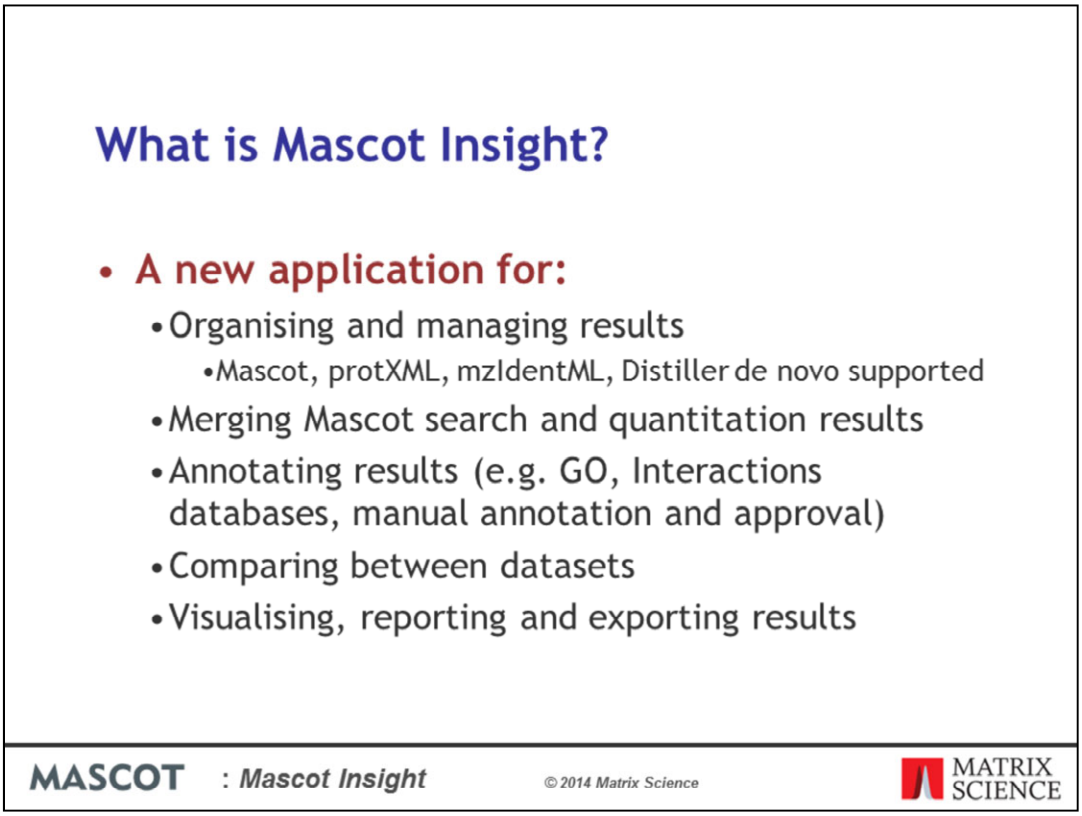
Mascot Insight



What is Mascot Insight?

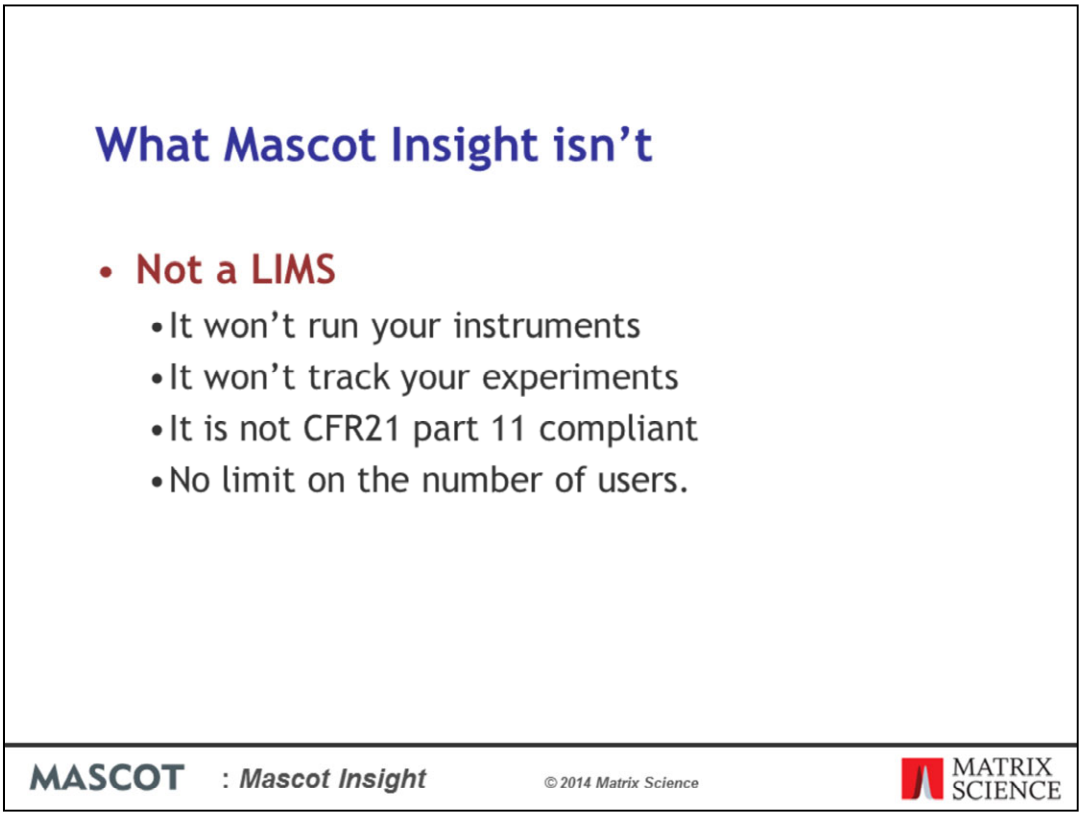
- A new application for:
  - Organising and managing results
    - Mascot, protXML, mzIdentML, Distiller de novo supported
  - Merging Mascot search and quantitation results
  - Annotating results (e.g. GO, Interactions databases, manual annotation and approval)
  - Comparing between datasets
  - Visualising, reporting and exporting results

Mascot Insight is a new application designed to help you to organise and manage your Mascot search and quantitation results.

Mascot Insight provides ways to flexibly merge your Mascot search and quantitation results, including MS-1 based quantitation data from Mascot Distiller, such as SILAC data, and MS2 based quantitation such as iTRAQ and TMT.

Mascot Insight allows you to further annotate your results, including areas such as Gene Ontology annotation of results, use of molecular interactions databases and manual annotation and approval of protein hits

And Mascot Insight provides a wide range of reports covering areas such as comparing datasets, quantitation analysis, plotting charts etc, and provides exports for these to allow you to easily share those results and to export data in machine readable formats



What Mascot Insight isn’t

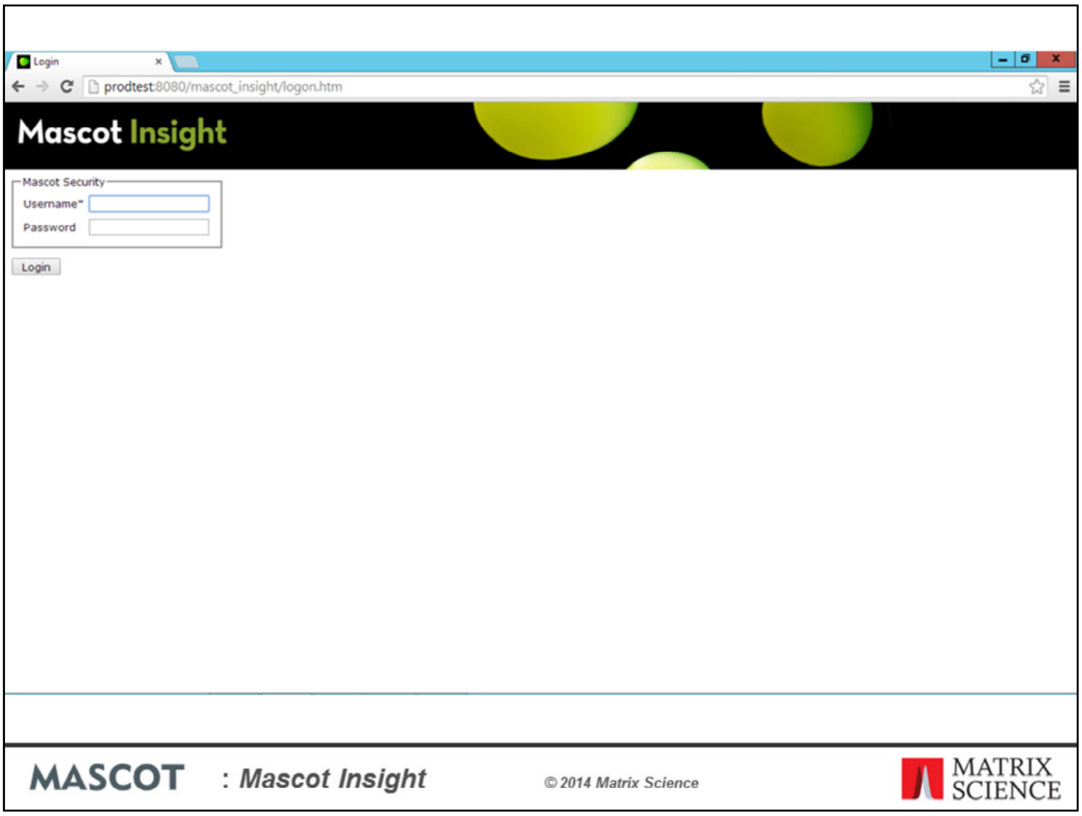
- Not a LIMS
  - It won’t run your instruments
  - It won’t track your experiments
  - It is not CFR21 part 11 compliant
  - No limit on the number of users.

MASCOT : *Mascot Insight*

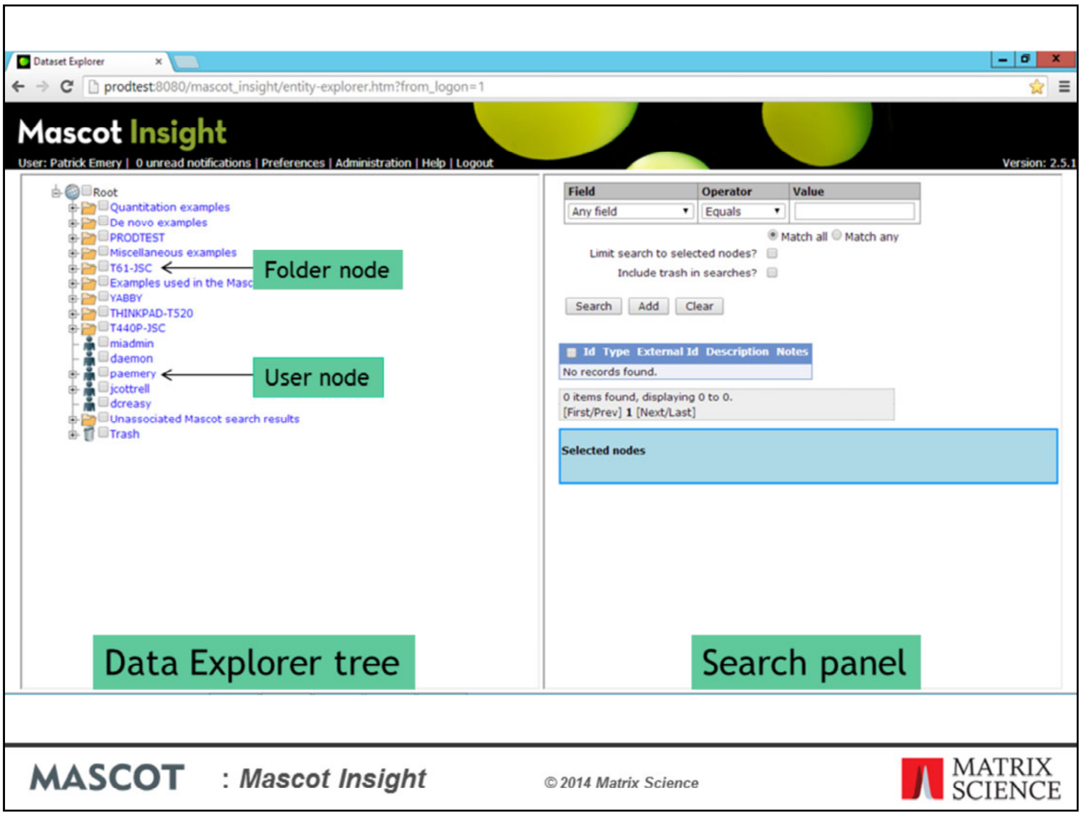
MATRIX SCIENCE

It is also important to mention one thing that Mascot Insight isn’t. It is not a LIMS system. Although you can use it to organise your results and report on them, and it also allows you to record details about your experiments if you wish, it will not carry out many of the functions you would associated with a LIMS such as instrument integration, sample tracking and FDA regulatory compliance.

