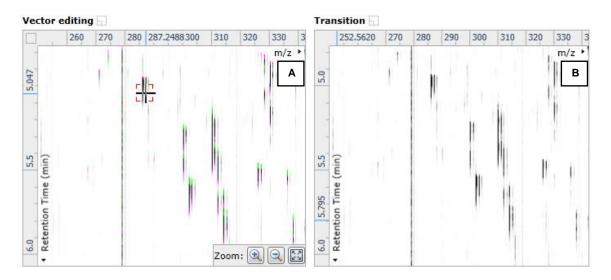
Note: the lon Intensity Map gives you a colour metric, visually scoring of the current alignment and an overall score is placed next to the Vectors column in the table. With each additional vector this score will update to reflect the overall quality of the alignment. The colour coding on the lon intensity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

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

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

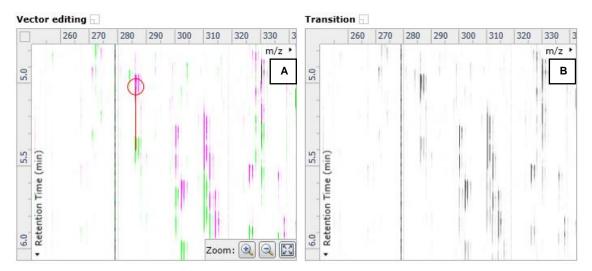


5. As you are holding down the left mouse button (depending on the severity of the misalignment), the alignment vector will automatically find the correct lock. If not, drag the green feature over the corresponding magenta feature of the reference run. The red box will appear as shown below indicating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



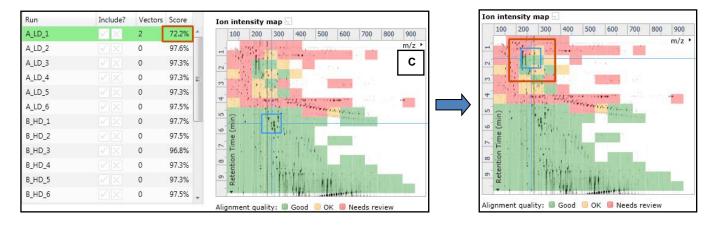
Tip: while holding down the mouse button hold down the **Alt** button. This will allow smooth movement of the cursor as the **Alt** key allows you to override the 'automatic alignment' performed as you depress the mouse button.

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.



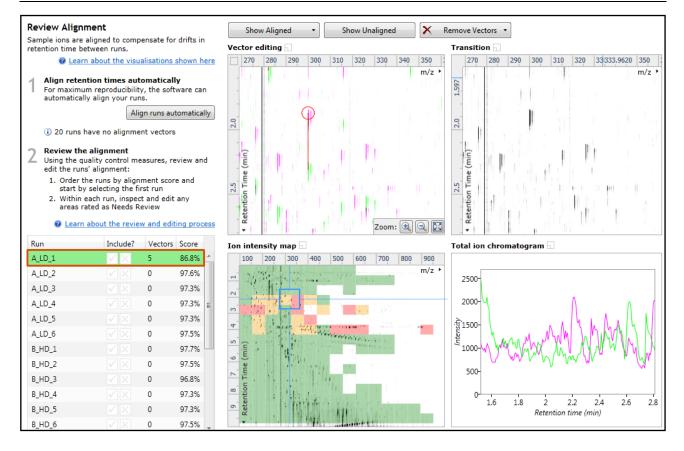
8. Adding an additional vector will improve the alignment further as shown below.



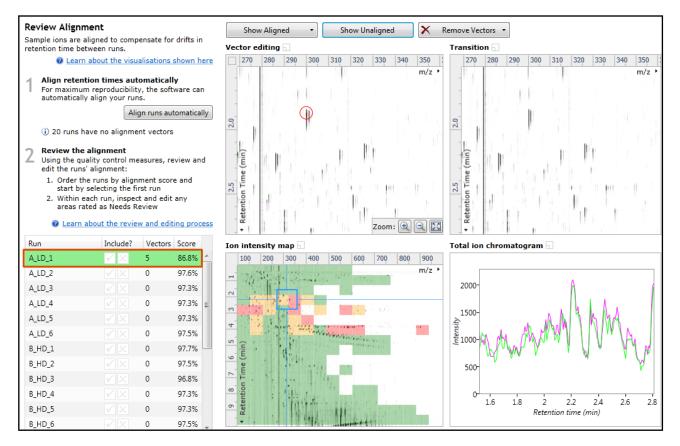
9. Repeat this process moving the focus from top to bottom on the **Ion Intensity Map** view the number you add is dependent on the length of gradient used and the severity of the misalignment. In many cases a single vector per alignment is all that is required to improve the performance of the Automatic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