Mascot Insight is not licensed ‘per seat’, and there is no limit on the number of users that can access the system. The price of Mascot Insight depends instead on the number of CPUs licensed for the associated Mascot server



Mascot security provides a single login to both your Mascot server and to the Mascot Insight web application to identify users and to control user access levels.



- ☐ Root
  - ☐ Quantitation examples
  - ☐ De novo examples
  - ☐ PRODTEST
  - ☐ Miscellaneous examples
  - ☐ T61-JSC
  - ☐ Examples used in the Masc
  - ☐ YABBY
  - ☐ THINKPAD-T520
  - ☐ T440P-JSC
  - ☐ miadmin
  - ☐ daemon
  - ☐ paemery
  - ☐ jcottrell
  - ☐ dcreasy
  - ☐ Unassociated Mascot search results
  - ☐ Trash

Folder node

User node

Field	Operator	Value
Any field	Equals	

● Match all ○ Match any

Limit search to selected nodes? ☐

Include trash in searches? ☐

☐	Id	Type	External Id	Description	Notes
No records found.					

0 items found, displaying 0 to 0.
[First/Prev] **1** [Next/Last]

Selected nodes

Data Explorer tree

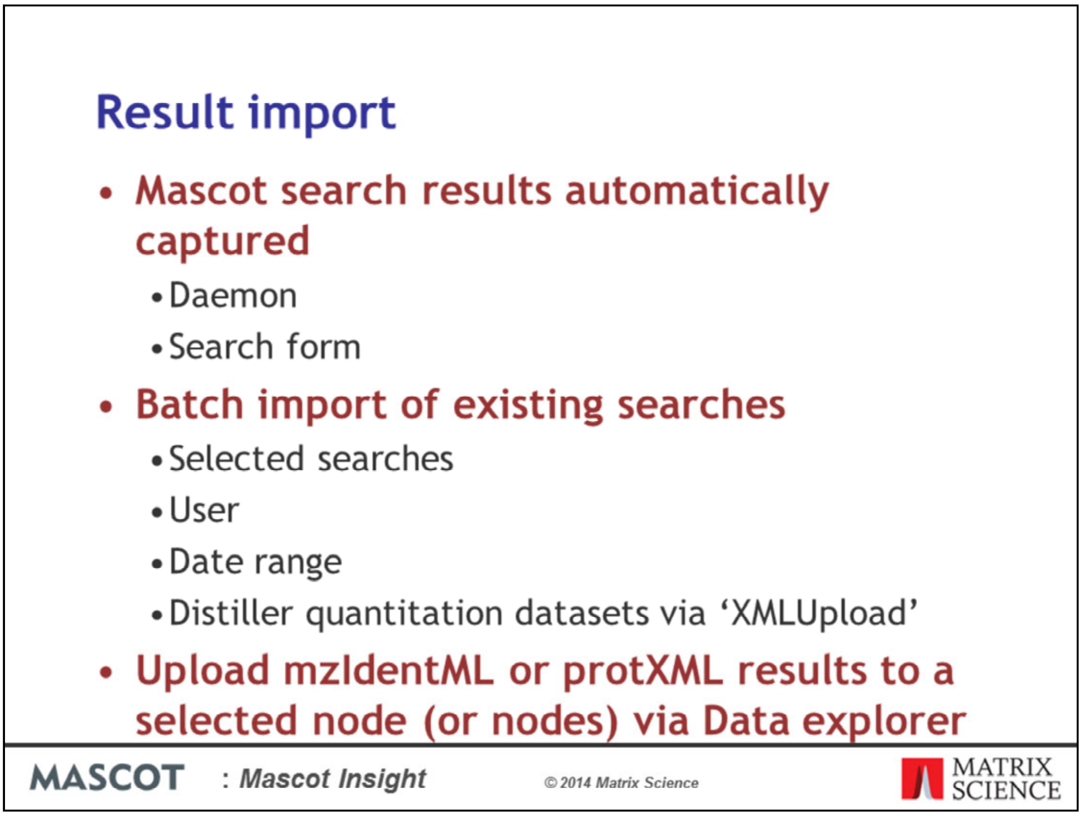
Search panel

MASCOT : *Mascot Insight* © 2014 *Matrix Science* MATRIX SCIENCE

Once you’ve logged in, this is the home page of the application. On the left we have the Data explorer tree. This is where the data available in the system is organised and displayed, with different types of node for storing different types of information. Data can be manually position on the tree, or structures can be automatically generated to your own specifications

On the current view, we can see two main types of node – User nodes and Folder nodes. Each user is given a User node automatically, and this is for them to organise their own data and experiments. The folder node is the simplest form a grouping node and is simply used to group together other related nodes on the tree. There are other node types available in the system, including Project, Study, Experiment, Sample and Result type nodes. These different node types allow you to record different pieces of information (meta-data), and you can always add your own values.

On the right hand side of the home page, we have the ‘Search panel’ which allows us to rapidly search the tree to find a particular node or set of nodes. Security is role-based and handled by Mascot security. Exactly what you see on the dataset explorer tree when you log in depends on your users roles, and on what access settings have been given to particular nodes on the tree by their owners.



Result import

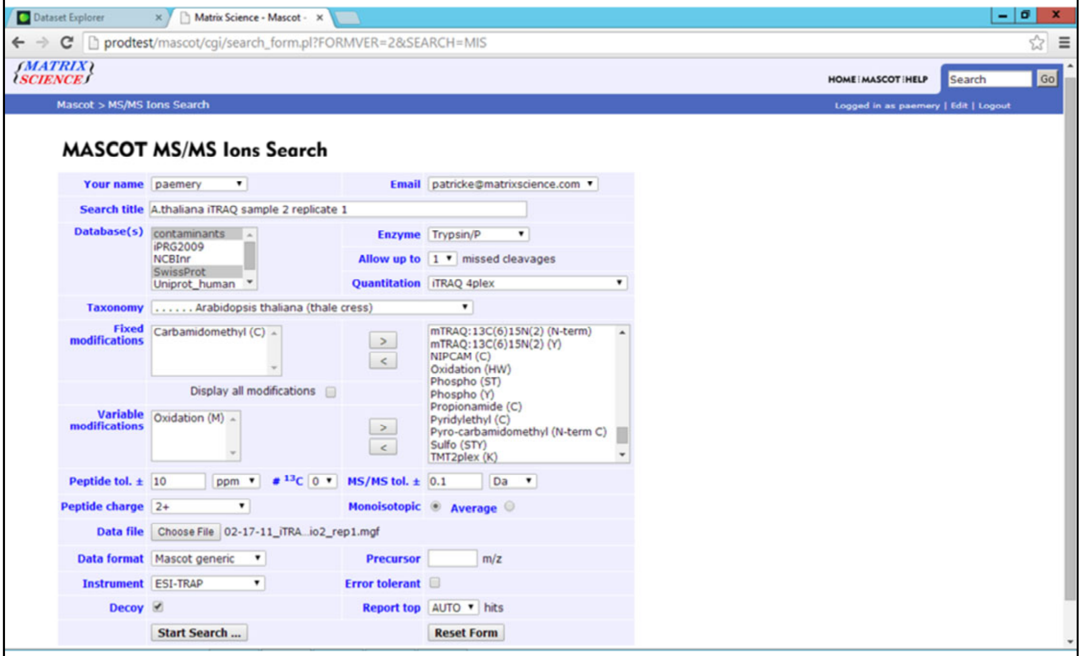
- **Mascot search results automatically captured**
  - Daemon
  - Search form
- **Batch import of existing searches**
  - Selected searches
  - User
  - Date range
  - Distiller quantitation datasets via ‘XMLUpload’
- **Upload mzIdentML or protXML results to a selected node (or nodes) via Data explorer**

MASCOT : *Mascot Insight* *© 2014 Matrix Science* MATRIX SCIENCE

Mascot search results are automatically captured and imported into the system. For the highest level of integration, we’d recommend using Mascot Daemon to submit your searches, and Mascot Distiller for your raw data reduction and MS-1 based quantitation. Mascot Daemon comes with your Mascot server, but Mascot Distiller is a separate product. However, neither is required and you can simply submit a search through the standard search form, or 3rd party client system, if you wish.

You can also carry out a batch import of existing search results from your Mascot server. You can easily limit this to a selected set of searches, to the results for a selected user or by date range. If you have existing Mascot Distiller based quantitation datasets, Mascot Insight comes with a utility ‘XmlUpload’ which will allow you to easily import the required search results and Distiller quantitation data in a hierarchy based on the file hierarchy from the file system.

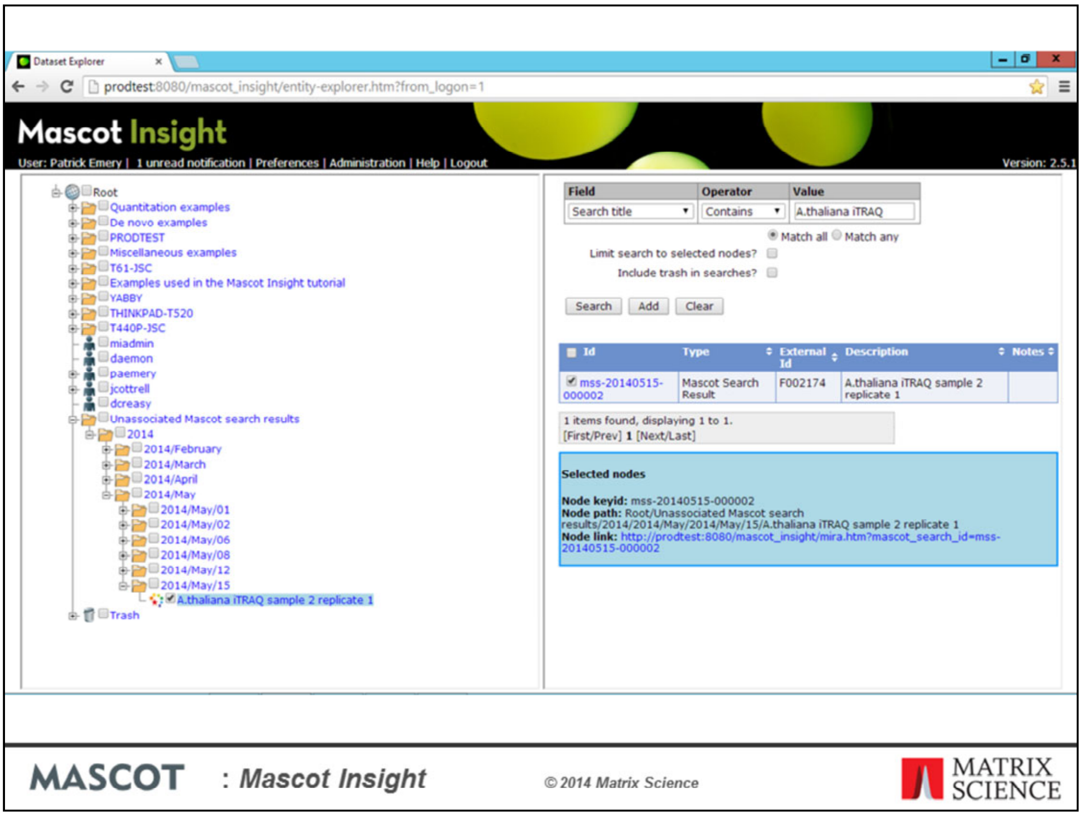
Other result types, such as data in the mzIdentML or protXML formats, can be uploaded to a selected node (or set of nodes) via the data explorer. Multiple files can be zipped together and the archive uploaded if you have multiple searches you want to import in a batch.



MASCOT MS/MS Ions Search

Your name paemery
Email patricke@matrixscience.com
Search title A.thaliana iTRAQ sample 2 replicate 1
Database(s) contaminants, iPRG2009, NCBInr, SwissProt, Uniprot_human
Enzyme Trypsin/P
Allow up to 1 missed cleavages
Quantitation iTRAQ 4plex
Taxonomy . . . . . . Arabidopsis thaliana (thale cress)
Fixed modifications Carbamidomethyl (C)
mTRAQ:13C(6)15N(2) (N-term), mTRAQ:13C(6)15N(2) (Y), NIPCAM (C), Oxidation (HW), Phospho (ST), Phospho (Y), Propionamide (C), Pyridylethyl (C), Pyro-carbamidomethyl (N-term C), Sulfo (STY), TMT2plex (K)
Display all modifications ☐
Variable modifications Oxidation (M)
Peptide tol. ± 10 ppm **# ^{13}C** 0
MS/MS tol. ± 0.1 Da
Peptide charge 2+
Monoisotopic ● **Average** ○
Data file Choose File 02-17-11_iTRA...io2_rep1.mgf
Data format Mascot generic
Precursor m/z
Instrument ESI-TRAP
Error tolerant ☐
Decoy ☑
Report top AUTO hits
Start Search ...
Reset Form

Here I’m carrying out a search of an Arabidopsis iTRAQ 4-plex dataset against the contaminants and SwissProt database on a Mascot server which is registered with a Mascot Insight server. I just submit the search as normal…



Field	Operator	Value
Search title	Contains	A.thaliana iTRAQ

● Match all ○ Match any

Limit search to selected nodes? ☐

Include trash in searches? ☐

☐ Id	Type	External Id	Description	Notes
☑ mss-20140515-000002	Mascot Search Result	F002174	A.thaliana iTRAQ sample 2 replicate 1	

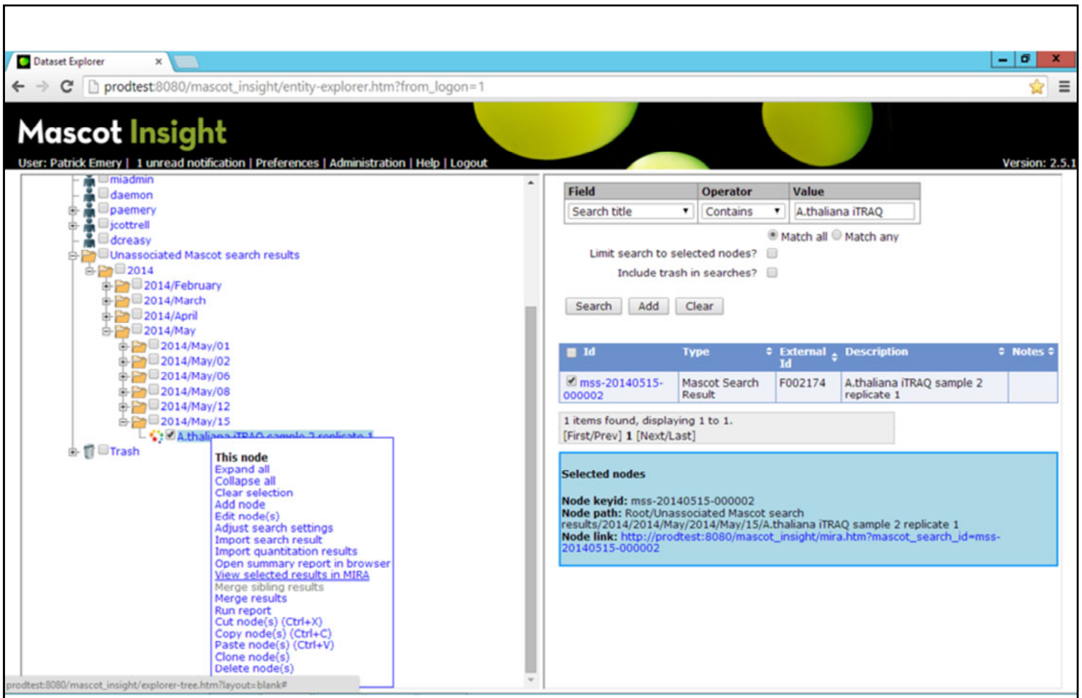
1 items found, displaying 1 to 1.
[First/Prev] **1** [Next/Last]

Selected nodes

Node keyid: mss-20140515-000002
Node path: Root/Unassociated Mascot search results/2014/2014/May/2014/May/15/A.thaliana iTRAQ sample 2 replicate 1
Node link: http://prodtest:8080/mascot_insight/mira.htm?mascot_search_id=mss-20140515-000002

Shortly after the search has completed, the result will also be available in Mascot Insight. In this case the result has been imported to the default location under the ‘Unassociated Mascot search results’ node in a date folder. I could move the result to a different location on the tree by either dragging and dropping it, or by using the cut and paste options from the dataset explorer tree’s context menu. However, you can also set up filters to import the result into a tree structure of your own choosing. These filters can automatically create new folders on the data explorer tree as required. We’ll take a look at an example of this later on in the presentation.

If you’re not sure where a search result has gone on the tree, then the tree is fully searchable using any available field – for example, on the search title



Field	Operator	Value
Search title	Contains	A.thaliana iTRAQ

● Match all ○ Match any

Limit search to selected nodes? ☐

Include trash in searches? ☐

☐ Id	Type	External Id	Description	Notes
☑ mss-20140515-000002	Mascot Search Result	F002174	A.thaliana iTRAQ sample 2 replicate 1	

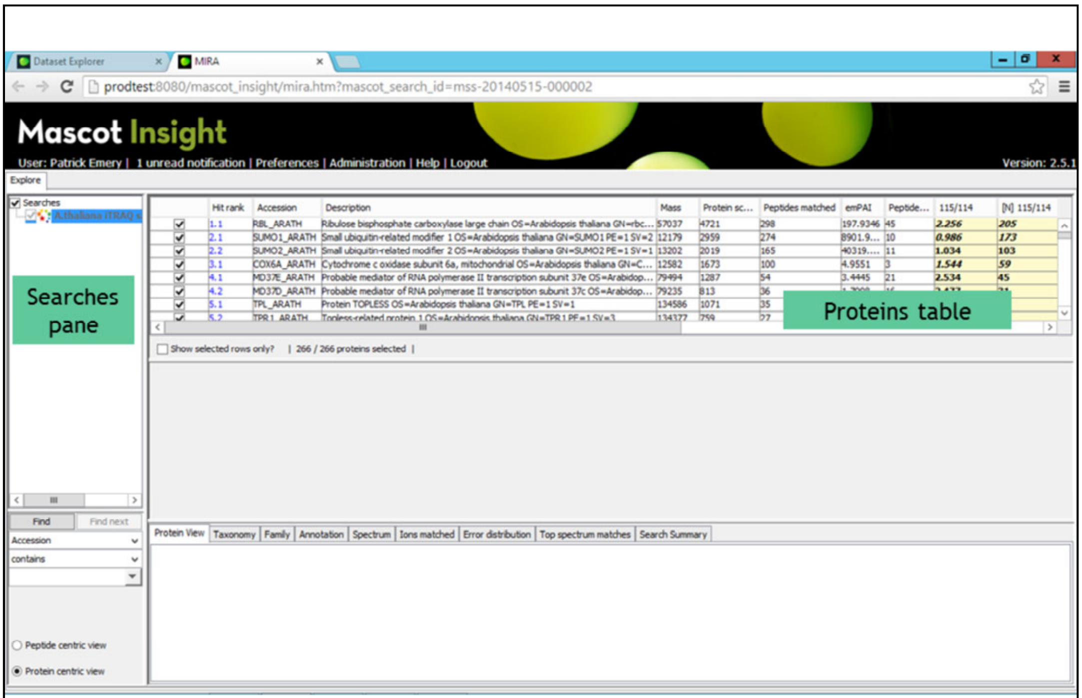
1 items found, displaying 1 to 1.
[First/Prev] **1** [Next/Last]

Selected nodes

Node keyid: mss-20140515-000002
Node path: Root/Unassociated Mascot search results/2014/2014/May/2014/May/15/A.thaliana iTRAQ sample 2 replicate 1
Node link: http://prodtest:8080/mascot_insight/mira.htm?mascot_search_id=mss-20140515-000002

MASCOT : *Mascot Insight* © 2014 Matrix Science MATRIX SCIENCE

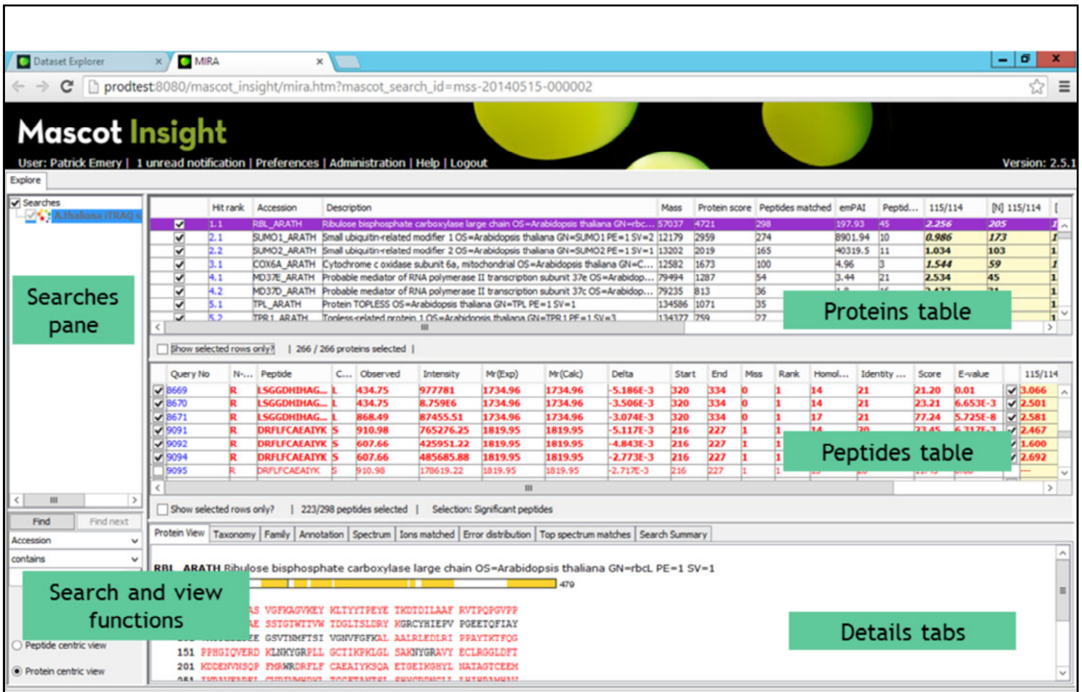
If I right click over a node, or a group of selected nodes, then I get a pop up context menu which allows me to carry out a number of operations, including editing the node, editing default search settings, cut, copy clone and paste operations. For a search result, I also get the option of viewing the original summary report in the browser, or to view the selected result in MIRA – the Mascot Insight Result Applet



	Hit rank	Accession	Description	Mass	Protein sc...	Peptides matched	emPAI	Peptide...	115/114	[N] 115/114
☑	1.1	RBL_ARATH	Ribulose bisphosphate carboxylase large chain OS=Arabidopsis thaliana GN=rbc...	57037	4721	298	197.9346	45	***2.256***	***205***
☑	2.1	SUMO1_ARATH	Small ubiquitin-related modifier 1 OS=Arabidopsis thaliana GN=SUMO1 PE=1 SV=2	12179	2959	274	8901.9...	10	***0.986***	***173***
☑	2.2	SUMO2_ARATH	Small ubiquitin-related modifier 2 OS=Arabidopsis thaliana GN=SUMO2 PE=1 SV=1	13202	2019	165	40319....	11	**1.034**	**103**
☑	3.1	COX6A_ARATH	Cytochrome c oxidase subunit 6a, mitochondrial OS=Arabidopsis thaliana GN=C...	12582	1673	100	4.9551	3	***1.544***	***59***
☑	4.1	MD37E_ARATH	Probable mediator of RNA polymerase II transcription subunit 37e OS=Arabidop...	79494	1287	54	3.4445	21	**2.534**	**45**
☑	4.2	MD37D_ARATH	Probable mediator of RNA polymerase II transcription subunit 37c OS=Arabidop...	79235	813	36				
☑	5.1	TPL_ARATH	Protein TOPLESS OS=Arabidopsis thaliana GN=TPL PE=1 SV=1	134586	1071	35				
☑	5.2	TPR1_ARATH	Topless-related protein 1 OS=Arabidopsis thaliana GN=TPR1 PE=1 SV=3	134377	759	27				