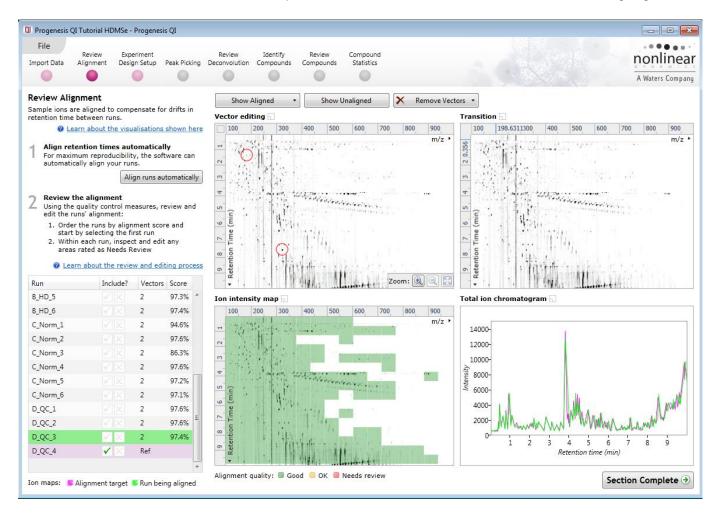
The number of manual vectors that you add at this stage is dependent on the misalignment between the current run and the Reference run. The effect of adding the manual vectors can be seen when you press show aligned



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.



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.

Automatic Ali <u>c</u>	Inment		3
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 🚽	
		OK Cancel]

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.

l Progenesis QI Tutorial HDMSe - Progenesis QI											
File Review Experiment Import Data Alignment Design Setup Peak Pic	Review king Deconvolution C		Review ompounds	Compound Statistics					D	Waters Co	C C
Conditions Minus QC	New									0) Help
Which experiment design type do yo	ou want to use fo	or this expe	riment?								
Do samples from a given subject appear in only one condition? Then	A	Del	lete	0-0 Have you take	n-subject Des n samples from a different conditi	given		Before	During	After	
use the between-subject design.	Q Create Ne	w Experiment D	_	Subject under	×			Delore	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	3 day treat	ne for the exper ment of patients	_	n:	(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.	Group	u want to group the runs manua an existing desig	illy			nis nas	Patient Y	Y1	Y2	Y3	
	Add con	an existing desig	n:	Create desig	gn Cancel	but The	Patient Z	Z1	Z2	Z3	
				Measures ANO A standard ANO because the da	hen perform a rep /A. DVA is not approp ata violates the AN independence. Wi	riate NOVA					I
				repeated meas differences car reduced as a so	ures ANOVA indivi be eliminated or burce of between rences (which hel	idual r					
				thought of as a paired-samples	ject design can be in extension of the s t-test to include tween more than ures.	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.

I Progenesis QI Tutorial HDMSe - Progenesis QI							
File Review Experiment Import Data Alignment Design Setup Peak Picking	Review I Deconvolution Cor	dentify Review mpounds Compou					nonlinea
	•	• •					A Waters Company
Conditions Minus QC 3 day	/ treatment of	patients I ×	C New				🕜 Help
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.	Patient 1	A_LD_1	A_LD_2	A_LD_3	A_LD_4		
Filter samples: 🔎				1999 B. C. C.			
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	Select Sample	Select Sample	Select Sample	•	
C_Norm_4	Add Subject]	
C_Norm_5						ſ	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 LipdBlast databases. Click on the download link.

QI TG_Lipid_Test_pos - Progenesis QI		_ • •
File Import Data Alignment Design Setup Peak Picking	Review Identify Review Compound Deconvolution Compounds Compounds Statistics	A Waters Company
Identify Compounds	Compound 11.39_878.7341n	🕢 Help 🔻
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 Filter the compounds Using the list below, filter the compounds to show only those you want to identify.	Intensity	
2 Set the search parameters Enter the mass error tolerances for matching each compound and its fragments: Precursor tolerance: 5 ppm ▼ Fragment tolerance: 5 ppm ▼ Report isomers as a single search hit @	QI LipidBlast database download Image: Comparison of the	900
3 Search 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.

Downloading LipidBlast from Metabolomics Fiehn La	b
	Cancel

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

	entify Compounds ect your identification me	thod:	
9	LipidBlast		•
	About this method	nod 🖊 Down	load others
1	Filter the compounds Using the list below, <u>filt</u> show only those you wa		inds to
2	Set the search param Enter the mass error to each compound and its	erances for m	atching
	Precursor tolerance:	5	ppm 🔻
	Fragment tolerance:	5	ppm 🔹
	👿 Report isomers as a	single search	hit 🕜
3	Search for identificati After searching, identific to the relevant compour	cations will be	ally.

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.

File Review mport Data Alignment	Experiment Design Setup	Peak Picking	Review Deconvolution	Identify Compounds	Review Compounds	Compound Statistics					nonlinea
											A Waters Compa
dentify Compounds elect your identification			Compour	nd 11.76_8	352.7195n						🔞 Hel
About this m Filter the compoun Using the list below, show only those you Set the search para Enter the mass error each compound and	nds filter the compr want to identifi ameters tolerances for its fragments:	ounds to y. matching	Database Measured		Inten Datab	m/z: 597.4 isity: 11602 base m/z: 5 z: 0.0008	.92 (100.0	00%)			
Precursor toleranc Fragment toleranc Report isomers a Search for identific After searching, iden	e: 5 as a single searc		520 Legend:	Matched	580 600 620 fragment 🔳 Ur			00 720 m/z	740 760 78	0 800 820	840 860 880 900
Fragment toleranc	e: 5 as a single search tifications wounds automat	ppm	Legend: Possible i	Matched dentification	fragment 🔲 Ur ns: 8		agment	m/z		0 800 820	840 860 880 900
Fragment toleranc Report isomers a Search for identific After searching, iden	e: 5 as a single search tifications wounds automat	h hit @	Legend: Possible id	Matched dentification	fragment Ur ns: 8 Description	nmatched fra	agment Adducts	m/z Formula	Retention tir	0 800 820	840 860 880 900
Fragment toleranc Report isomers a Search for identific After searching, iden	e: 5 as a single search tifications wounds automat	ppm	Legend: Possible in Comp TG(16	Matched dentification ound ID 5:0/18:2/18:3)	fragment Ur us: 8 Description TG(16:0/18:2/18	nmatched fra	agment	m/z Formula CssH96O6	Retention tir	0 800 820	840 860 880 900
Fragment toleranc Report isomers a Search for identific After searching, iden to the relevant comp	e: 5 as a single search tifications wounds automat	ppm	Legend: Possible in Comp TG(16 TG(16	Matched dentification round ID 5:0/18:2/18:3) 5:0/16:0/20:5)	fragment Ur ur Ur Ur Ur Ur Ur Ur	nmatched fra 8:3) 0:5)	agment Adducts M+H,	m/z Formula	Retention tir	0 800 820	840 860 880 900
Fragment toleranc Report isomers a Search for identific After searching, iden to the relevant comp No filter applied	e: 5 as a single search cations tifications will b iounds automat Search for ic	ppm ppm part of the second	Legend: Possible in Comp TG(16 Grave TG(16 TG(16	Matched dentification iound ID 5:0/18:2/18:3) 5:0/16:0/20:5) 5:0/16:1/20:4)	fragment Ur us: 8 Description TG(16:0/18:2/18	nmatched fra 8:3) 0:5) 0:4)	Adducts M+H, M+H,	m/z Formula C55H96O6 C55H96O6	Retention tir	0 800 820	840 860 880 900
Fragment tolerance Report isomers a Search for identifice After searching, iden to the relevant comp No filter applied ompound Acce	e: 5 as a single search cations tifications will b iounds automat Search for ic	ppm • h hit @ be assigned ically. dentifications Create	Legend: Possible in ☆ Comp ★ TG(10 ☆ TG(10 ☆ TG(10	Matched dentification iound ID 5:0/18:2/18:3) 5:0/16:0/20:5) 5:0/16:1/20:4) 7:1/17:1/18:3)	fragment Ur Ur Ur Ur Ur Ur Ur Ur	8:3) 0:5) 0:4) 8:3)	Adducts M+H, M+H, M+H,	<i>m/z</i> Formula CssH96O6 CssH96O6 CssH96O6	Retention tir		