☐ Show selected rows only? | 266 / 266 proteins selected |

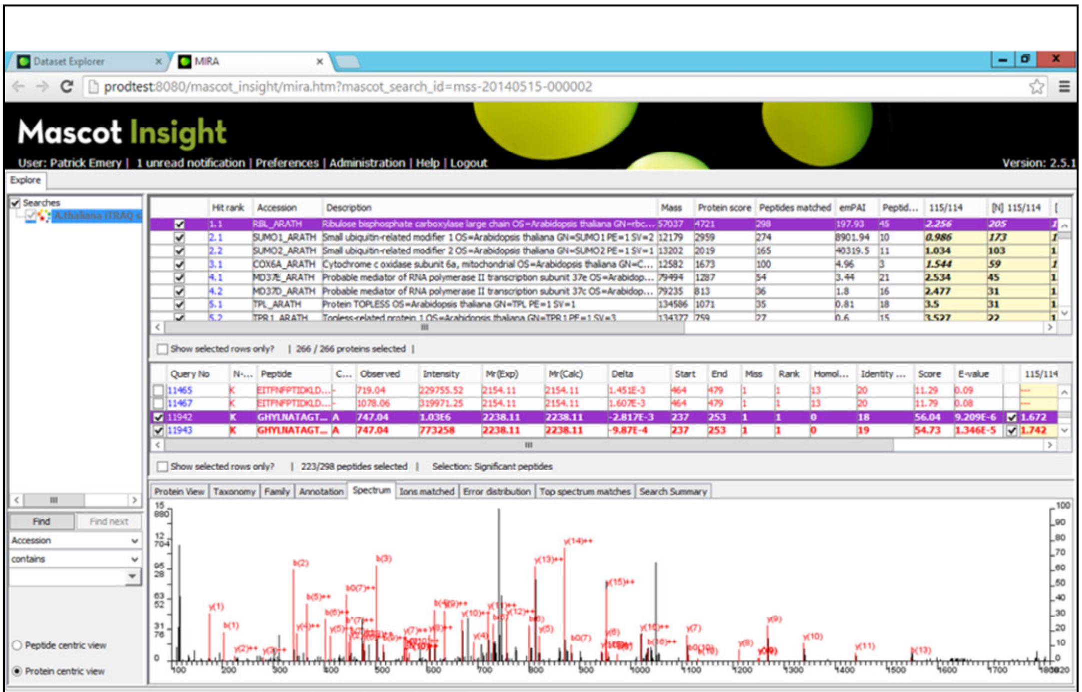
MIRA provides a tabular view of a single result or merged set of results. You can view results in either the Protein centric mode (as we are doing here), or in a peptide centric view. On the upper left we have the searches pane, which is used to change between the selected results, changing which result is being displayed in the main body of MIRA. It is also used for running search level reports from within MIRA and for setting result display parameters across multiple searches. The upper panel on the right contains the proteins table. This is the protein hit list for the currently selected result on the searches pane, and resembles the ‘Report builder’ view in the standard Mascot protein family report. If we click on a row in the Proteins table:



MASCOT : *Mascot Insight*

MATRIX SCIENCE

The peptides table, containing the list of matching peptide hits for the selected protein, and various details tabs (such as a protein view) are populated. Clicking on a table column header for either the proteins or peptides table sorts the table on that field. You can also re-order the columns by dragging them and also show and hide columns as you wish – you can save your default column selection, order and column widths if you wish; your choices are saved to the Mascot Insight database, so are available on any client PC you log in from. In the bottom left of the main pane is a panel with search functions, and options to switch between this view and a peptide (or spectrum) centric view.



	Hit rank	Accession	Description	Mass	Protein score	Peptides matched	emPAI	Peptid...	115/114	[N] 115/114	[
☑	1.1	RBL_ARATH	Ribulose bisphosphate carboxylase large chain OS=Arabidopsis thaliana GN=rbc...	57037	4721	298	197.93	45	*2.256*	*205*	1
☑	2.1	SUMO1_ARATH	Small ubiquitin-related modifier 1 OS=Arabidopsis thaliana GN=SUMO1 PE=1 SV=2	12179	2959	274	8901.94	10	*0.986*	*173*	1
☑	2.2	SUMO2_ARATH	Small ubiquitin-related modifier 2 OS=Arabidopsis thaliana GN=SUMO2 PE=1 SV=1	13202	2019	165	40319.5	11	**1.034**	**103**	1
☑	3.1	COX6A_ARATH	Cytochrome c oxidase subunit 6a, mitochondrial OS=Arabidopsis thaliana GN=C...	12582	1673	100	4.96	3	*1.544*	*59*	1
☑	4.1	MD37E_ARATH	Probable mediator of RNA polymerase II transcription subunit 37e OS=Arabidop...	79494	1287	54	3.44	21	**2.534**	**45**	1
☑	4.2	MD37D_ARATH	Probable mediator of RNA polymerase II transcription subunit 37c OS=Arabidop...	79235	813	36	1.8	16	**2.477**	**31**	1
☑	5.1	TPL_ARATH	Protein TOPLESS OS=Arabidopsis thaliana GN=TPL PE=1 SV=1	134586	1071	35	0.81	18	**3.5**	**31**	1
☑	5.2	TPR1_ARATH	Topless-related protein 1 OS=Arabidopsis thaliana GN=TPR1 PE=1 SV=3	134377	759	27	0.6	15	**3.527**	**22**	1

☐ Show selected rows only? | 266 / 266 proteins selected |

	Query No	N-...	Peptide	C...	Observed	Intensity	Mr(Exp)	Mr(Calc)	Delta	Start	End	Miss	Rank	Homol...	Identity ...	Score	E-value		115/114
☐	11465	K	EITFNFPTIDKLD...	-	719.04	229755.52	2154.11	2154.11	1.451E-3	464	479	1	1	13	20	11.29	0.09		---
☐	11467	K	EITFNFPTIDKLD...	-	1078.06	319971.25	2154.11	2154.11	1.607E-3	464	479	1	1	13	20	11.79	0.08		---
☑	11942	K	**GHYLNATAGT...**	**A**	**747.04**	**1.03E6**	**2238.11**	**2238.11**	**-2.817E-3**	**237**	**253**	**1**	**1**	**0**	**18**	**56.04**	**9.209E-6**	☑	**1.672**
☑	11943	K	**GHYLNATAGT...**	**A**	**747.04**	**773258**	**2238.11**	**2238.11**	**-9.87E-4**	**237**	**253**	**1**	**1**	**0**	**19**	**54.73**	**1.346E-5**	☑	**1.742**

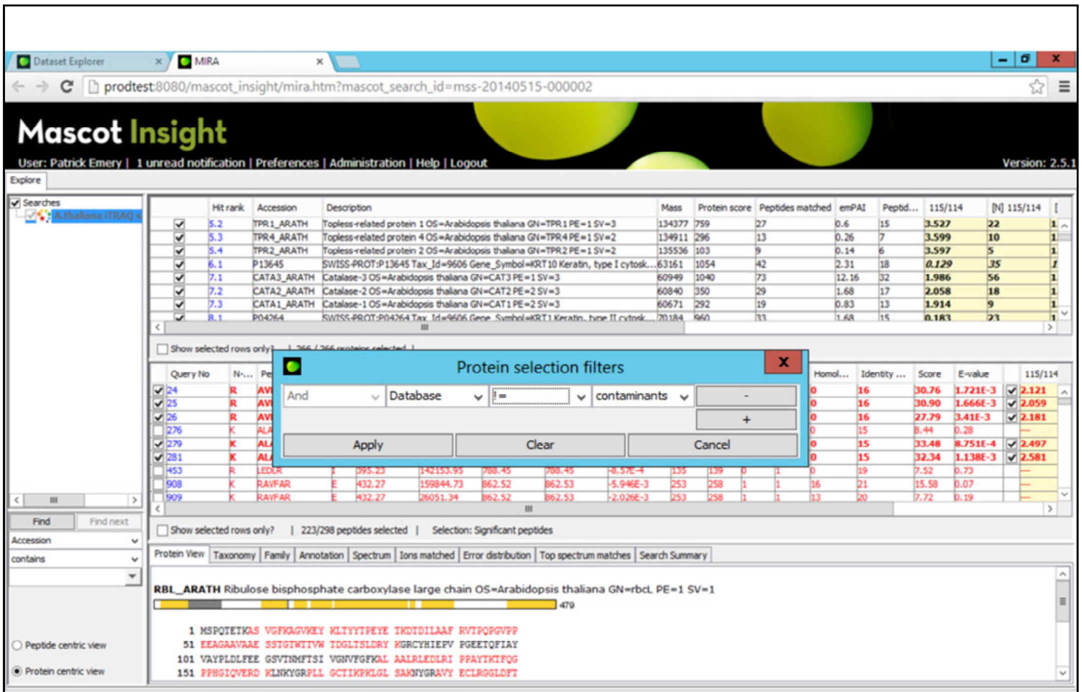
☐ Show selected rows only? | 223/298 peptides selected | Selection: Significant peptides

MASCOT : *Mascot Insight*

© 2014 Matrix Science

MATRIX SCIENCE

Clicking on a peptide row allows you to drill down and see peptide specific information, such as an annotated spectrum.



	Hit rank	Accession	Description	Mass	Protein score	Peptides matched	emPAI	Peptid...	115/114	[N] 115/114	[
☑	5.2	TPR1_ARATH	Topless-related protein 1 OS=Arabidopsis thaliana GN=TPR1 PE=1 SV=3	134377	759	27	0.6	15	3.527	22	1
☑	5.3	TPR4_ARATH	Topless-related protein 4 OS=Arabidopsis thaliana GN=TPR4 PE=1 SV=2	134911	296	13	0.26	7	3.599	10	1
☑	5.4	TPR2_ARATH	Topless-related protein 2 OS=Arabidopsis thaliana GN=TPR2 PE=1 SV=2	135536	103	9	0.14	6	3.597	5	1
☑	6.1	P13645	SWISS-PROT:P13645 Tax_Id=9606 Gene_Symbol=KRT10 Keratin, type I cytosk...	63161	1054	42	2.31	18	*0.129*	*35*	*1*
☑	7.1	CATA3_ARATH	Catalase-3 OS=Arabidopsis thaliana GN=CAT3 PE=1 SV=3	60949	1040	73	12.16	32	1.986	56	1
☑	7.2	CATA2_ARATH	Catalase-2 OS=Arabidopsis thaliana GN=CAT2 PE=2 SV=3	60840	350	29	1.68	17	2.058	18	1
☑	7.3	CATA1_ARATH	Catalase-1 OS=Arabidopsis thaliana GN=CAT1 PE=2 SV=3	60671	292	19	0.83	13	1.914	9	1
☑	8.1	P04264	SWISS-PROT:P04264 Tax_Id=9606 Gene_Symbol=KRT1 Keratin, type II cytosk...	70184	960	33	1.68	15	0.183	23	1

☐ Show selected rows only? | 266 / 266 proteins selected |

Protein selection filters

And | Database | != | contaminants

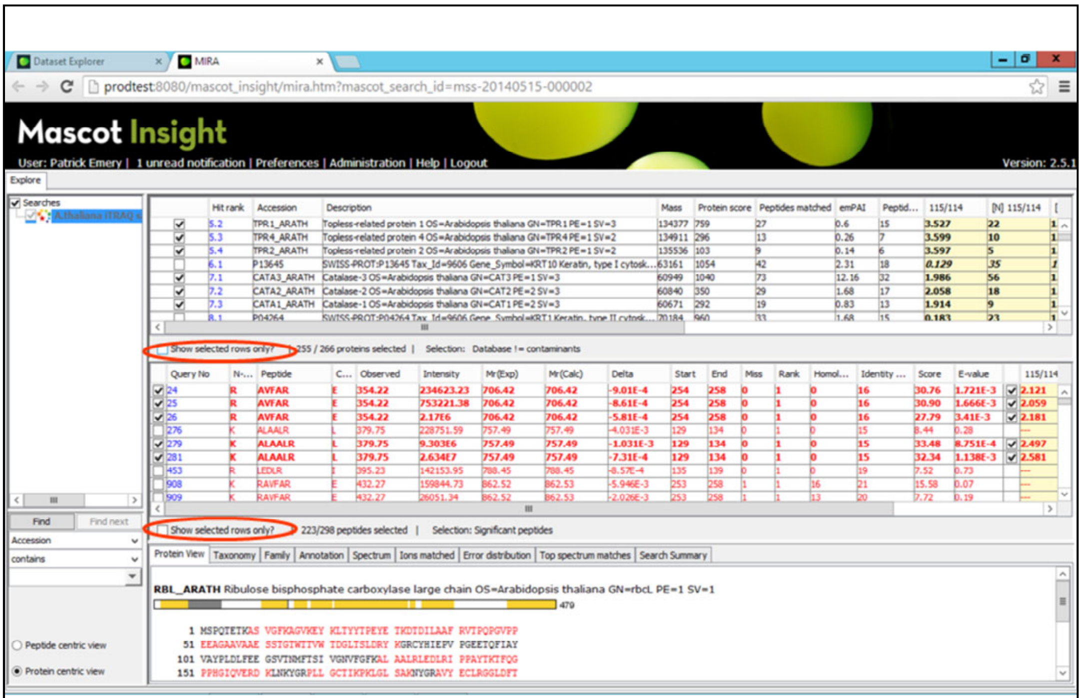
	Query No	N-...	Pe											Homol...	Identity ...	Score	E-value		115/114
☑	24	R	AV											0	16	30.76	1.721E-3	☑	2.121
☑	25	R	AV											0	16	30.90	1.666E-3	☑	2.059
☑	26	R	AV											0	16	27.79	3.41E-3	☑	2.181
☐	276	K	ALA											0	15	8.44	0.28		---
☑	279	K	ALA											0	15	33.48	8.751E-4	☑	2.497
☑	281	K	ALA											0	15	32.34	1.138E-3	☑	2.581
☐	453	R	LEDLR	I	395.23	142153.95	788.45	788.45	-8.57E-4	135	139	0	1	0	19	7.52	0.73		---
☐	908	K	RAVFAR	E	432.27	159844.73	862.52	862.53	-5.946E-3	253	258	1	1	16	21	15.58	0.07		---
☐	909	K	RAVFAR	E	432.27	26051.34	862.52	862.53	-2.026E-3	253	258	1	1	13	20	7.72	0.19		---

☐ Show selected rows only? | 223/298 peptides selected | Selection: Significant peptides

RBL_ARATH Ribulose bisphosphate carboxylase large chain OS=Arabidopsis thaliana GN=rbcL PE=1 SV=1

```
  1 MSPQTETKAS VGFKAGVKEY KLTYYTPEYE TKDTDILAAF RVTPQPGVPP
 51 EEAGAAVAAE SSTGTWTTVW TDGLTSLDRY KGRCYHIEPV PGEETQFIAY
101 VAYPLDLFEE GSVTNMFTSI VGNVFGFKAL AALRLEDLRI PPAYTKTFQG
151 PPHGIQVERD KLNKYGRPLL GCTIKPKLGL SAKNYGRAVY ECLRGGLDFT
```

Proteins and peptides are selected using the checkboxes in the first column of the tables. If you generate a protein or peptide level report, only the selected proteins or peptides are used. You can easily apply protein and peptide selection level filters in MIRA, which will automatically control which proteins or peptides are selected in the tables. Here, we are going to deselect out any matches to the contaminants database. Set up the filter and click ‘Apply’



	Hit rank	Accession	Description	Mass	Protein score	Peptides matched	emPAI	Peptid...	115/114	[N] 115/114	[
☑	5.2	TPR1_ARATH	Topless-related protein 1 OS=Arabidopsis thaliana GN=TPR1 PE=1 SV=3	134377	759	27	0.6	15	**3.527**	**22**	1
☑	5.3	TPR4_ARATH	Topless-related protein 4 OS=Arabidopsis thaliana GN=TPR4 PE=1 SV=2	134911	296	13	0.26	7	**3.599**	**10**	1
☑	5.4	TPR2_ARATH	Topless-related protein 2 OS=Arabidopsis thaliana GN=TPR2 PE=1 SV=2	135536	103	9	0.14	6	**3.597**	**5**	1
☐	6.1	P13645	SWISS-PROT:P13645 Tax_Id=9606 Gene_Symbol=KRT10 Keratin, type I cytosk...	63161	1054	42	2.31	18	***0.129***	***35***	*1*
☑	7.1	CATA3_ARATH	Catalase-3 OS=Arabidopsis thaliana GN=CAT3 PE=1 SV=3	60949	1040	73	12.16	32	**1.986**	**56**	1
☑	7.2	CATA2_ARATH	Catalase-2 OS=Arabidopsis thaliana GN=CAT2 PE=2 SV=3	60840	350	29	1.68	17	**2.058**	**18**	1
☑	7.3	CATA1_ARATH	Catalase-1 OS=Arabidopsis thaliana GN=CAT1 PE=2 SV=3	60671	292	19	0.83	13	**1.914**	**9**	1
☐	8.1	P04264	SWISS-PROT:P04264 Tax_Id=9606 Gene_Symbol=KRT1 Keratin, type II cytosk...	70184	960	33	1.68	15	**0.183**	**23**	1

☐ Show selected rows only? 255 / 266 proteins selected | Selection: Database != contaminants

	Query No	N-...	Peptide	C...	Observed	Intensity	Mr(Exp)	Mr(Calc)	Delta	Start	End	Miss	Rank	Homol...	Identity ...	Score	E-value		115/114
☑	24	**R**	**AVFAR**	**E**	**354.22**	**234623.23**	**706.42**	**706.42**	**-9.01E-4**	**254**	**258**	**0**	**1**	**0**	**16**	**30.76**	**1.721E-3**	☑	**2.121**
☑	25	**R**	**AVFAR**	**E**	**354.22**	**753221.38**	**706.42**	**706.42**	**-8.61E-4**	**254**	**258**	**0**	**1**	**0**	**16**	**30.90**	**1.666E-3**	☑	**2.059**
☑	26	**R**	**AVFAR**	**E**	**354.22**	**2.17E6**	**706.42**	**706.42**	**-5.81E-4**	**254**	**258**	**0**	**1**	**0**	**16**	**27.79**	**3.41E-3**	☑	**2.181**
☐	276	K	ALAALR	L	379.75	228751.59	757.49	757.49	-4.031E-3	129	134	0	1	0	15	8.44	0.28		---
☑	279	**K**	**ALAALR**	**L**	**379.75**	**9.303E6**	**757.49**	**757.49**	**-1.031E-3**	**129**	**134**	**0**	**1**	**0**	**15**	**33.48**	**8.751E-4**	☑	**2.497**
☑	281	**K**	**ALAALR**	**L**	**379.75**	**2.634E7**	**757.49**	**757.49**	**-7.31E-4**	**129**	**134**	**0**	**1**	**0**	**15**	**32.34**	**1.138E-3**	☑	**2.581**
☐	453	R	LEDLR	I	395.23	142153.95	788.45	788.45	-8.57E-4	135	139	0	1	0	19	7.52	0.73		---
☐	908	K	RAVFAR	E	432.27	159844.73	862.52	862.53	-5.946E-3	253	258	1	1	16	21	15.58	0.07		---
☐	909	K	RAVFAR	E	432.27	26051.34	862.52	862.53	-2.026E-3	253	258	1	1	13	20	7.72	0.19		---

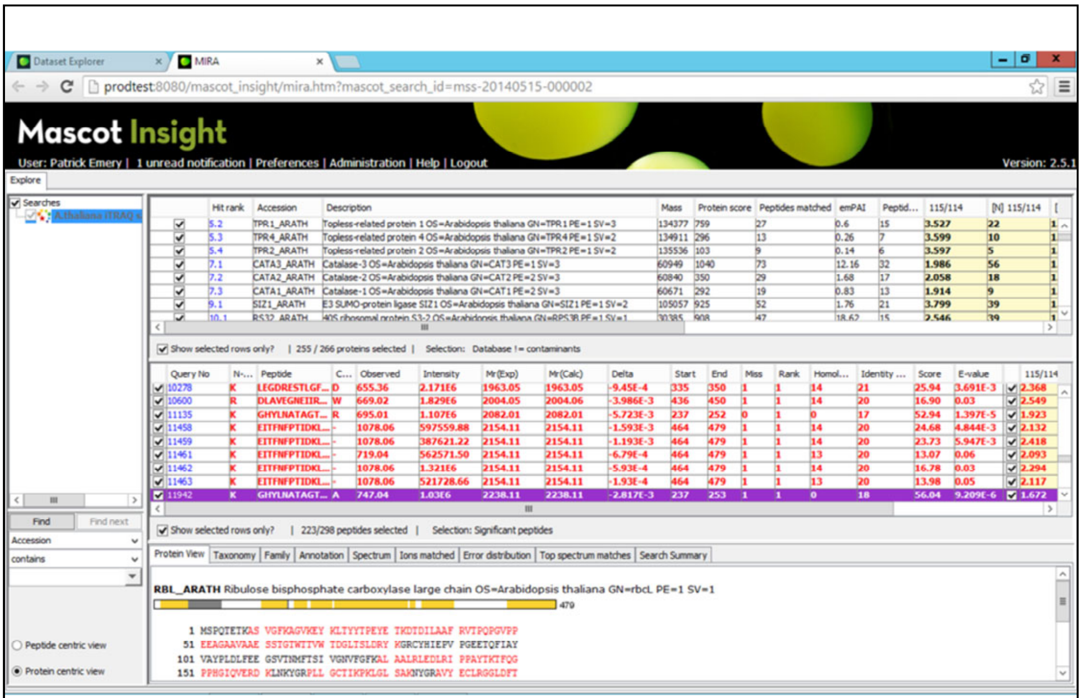
☐ Show selected rows only? 223/298 peptides selected | Selection: Significant peptides

RBL_ARATH Ribulose bisphosphate carboxylase large chain OS=Arabidopsis thaliana GN=rbcL PE=1 SV=1

479

```
  1 MSPQTETKAS VGFKAGVKEY KLTYYTPEYE TKDTDILAAF RVTPQPGVPP
 51 EEAGAAVAAE SSTGTWTTVW TDGLTSLDRY KGRCYHIEPV PGEETQFIAY
101 VAYPLDLFEE GSVTNMFTSI VGNVFGFKAL AALRLEDLRI PPAYTKTFQG
151 PPHGIQVERD KLNKYGRPLL GCTIKPKLGL SAKNYGRAVY ECLRGGLDFT
```

And we've now deselected the 11 matches to the contaminants database for this search result. I can hide unselected protein and peptide rows by checking the 'Show selected rows only?' checkboxes



	Hit rank	Accession	Description	Mass	Protein score	Peptides matched	emPAI	Peptid...	115/114	[N] 115/114	[
☑	5.2	TPR1_ARATH	Topless-related protein 1 OS=Arabidopsis thaliana GN=TPR1 PE=1 SV=3	134377	759	27	0.6	15	**3.527**	**22**	1
☑	5.3	TPR4_ARATH	Topless-related protein 4 OS=Arabidopsis thaliana GN=TPR4 PE=1 SV=2	134911	296	13	0.26	7	**3.599**	**10**	1
☑	5.4	TPR2_ARATH	Topless-related protein 2 OS=Arabidopsis thaliana GN=TPR2 PE=1 SV=2	135536	103	9	0.14	6	**3.597**	**5**	1
☑	7.1	CATA3_ARATH	Catalase-3 OS=Arabidopsis thaliana GN=CAT3 PE=1 SV=3	60949	1040	73	12.16	32	**1.986**	**56**	1
☑	7.2	CATA2_ARATH	Catalase-2 OS=Arabidopsis thaliana GN=CAT2 PE=2 SV=3	60840	350	29	1.68	17	**2.058**	**18**	1
☑	7.3	CATA1_ARATH	Catalase-1 OS=Arabidopsis thaliana GN=CAT1 PE=2 SV=3	60671	292	19	0.83	13	**1.914**	**9**	1
☑	9.1	SIZ1_ARATH	E3 SUMO-protein ligase SIZ1 OS=Arabidopsis thaliana GN=SIZ1 PE=1 SV=2	105057	925	52	1.76	21	**3.799**	**39**	1
☑	10.1	RS32_ARATH	40S ribosomal protein S3-2 OS=Arabidopsis thaliana GN=RPS3B PE=1 SV=1	30385	908	47	18.62	15	**2.546**	**39**	1

☑ Show selected rows only? | 255 / 266 proteins selected | Selection: Database != contaminants