Fragment toleranc Report isomers a Search for identific After searching, iden to the relevant comp No filter applied iompound Acce 2.12.11_854.7374n	e: 5 as a single search cations tifications will b iounds automat Search for ic	ppm • h hit @ be assigned ically. dentifications Create	Legend: ☆ Comp ★ TG(10) ☆ TG(10) ☆ TG(10) ☆ TG(10) ☆ TG(10) ☆ TG(10) ☆ TG(10)	Matched dentification iound ID 5:0/16:0/20:5) 5:0/16:1/20:4) 7:1/17:1/18:3) 7:0/17:2/18:3)	fragment Ur IS: 8 Description TG(16:0/18:2/18 TG(16:0/16:1/20 TG(16:1/17:1/18	nmatched fr 8:3) 0:5) 0:4) 8:3) 8:3)	Adducts M+H, M+H, M+H, M+H,	<i>m/z</i> Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention tir		840 860 880 900
Fragment toleranc Report isomers a Search for identific After searching, iden to the relevant comp No filter applied Compound Acce 12.11_854.7374n 1.176_852.7195n	e: 5 as a single search cations tifications will b iounds automat Search for ic	ppm • h hit @ be assigned ically. dentifications Create	Legend: ☆ Comp ★ TG(10) ☆ TG(10)	Matched dentification iound ID 5:0/18:2/18:3) 5:0/16:0/20:5) 5:0/16:1/20:4) 7:1/17:1/18:3) 7:0/17:2/18:3) 5:1/18:1/18:3)	fragment Ur Ur Ur Ur Ur Ur Ur Ur Ur Ur	nmatched fra 8:3) 0:5) 0:4) 8:3) 8:3) 8:3)	Adducts M+H, M+H, M+H, M+H, M+H,	<i>m/z</i> Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention tir		
Fragment toleranc Report isomers a Search for identific After searching, iden to the relevant comp No filter applied Compound Acce 12.11_854.7374n 11.76_852.7195n 11.12_876.7191n	e: 5 as a single search cations tifications will b iounds automat Search for ic	ppm • h hit @ be assigned ically. dentifications Create	Legend: ☆ Comp ★ TG(10) ☆ TG(10)	Matched dentification wound ID 5:0/18:2/18:3) 5:0/16:0/20:5) 5:0/16:1/20:4) 7:1/17:1/18:3) 7:0/17:2/18:3) 5:1/18:1/18:3) 5:1/18:2/18:2)	fragment Ur Ur Ur Ur Ur Ur Ur Ur Ur Ur	nmatched fra 8:3) 0:5) 0:4) 8:3) 8:3) 8:3) 8:3) 8:2)	Adducts M+H ₂ M+H ₂ M+H ₂ M+H ₂ M+H ₂	m/z Formula CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6 CssH96O6	Retention tir		
Fragment toleranc	e: 5 as a single search cations tifications will b iounds automat Search for ic	ppm • h hit @ be assigned ically. dentifications Create	Legend: ☆ Comp ★ TG(10) ☆ TG(10)	Matched dentification wound ID 5:0/18:2/18:3) 5:0/16:0/20:5) 5:0/16:1/20:4) 7:1/17:1/18:3) 7:0/17:2/18:3) 5:1/18:1/18:3) 5:1/18:2/18:2)	fragment Ur Ur Ur Ur Ur Ur Ur Ur Ur Ur	nmatched fra 8:3) 0:5) 0:4) 8:3) 8:3) 8:3) 8:3) 8:2)	Adducts M+H ₁ M+H ₁ M+H ₁ M+H ₁ M+H ₁ M+H ₁	m/z Formula CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6	Retention tir		
Fragment toleranc Report isomers a Search for identific After searching, iden to the relevant comp No filter applied	e: 5 as a single search cations tifications will b iounds automat Search for ic	ppm • h hit @ be assigned ically. dentifications Create	Legend: ☆ Comp ★ TG(10) ☆ TG(10)	Matched dentification wound ID 5:0/18:2/18:3) 5:0/16:0/20:5) 5:0/16:1/20:4) 7:1/17:1/18:3) 7:0/17:2/18:3) 5:1/18:1/18:3) 5:1/18:2/18:2)	fragment Ur Ur Ur Ur Ur Ur Ur Ur Ur Ur	nmatched fra 8:3) 0:5) 0:4) 8:3) 8:3) 8:3) 8:3) 8:2)	Adducts M+H ₁ M+H ₁ M+H ₁ M+H ₁ M+H ₁ M+H ₁	m/z Formula CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6 CssHe6O6	Retention tir		

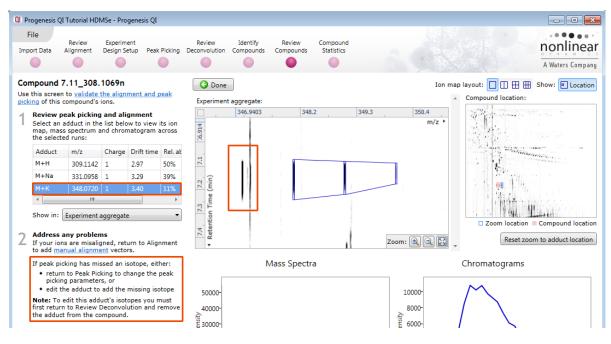
Note: a matched fragment is indicated in red, spanning the x axis. If you hold the cursor over a matched fragment details of the measured and database m/z are reported as well as the mass difference.

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: Editing an adduct to add a missing isotope

When using the Compound Validation view you find an example where the Peak picking has missed an isotope for an adduct ion as shown below. As shown below the monoisotopic peak for the M+K has not been detected.



For the purposes of explaining the processes involved in editing this adduct we have (on purpose) deconstructed the M+K adduct detection.

As the note on the screen suggests you must first return to the **Review Deconvolution** screen and remove the adduct from the compound.

Before doing this, Click **Done** and create a tag for the compound you are about to change. Then filter so that only this compound is displayed at the **Review Compounds** stage.

Q Progenesis QI Tutorial HDMSe - Progenesis QI		
File Import Data Alignment Design Setup Peak Picking	Review Deconvolution Compounds Compounds Compounds	A Waters Company
Review Compounds	Find a compound: Search	🕜 Help 🕈
Using this screen, you can find the compounds of interest in your experiment.	Compound Neutral mass m/z z Retention time Drift time Peak Width Tag 🔻	Accepted ID Identifica
 Create a shortlist to review In the table, sort and filter the compounds based on their measurements, to generate a shortlist for further review. Whow are the measurements calculated? To sort the table by a given value, simply click the relevant column header. 	• 7.11_308.1069n 308.1069 347.0700 1 7.11 3.40 0.18 • Image: Create a filter Image: Create a filter Image: Create the filter, For more guidance, please see the online reference. Image: Create the filter, For more guidance, please see the online reference.	4702 1
2 Review the compounds For each compound of interest, inspect the ions' alignment and peak picking: Review selected compound	Available tags: Available tags: Not identified (5086 compounds) Anova p-value ≤ 0.05 (1243 compounds) My Compound (1 compound)	Þ
You can also double-click to review a compound. 3 Choose the correct identifications For each compound, select one of its possible identifications as the accepted one.	● Max fold change ≥ 2 (444 compounds) ● Separated by drift time (130 compounds) ● S.Plot A (190 compounds) ● S.Plot B (47 compounds)	B_HD I►I
To speed this up, you can automatically accept identifications in compounds where only one of the possible identifications has: Score ≥ 50.0 Accept identifications	Confirmed Compounds (9 compounds) Hide compounds that have any of these tags:	
4 Export data for further processing By exporting your data to external tools, there's no limit to your analysis.	Clear the filter OK Cancel	
Export to pathways tool	₩102	
Experiment design Review your data from a different perspective:	*	Þ
Current design: Minus QC 🔹	Sec	tion Complete 🏵

Now click Review Deconvolution on the workflow

In this view the 3 Adducts are displayed. Right click on the M+K adduct and select **Remove from Compound**.



The table will update to show 2 compounds (both tagged). The second one is based on the single adduct M+K.

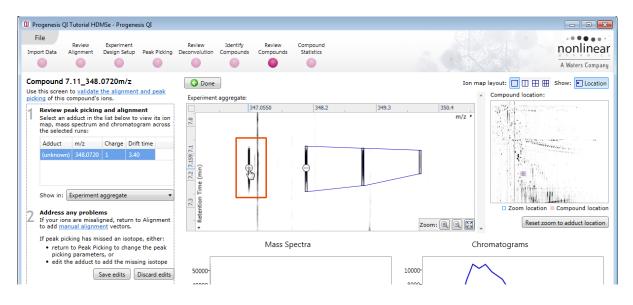


Now click on Review Compounds.

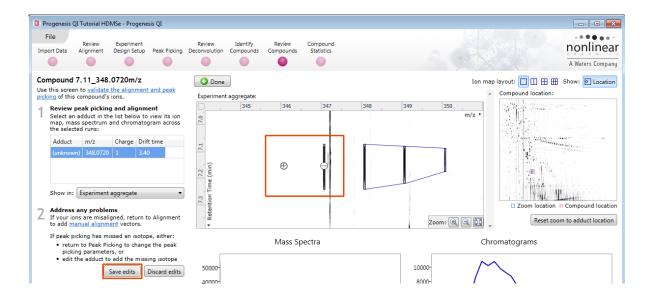
Select the compound with the 'unknown' Neutral Mass and click Review selected compound