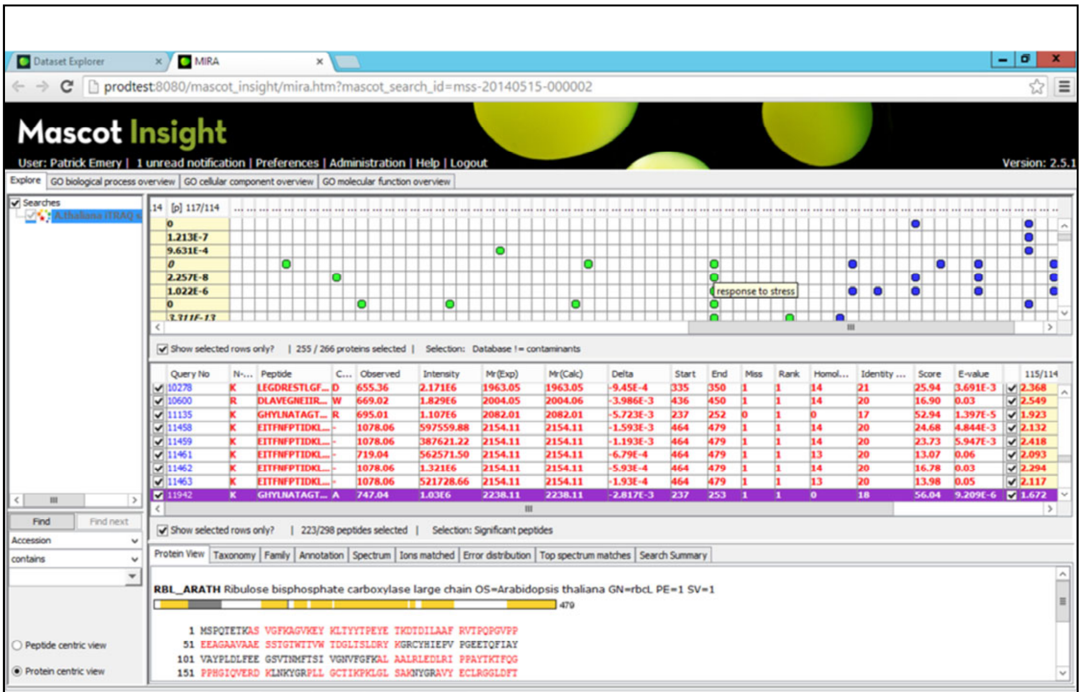
	Query No	N-...	Peptide	C...	Observed	Intensity	Mr(Exp)	Mr(Calc)	Delta	Start	End	Miss	Rank	Homol...	Identity ...	Score	E-value		115/114
☑	10278	K	LEGDRESTLGF...	D	655.36	2.171E6	1963.05	1963.05	-9.45E-4	335	350	1	1	14	21	25.94	3.691E-3	☑	2.368
☑	10600	R	DLAVEGNEIIR...	W	669.02	1.829E6	2004.05	2004.06	-3.986E-3	436	450	1	1	14	20	16.90	0.03	☑	2.549
☑	11135	K	GHYLNATAGT...	R	695.01	1.107E6	2082.01	2082.01	-5.723E-3	237	252	0	1	0	17	52.94	1.397E-5	☑	1.923
☑	11458	K	EITFNFPTIDKL...	-	1078.06	597559.88	2154.11	2154.11	-1.593E-3	464	479	1	1	14	20	24.68	4.844E-3	☑	2.132
☑	11459	K	EITFNFPTIDKL...	-	1078.06	387621.22	2154.11	2154.11	-1.193E-3	464	479	1	1	14	20	23.73	5.947E-3	☑	2.418
☑	11461	K	EITFNFPTIDKL...	-	719.04	562571.50	2154.11	2154.11	-6.79E-4	464	479	1	1	13	20	13.07	0.06	☑	2.093
☑	11462	K	EITFNFPTIDKL...	-	1078.06	1.321E6	2154.11	2154.11	-5.93E-4	464	479	1	1	14	20	16.78	0.03	☑	2.294
☑	11463	K	EITFNFPTIDKL...	-	1078.06	521728.66	2154.11	2154.11	-1.93E-4	464	479	1	1	13	20	13.98	0.05	☑	2.117
☑	11942	K	GHYLNATAGT...	A	747.04	1.03E6	2238.11	2238.11	-2.817E-3	237	253	1	1	0	18	56.04	9.209E-6	☑	1.672

☑ Show selected rows only? | 223/298 peptides selected | Selection: Significant peptides

RBL_ARATH Ribulose bisphosphate carboxylase large chain OS=Arabidopsis thaliana GN=rbcL PE=1 SV=1

```
  1 MSPQTETKAS VGFKAGVKEY KLTYYTPEYE TKDTDILAAF RVTPQPGVPP
 51 EEAGAAVAAE SSTGTWTTVW TDGLTSLDRY KGRCYHIEPV PGEETQFIAY
101 VAYPLDLFEE GSVTNMFTSI VGNVFGFKAL AALRLEDLRI PPAYTKTFQG
151 PPHGIQVERD KLNKYGRPLL GCTIKPKLGL SAKNYGRAVY ECLRGGLDFT
```

So now I'm only viewing proteins which matched SwissProt and significant (above homology threshold) peptide matches



☑ Show selected rows only? | 255 / 266 proteins selected | Selection: Database != contaminants

	Query No	N-...	Peptide	C...	Observed	Intensity	Mr(Exp)	Mr(Calc)	Delta	Start	End	Miss	Rank	Homol...	Identity ...	Score	E-value		115/114
☑	10278	K	LEGDRESTLGF...	D	655.36	2.171E6	1963.05	1963.05	-9.45E-4	335	350	1	1	14	21	25.94	3.691E-3	☑	2.368
☑	10600	R	DLAVEGNEIIR...	W	669.02	1.829E6	2004.05	2004.06	-3.986E-3	436	450	1	1	14	20	16.90	0.03	☑	2.549
☑	11135	K	GHYLNATAGT...	R	695.01	1.107E6	2082.01	2082.01	-5.723E-3	237	252	0	1	0	17	52.94	1.397E-5	☑	1.923
☑	11458	K	EITFNFPTIDKL...	-	1078.06	597559.88	2154.11	2154.11	-1.593E-3	464	479	1	1	14	20	24.68	4.844E-3	☑	2.132
☑	11459	K	EITFNFPTIDKL...	-	1078.06	387621.22	2154.11	2154.11	-1.193E-3	464	479	1	1	14	20	23.73	5.947E-3	☑	2.418
☑	11461	K	EITFNFPTIDKL...	-	719.04	562571.50	2154.11	2154.11	-6.79E-4	464	479	1	1	13	20	13.07	0.06	☑	2.093
☑	11462	K	EITFNFPTIDKL...	-	1078.06	1.321E6	2154.11	2154.11	-5.93E-4	464	479	1	1	14	20	16.78	0.03	☑	2.294
☑	11463	K	EITFNFPTIDKL...	-	1078.06	521728.66	2154.11	2154.11	-1.93E-4	464	479	1	1	13	20	13.98	0.05	☑	2.117
☑	11942	K	GHYLNATAGT...	A	747.04	1.03E6	2238.11	2238.11	-2.817E-3	237	253	1	1	0	18	56.04	9.209E-6	☑	1.672

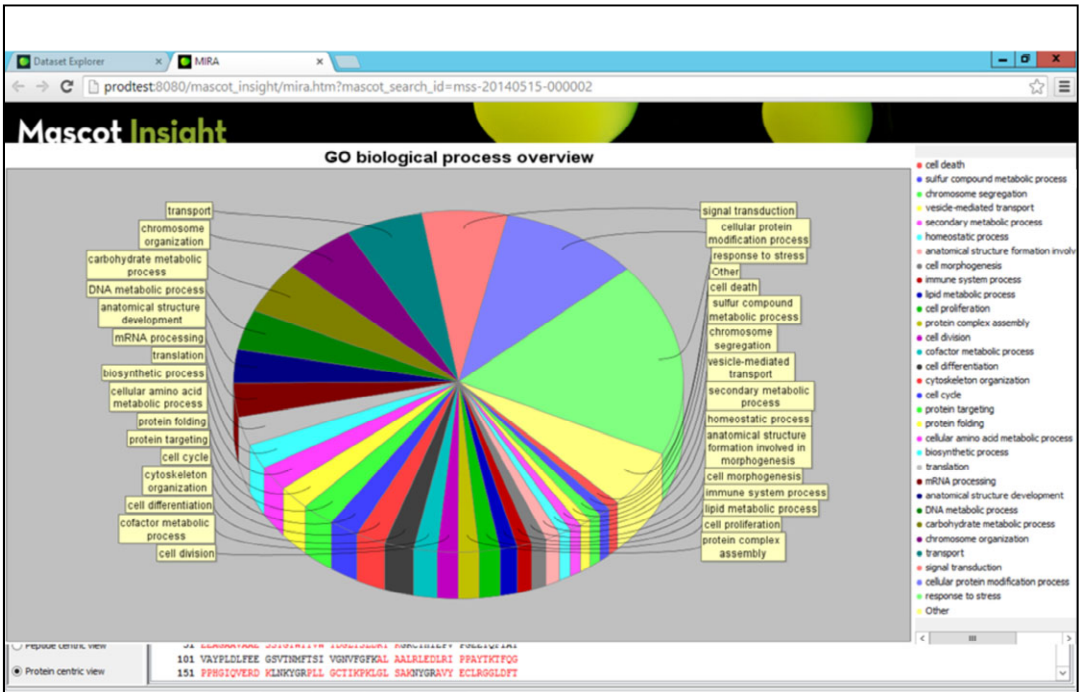
☑ Show selected rows only? | 223/298 peptides selected | Selection: Significant peptides

RBL_ARATH Ribulose bisphosphate carboxylase large chain OS=Arabidopsis thaliana GN=rbcL PE=1 SV=1

479

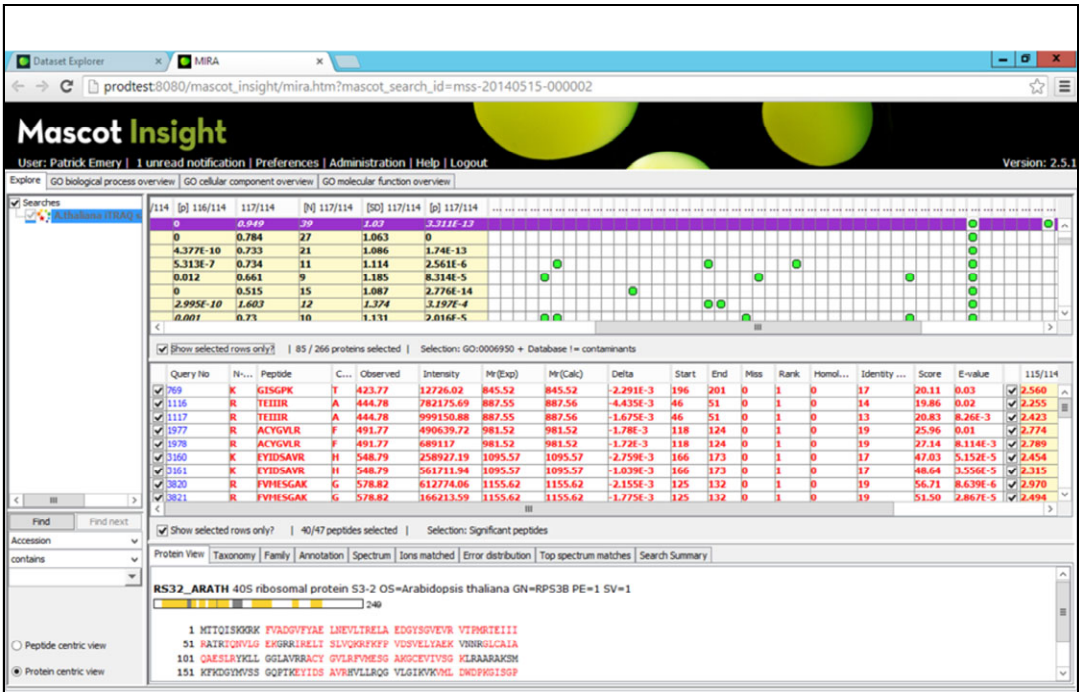
```
  1 MSPQTETKAS VGFKAGVKEY KLTYYTPEYE TKDTDILAAF RVTPQPGVPP
 51 EEAGAAVAAE SSTGTWTTVW TDGLTSLDRY KGRCYHIEPV PGEETQFIAY
101 VAYPLDLFEE GSVTNMFTSI VGNVFGFKAL AALRLEDLRI PPAYTKTFQG
151 PPHGIQVERD KLNKYGRPLL GCTIKPKLGL SAKNYGRAVY ECLRGGLDFT
```

You can enable GO Analysis in MIRA. GO annotations are shown in the proteins table, as different coloured dots. Hover over a cell for a tooltip.



```
101 VAYPLDLFEE GSVTNMFTSI VGNVFGFKAL AALRLEDLRI PPAYTKTFQG
151 PPHGIQVERD KLNKYGRPLL GCTIKPKLGL SAKNYGRAVY ECLRGGLDFT
```

GO assignments are also presented in additional tabs as piecharts. Click on a wedge of the pie-chart, and the protein table is filtered on that GO assignment



/114	[p] 116/114	117/114	[N] 117/114	[SD] 117/114	[p] 117/114
	0	0.949	39	1.03	3.311E-13
	0	0.784	27	1.063	0
	4.377E-10	0.733	21	1.086	1.74E-13
	5.313E-7	0.734	11	1.114	2.561E-6
	0.012	0.661	9	1.185	8.314E-5
	0	0.515	15	1.087	2.776E-14
	2.995E-10	1.603	12	1.374	3.197E-4
	0.001	0.73	10	1.131	2.016E-5

☑ Show selected rows only? | 85 / 266 proteins selected | Selection: GO:0006950 + Database != contaminants

	Query No	N-...	Peptide	C...	Observed	Intensity	Mr(Exp)	Mr(Calc)	Delta	Start	End	Miss	Rank	Homol...	Identity ...	Score	E-value		115/114
☑	769	K	GISGPK	T	423.77	12726.02	845.52	845.52	-2.291E-3	196	201	0	1	0	17	20.11	0.03	☑	2.560
☑	1116	R	TEIIIR	A	444.78	782175.69	887.55	887.56	-4.435E-3	46	51	0	1	0	14	19.86	0.02	☑	2.255
☑	1117	R	TEIIIR	A	444.78	999150.88	887.55	887.56	-1.675E-3	46	51	0	1	0	13	20.83	8.26E-3	☑	2.423
☑	1977	R	ACYGVLR	F	491.77	490639.72	981.52	981.52	-1.78E-3	118	124	0	1	0	19	25.96	0.01	☑	2.774
☑	1978	R	ACYGVLR	F	491.77	689117	981.52	981.52	-1.72E-3	118	124	0	1	0	19	27.14	8.114E-3	☑	2.789
☑	3160	K	EYIDSAVR	H	548.79	258927.19	1095.57	1095.57	-2.759E-3	166	173	0	1	0	17	47.03	5.152E-5	☑	2.454
☑	3161	K	EYIDSAVR	H	548.79	561711.94	1095.57	1095.57	-1.039E-3	166	173	0	1	0	17	48.64	3.556E-5	☑	2.315
☑	3820	R	FVMESGAK	G	578.82	612774.06	1155.62	1155.62	-2.155E-3	125	132	0	1	0	19	56.71	8.639E-6	☑	2.970
☑	3821	R	FVMESGAK	G	578.82	166213.59	1155.62	1155.62	-1.775E-3	125	132	0	1	0	19	51.50	2.867E-5	☑	2.494

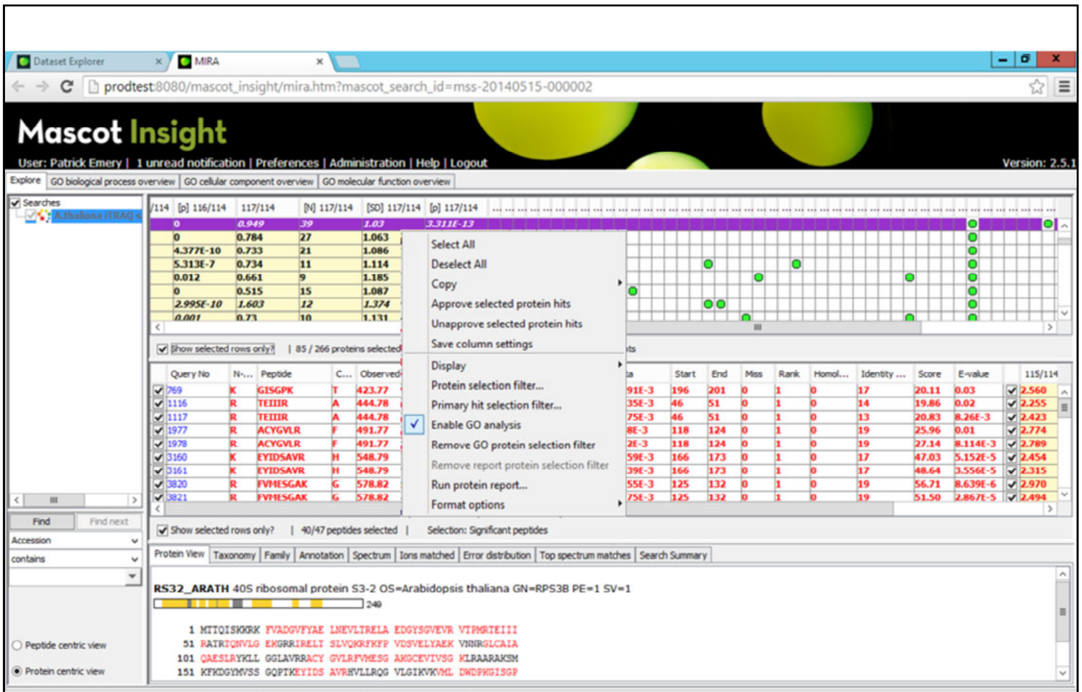
☑ Show selected rows only? | 40/47 peptides selected | Selection: Significant peptides

RS32_ARATH 40S ribosomal protein S3-2 OS=Arabidopsis thaliana GN=RPS3B PE=1 SV=1

249

```
  1 MTTQISKKRK FVADGVFYAE LNEVLTRELA EDGYSGVEVR VTPMRTEIII
 51 RAIRIQNVLG EKGRRIRELT SLVQKRFKFP VDSVELYAEK VNNRGLCAIA
101 QAESLRYKLL GGLAVRRACY GVLRFVMESG AKGCEVIVSG KLRAARAKSM
151 KFKDGYMVSS GQPTKEYIDS AVRHVLLRQG VLGIKVKVML DWDPKGISGP
```

Now we are just looking at the 85 protein matches with the GO biological function assignment of ‘Response to stress’ and which are not matches from the contaminants database



/114	[p] 116/114	117/114	[N] 117/114	[SD] 117/114	[p] 117/114
	0	0.949	39	1.03	3.311E-13
	0	0.784	27	1.063	
	4.377E-10	0.733	21	1.086	
	5.313E-7	0.734	11	1.114	
	0.012	0.661	9	1.185	
	0	0.515	15	1.087	
	2.995E-10	1.603	12	1.374	
	0.001	0.73	10	1.131	

☑ Show selected rows only? | 85 / 266 proteins selected

	Query No	N-...	Peptide	C...	Observed
☑	769	K	GISGPK	T	423.77
☑	1116	R	TEIIIR	A	444.78
☑	1117	R	TEIIIR	A	444.78
☑	1977	R	ACYGVLR	F	491.77
☑	1978	R	ACYGVLR	F	491.77
☑	3160	K	EYIDSAVR	H	548.79
☑	3161	K	EYIDSAVR	H	548.79
☑	3820	R	FVMESGAK	G	578.82
☑	3821	R	FVMESGAK	G	578.82

...ta	Start	End	Miss	Rank	Homol...	Identity ...	Score	E-value		115/114
91E-3	196	201	0	1	0	17	20.11	0.03	☑	2.560
35E-3	46	51	0	1	0	14	19.86	0.02	☑	2.255
75E-3	46	51	0	1	0	13	20.83	8.26E-3	☑	2.423
8E-3	118	124	0	1	0	19	25.96	0.01	☑	2.774
2E-3	118	124	0	1	0	19	27.14	8.114E-3	☑	2.789
59E-3	166	173	0	1	0	17	47.03	5.152E-5	☑	2.454
39E-3	166	173	0	1	0	17	48.64	3.556E-5	☑	2.315
55E-3	125	132	0	1	0	19	56.71	8.639E-6	☑	2.970
75E-3	125	132	0	1	0	19	51.50	2.867E-5	☑	2.494

☑ Show selected rows only? | 40/47 peptides selected | Selection: Significant peptides

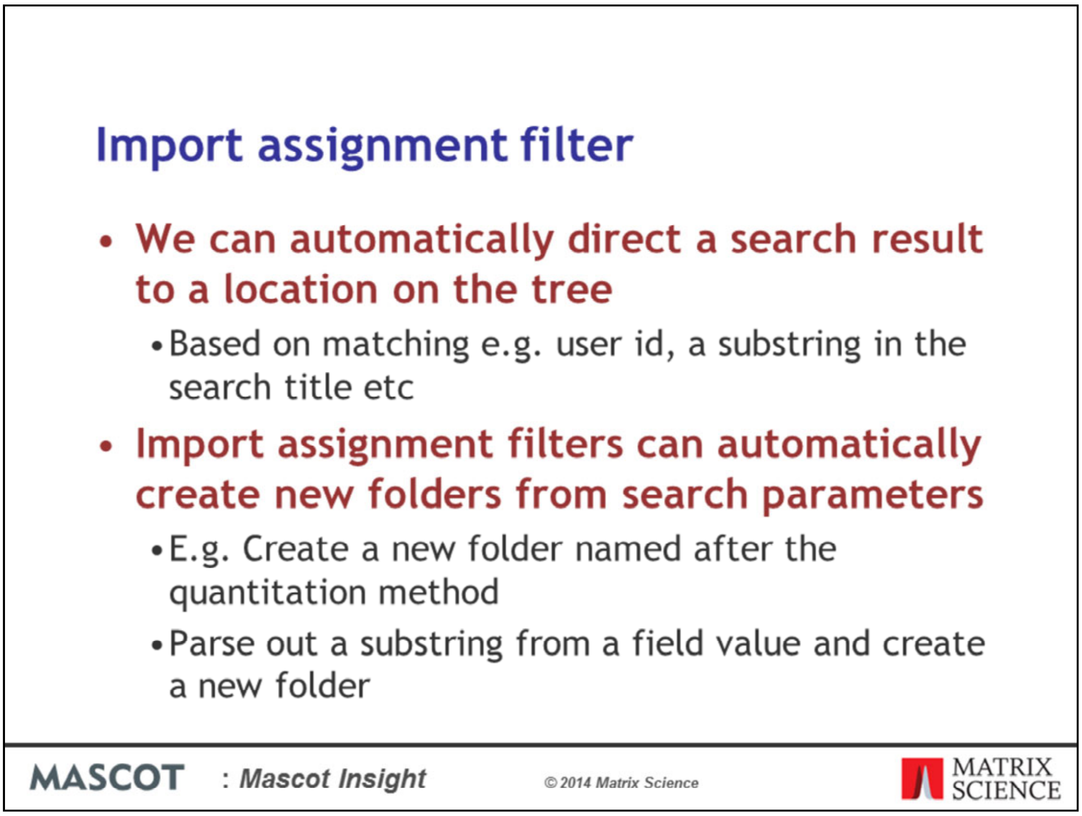
RS32_ARATH 40S ribosomal protein S3-2 OS=Arabidopsis thaliana GN=RPS3B PE=1 SV=1

249

```
  1 MTTQISKKRK FVADGVFYAE LNEVLTRELA EDGYSGVEVR VTPMRTEIII
 51 RAIRIQNVLG EKGRRIRELI SLVQKRFKFP VDSVELYAEK VNNRGLCAIA
101 QAESLRYKLL GGLAVRRACY GVLRFVMESG AKGCEVIVSG KLRAARAKSM
151 KFKDGYMVSS GQPTKEYIDS AVRHVLLRQG VLGIKVKVML DWDPKGISGP
```

MASCOT : *Mascot Insight* © 2013 Matrix Science MATRIX SCIENCE

Once you’re happy with a particular set of protein matches, you can approve the protein hits and make manual annotations if you wish. Approving a protein hit flags the protein in the Mascot Insight database. You can use this to easily highlight and filter out just those protein hits you think are present in the search sample, or are of particular interest.