File Import Data	Review Experiment a Alignment Design Setup Peak Picking I	Review Identify Deconvolution Compoun										n	onlinea Waters Compan
	Compounds	Find a compound: See	arch		p	Filter cor	npounds	 Filter is a 	active				🕜 Help
	screen, you can find the compounds of your experiment.	Compound	Neutral mass	m/z	z	Retention time	Drift time	Peak Width	Tag 💌	Accepted ID	Identifications	Anova (p)	Max fold chan
1 Creat	te a shortlist to review	• 7.11_348.0720m/z		348.0720					\$				
on the	e table, sort and <u>filter the compounds</u> based eir measurements, to generate a shortlist rther review.	• 7.11_308.1069n	308.1069	309.1142	1	7.11	2.97	0.18	•	4702	1	4.44E-16	1.09E+04
	How are the measurements calculated?												
To so the re	rt the table by a given value, simply click elevant column header.	٠											Þ
	ew the compounds	Compound 7.11_348.	0720m/z:										
	ach compound of interest, inspect the ions' ment and peak picking:	Compound abundan	ce Possible id	dentification	ns	3D Montage D	rift time mo	ontage					
	Review selected compound		C No	rm	-			A_LD			BH	-D	

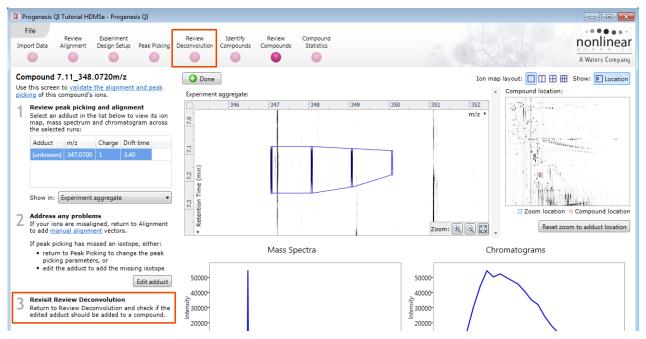
Now click on **Edit Adduct**. A **plus and minus** will appear on the experiment aggregate view. Click on the **plus** to add the missing isotope.



The view updates to show the minus sign now on the added isotope. Click on Save edits to accept the edit



The view will update to show the added isotope and a note will appear indicating that you should return to Review Deconvolution to check if the edited adduct should be added to a compound. Click on **Review Deconvolution** on the workflow.



Right click on the edited adduct and select Add to Compound

Review Deconvolution	Compound	7.11_308.106	9n (4702)				Grid size:	🗆 🗌 🛈 Abo	Welp ▼ ut this
After peak picking, ions are grouped by compound in a process called deconvolution.	+H+Na	M+H-2H ₂ O	M+H-H ₂ O	M+H	M+NH4	M+Na	M+CH₃OH+H	M+K	M+ACN+H
How does deconvolution work? Review the compounds Select compounds below to view: areas of the ion map showing the different adduct forms the mass and retention time profiles of their ions	1						dd to Compound dd to Clip Gallery		

The compound and its details will update to and the compound table will now show a single compound.

QI Progenesis QI Tutorial HDMSe - Progenesis QI		
File Review Experiment Import Data Alignment Design Setup Peak Picking	Review Identify Deconvolution Compounds	nonlinear A Waters Company
Review Deconvolution	Compound 7.11_308.1069n (4702)	Grid size: 🔹 🗆 🗌 🛈 About this
After peak picking, ions are grouped by compound in a process called deconvolution. <u>O How does deconvolution work?</u>	-H+Na M+H-2H2O M+H-H2O M+H M+NH4	M+Na M+CH3OH+H M+K M+A
 Review the compounds Select compounds below to view: areas of the ion map showing the different adduct forms the mass and retention time profiles of their ions These views can be used to confirm the validity of your interesting compounds. Wow do I use this screen? 		
2 Optimise for your samples If any compounds have an ion whose profile doesn't match the majority in the compound, it can be removed. Likewise, missing compound ions can be added.	۲ ساله ۲۰۰۰ Mass Spectra	Chromatograms Peak
Tag filter applied compounds may be hidden Edit Compound Accepted ID Tag ▼ Ad • 7.11_308.1069n 4702 №	100- 100-	

Now click on **Review Compounds** and click **Review selected compound** with the edited compound selected.

When you now click on the M+K adduct it contains 4 isotopes and its m/z has altered accordingly to reflect the addition of the isotope.



Appendix 8: 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 local licensed copy, for licensing details please visit the IPA website.

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.

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		A Waters Company
Review Compounds Using this screen, you can find the compounds of interest in your experiment.		r compounds Filter is active	Help ▼ Accepted ID Ide
In the table, sort and <u>filter the com</u> on their measurements, to generat for further review. <u>O How are the measuremer</u> To set the table by a given value	hways Information (thways tool thways tool from the list below. You can find out more or download no g the links below. vays tool do you want to use?	ew 0.14 @ 0.30 @ 0.27 @	49854487 4 * 14708992 1 E 9357 1 7847292 1 7847488 1 *
	ity Pathway Analysis 🕜 About this plugin 🛛 🐳 Download othe	er plugins ntage	Þ
You can also double-click to review Choose the correct identificatio For each compound, select one of i identifications as the accepted one.		•	•
To speed this up, you can <i>automati</i> identifications in compounds where the possible identifications has: Score ≥ 50.0 Accept i			
4 Export data for further processi By exporting your data to external no limit to your analysis.		Cancel	
Export to pathways tool			
Experiment design Review your data from a different perspective: Current design: Minus QC •	4	Section	► 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.

QI 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)	
Select the identifier type used for compounds in the experiment:	
Human Metabolome Database (HMDB)	
CAS Registry Number	1
KEGG Human Metabolome Database (HMDB)	
PubChem CID	11
	"
< Back Export compounds to IPA Close]

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)
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. B_HD C_Norm A_LD B_HD
< Back 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 .

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

	(/apps.ingenuity.c ♀ ← ▲ ♂ Ingenuity – Ingenuity Login ×	- • × A t 3
	NUITY*	
Welcome! Email Password	Please login smith @work.com Remember my password LOG IN Sign Up Forgot Password	Contact Customer Support 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
<		>

IPA will open displaying the imported data from Progenesis.

<u>E</u> dit <u>V</u> iew <u>W</u> indow	<u>H</u> elp		Provide Feedback Support Janucz Nykiel Clos
	Genes and Chemicals	Diseases and Functions Pathways and Tox Lists	
N ¥	Enter gene names/symbol	ls/IDs or chemical/drug names here	SEARCH Advanced Search 🔽
set Upload - New Datas	et 2015-05-11 07:35 AM		Dataset Upload Workflow Instructions
Select File Format:	Ingenuity File Format A or B	 More Info 	
Contains Column Header	: O Yes No		Data Upload Workflow
Select Identifier Type:	Human Metabolome Database (H	MDB) Specify the identifier type found in the dataset.	Use Dataset Upload to import your dataset file into IPA.
		mbb) • specify the identifier type round in the dataset.	Once uploaded, many different analysis options exist including Biomarker Filter, Molecular Tox and Core Analyses. Review
Array platform used for e	xperiments: Not specified/applicable	 Select relevant array platform as a reference set for data analysis 	5. the different type of analyses and see which one best fits
Jse the dropdown menu	s to specify the columns that contain identifiers	and observations. For observations, select the appropriate expression value ty	your needs.
			1. To upload a dataset file, click here.
aw Data (567) \ Dataset S	ummary (182)		The spload a decase may <u>click merc</u> .
			🛞 Open
ID	Observation 1		Look (n: 🙆 Multiple Rank 🗸 🕼 🏠 🕼 😫 🚍
			Batch (mult-timepoint)
	Fold Change 🔻		Affy_with_p-value_and_fold.txt
1 HMDB01988	Infinity		LL_with_Fold and_Normalized.txt
2 HMDB05015	1.72884040538655		
3 HMDB40639	3.98236765102672		
4 HMDB40285	2.1144344578431		
5 HMDB39847	1.91509320582337		
6 HMDB05049 7 HMDB30934	1.36771507748685 -2.37287619375142		File Name:
8 HMDB37820	1.69294555757503		Files of Type: All Files
9 HMDB15294	-1.4569308616069		Open Cancel
10 HMDB39350	-1.46035570178035		
11 HMDB41326	1.89049486451889		2. Calest the dataset file from your computer and disk the
12 HMDB34382	1.75866062252427		 Select the dataset file from your computer and click the Open button.
13 HMDB35337 14 HMDB36734	3.98452117071048 2.38572167550178		
14 HMDB36734 15 HMDB13809	-1.25125092715796		3. Select Flexible format for the file format from the
16 HMDB14821	1.08777276460426		dropdown menu.
17 HMDB00782	1.14588290027386		4. Select an Identifier Type from the dropdown menu. IPA
18 HMDB40806	1.83433974738324		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 22 HMDB37712	1.17037352363088 1.50281207491669		the most appropriate one. If more than one type of identifier
22 HMDB37712 23 HMDB38057	1.43917459561637		exists in your dataset, select all appropriate ones.
24 HMDB40285	1.84463305905821		GenBank
25 HMDB02725	1.80218245845514		Select All
26 HMDB29978	2.11110785701812		
27 HMDB37272	1.27732135837318		Affymetrix SNP ID
28 HMDB39848	1.14877655967276		Agilent
29 HMDB29828 30 HMDB30932	-1.35134004957123 1.82757219666658		Applied Biosystems
30 HMDB30932 31 HMDB13300	1.38248130204499		CAS Registry Number
32 HMDB32791	1.21282984334921		
			Entrez Gene (LocusLink)

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 9: 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?

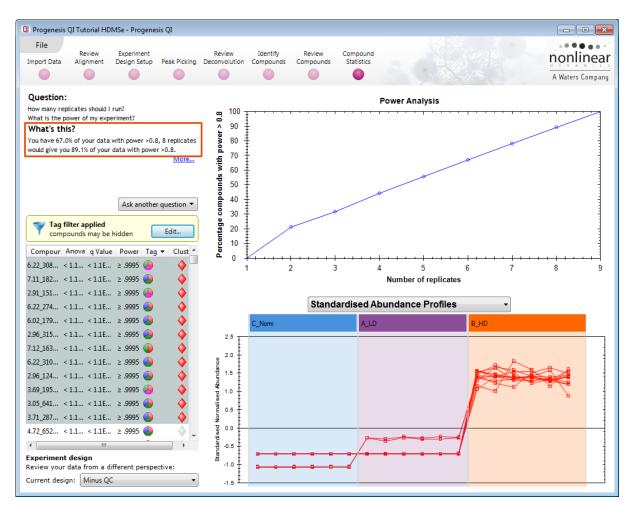
Select the option

'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. 1243 compounds),** as an example, view the power analysis.



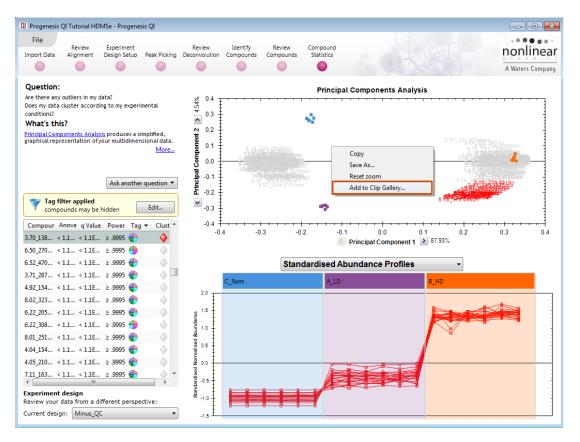
This is displayed graphically showing that 67.0% of the 1243 compounds have a power of 80% or that 8 replicates would give you 89% of your data with power > 0.8.

Appendix 10: 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....



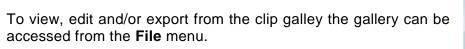
This will open a dialog displaying what is to be saved and allows you to alter the title and provides a description of the image for later reference.

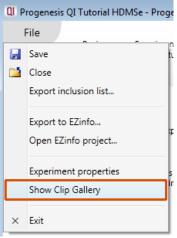
Enter a title and description for this clipping:

I Add to Clip Gallery

Enter details as required and click Add to clip gallery

×





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

Q Clip Gallery							
Clip Gallery	Item siz	ee: • 🗖 🗋 Search 🔎					
Actions Edit Copy Imag Copy Title Copy Organization Export Image Delete	rogenesis QI	Data analysis performed using: Progenesis QI www.nonlinear.com					
Compounds from EzInfo biplot PCA Graph Imported Biplot details from EZInfo	Progenesis QI logo A high-resolution image that you can use in presentations and posters to show the software was used in your research.	Progenesis QI analysis badge A high-resolution image that you can use in presentations and posters to show the software was used in your research.					
		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.

Appendix 11: Waters Machine Specifications and approx. timings

This appendix provides information on the time(s) taken at each stage and the total time taken to analyse a set of 22 (Phase 1) HDMSe runs on a Waters Demo Spec PC.

Machine Spec: Lenovo							
Processor: Intel® Xeon® CPU 2.66GHz 12core X5650 @ 2.67GHz							
RAM: 24.0 GB							
System Type: 64-bit Operating System							
File Folder Size: Each file folder (.RAW): 0.75 Gig							
Analysis Stages:	Per file		Total				
Auto processed:	Import and analysis of data to peak picking		30.0min				
Import Data:	Loading of Raw data per file	1min	22min Total for 22 files				
	(re-opening at Import Data)	5s					
Alignment:	Automatic alignment of data (re-opening at Alignment)		6.0min 10s				
Peak Detection:	Automatic Detection of data		4.5min				
	(re-opening at Peak Detectio	n)	10s				
Identify Compounds: Performing .SDF Compound Search		Search	5s				
	(re-opening at Identify Comp	ounds)	10s				
Total Analysis Time Ir		32.5min					
Restoring Progenesis QI tutorial HDMSe experiment from archive			20s				