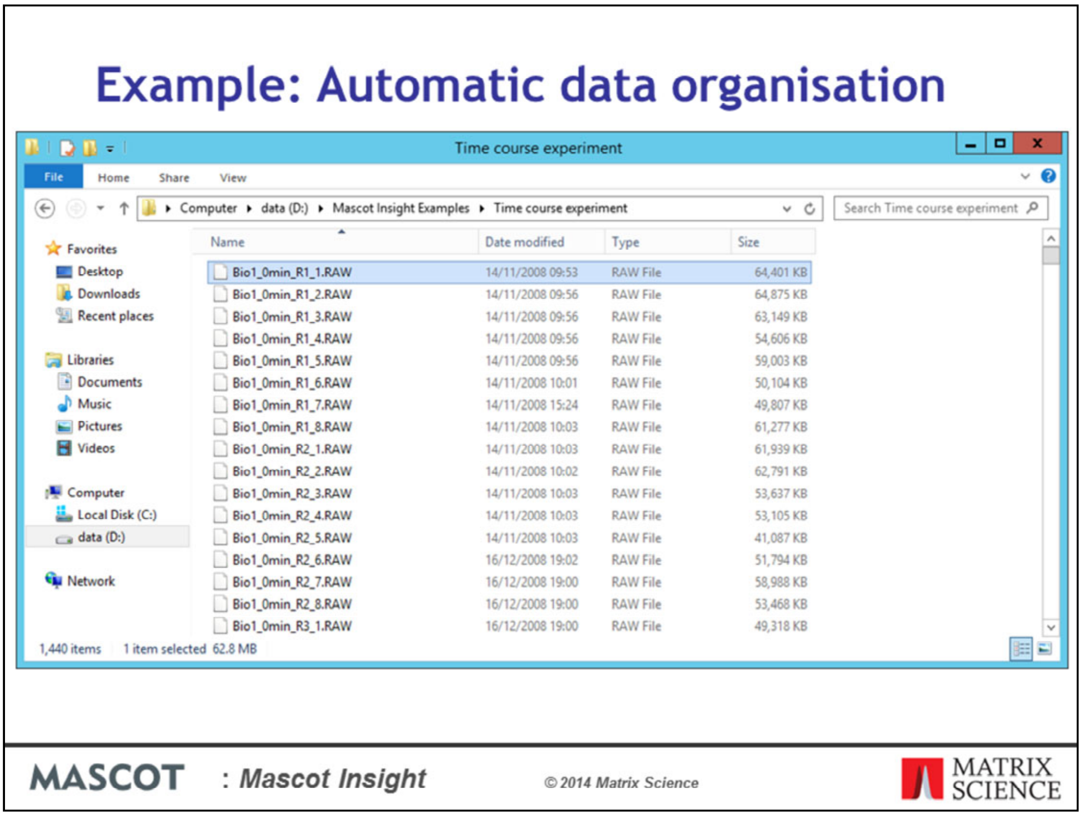
Import assignment filter

- We can automatically direct a search result to a location on the tree
  - Based on matching e.g. user id, a substring in the search title etc
- Import assignment filters can automatically create new folders from search parameters
  - E.g. Create a new folder named after the quantitation method
  - Parse out a substring from a field value and create a new folder

MASCOT : *Mascot Insight*

MATRIX SCIENCE

For that previous example, I simply allowed the search result to be imported into the default location on the dataset explorer. However, you can override the default result import location using an Import assignment filter. This is a filter that you can define in the system which will direct a matching search, to a specified node on the system explorer tree, overriding the default import location. In addition to that, Import assignment filters can automatically create new folders on the tree based on the values of search parameters, or by parsing out a string from a field.



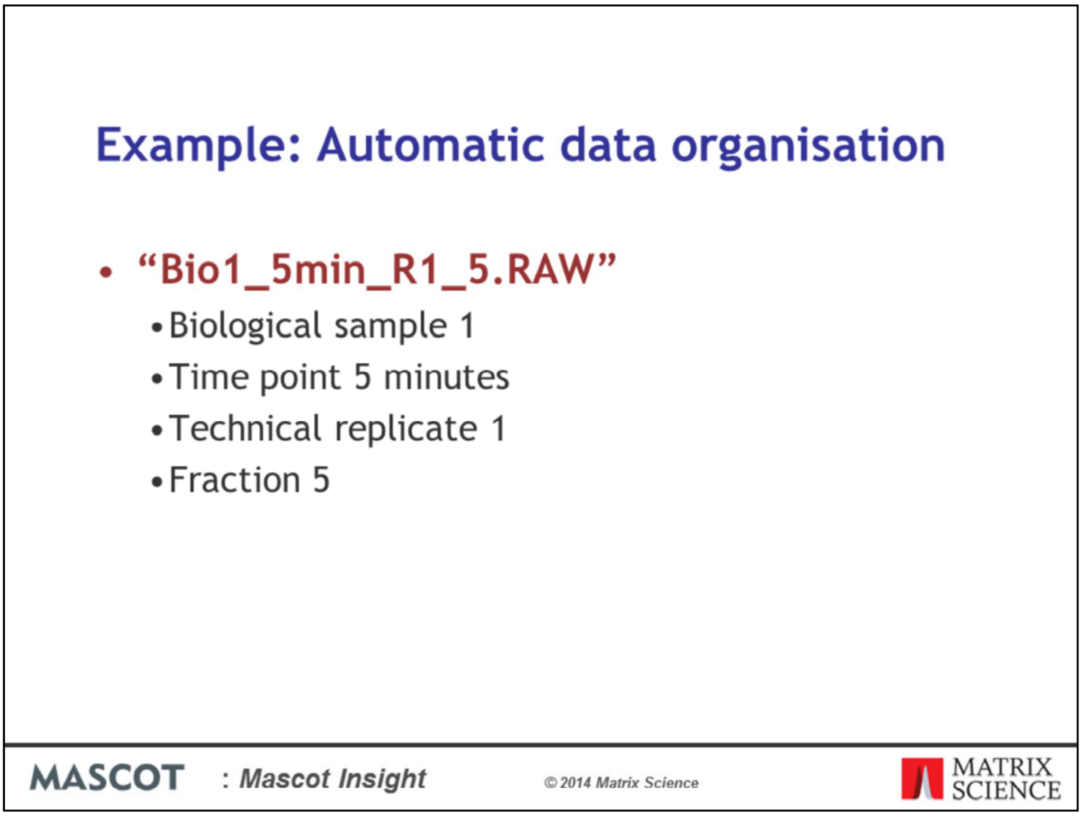
Example: Automatic data organisation

Name	Date modified	Type	Size
Bio1_0min_R1_1.RAW	14/11/2008 09:53	RAW File	64,401 KB
Bio1_0min_R1_2.RAW	14/11/2008 09:56	RAW File	64,875 KB
Bio1_0min_R1_3.RAW	14/11/2008 09:56	RAW File	63,149 KB
Bio1_0min_R1_4.RAW	14/11/2008 09:56	RAW File	54,606 KB
Bio1_0min_R1_5.RAW	14/11/2008 09:56	RAW File	59,003 KB
Bio1_0min_R1_6.RAW	14/11/2008 10:01	RAW File	50,104 KB
Bio1_0min_R1_7.RAW	14/11/2008 15:24	RAW File	49,807 KB
Bio1_0min_R1_8.RAW	14/11/2008 10:03	RAW File	61,277 KB
Bio1_0min_R2_1.RAW	14/11/2008 10:03	RAW File	61,939 KB
Bio1_0min_R2_2.RAW	14/11/2008 10:02	RAW File	62,791 KB
Bio1_0min_R2_3.RAW	14/11/2008 10:03	RAW File	53,637 KB
Bio1_0min_R2_4.RAW	14/11/2008 10:03	RAW File	53,105 KB
Bio1_0min_R2_5.RAW	14/11/2008 10:03	RAW File	41,087 KB
Bio1_0min_R2_6.RAW	16/12/2008 19:02	RAW File	51,794 KB
Bio1_0min_R2_7.RAW	16/12/2008 19:00	RAW File	58,988 KB
Bio1_0min_R2_8.RAW	16/12/2008 19:00	RAW File	53,468 KB
Bio1_0min_R3_1.RAW	16/12/2008 19:00	RAW File	49,318 KB

1,440 items 1 item selected 62.8 MB

Lets take a look at a real world example of how you might set up and use an Import assignment filter. The Data Explorer displays a tree structure of folders, results, and saved reports. You can make the structure as simple or as complicated as you wish, but the idea is to organise results logically so as to facilitate merging and reporting. For example, consider a quantitation experiment where 5 biological samples are analysed at 12 time intervals and each analysis is performed in triplicate (to give technical replicates). Each analysis involves 2D chromatography, requiring the mass spectrometry analysis of 8 fractions. This one experiment gives rise to 1440 Mascot search results. By grouping these results into a logical hierarchy of folders, we can minimise the work required to merge results and create meaningful reports.

In this case, the RAW data files are all in a single folder as shown, but the data file names contain all the information that we need to automatically generate a meaningful structure on the Data Explorer as the search results are imported.



Example: Automatic data organisation

- “Bio1_5min_R1_5.RAW”
  - Biological sample 1
  - Time point 5 minutes
  - Technical replicate 1
  - Fraction 5

MASCOT : *Mascot Insight*

MATRIX SCIENCE

So, here we have an example filename from the dataset. The name structure of the raw data file has been logically put together and breaks down like this:

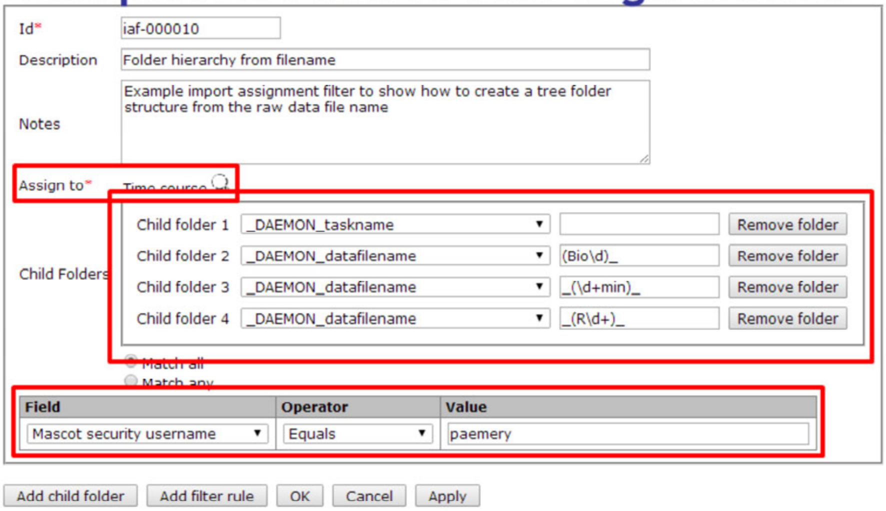
Bio – Biological sample number

Time interval

R – Technical replicate number

Fraction number

Example: Automatic data organisation



Id* iaf-000010

Description: Folder hierarchy from filename

Notes: Example import assignment filter to show how to create a tree folder structure from the raw data file name

Assign to*: Time course

Child Folders

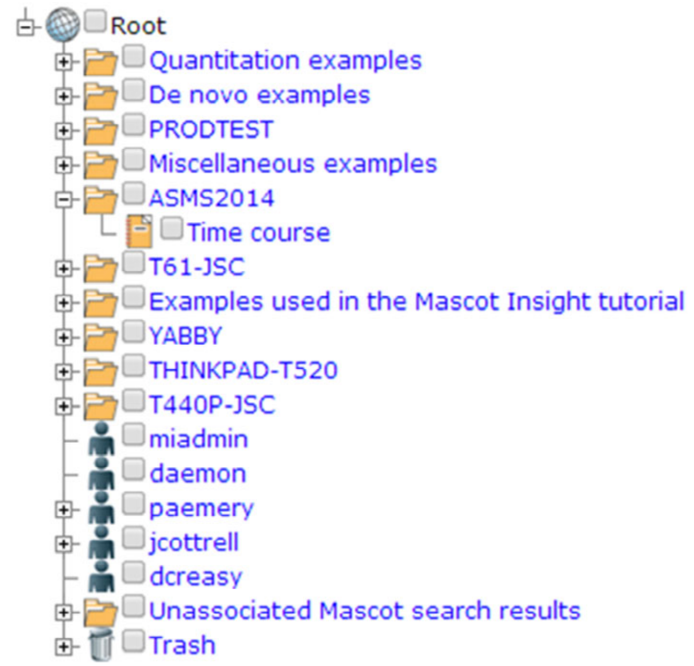
- Child folder 1: _DAEMON_taskname
- Child folder 2: _DAEMON_datafilename (Bio\d)_
- Child folder 3: _DAEMON_datafilename _(\d+min)_
- Child folder 4: _DAEMON_datafilename _(R\d+)_

● Match all
○ Match any

Field	Operator	Value
Mascot security username	Equals	paemery

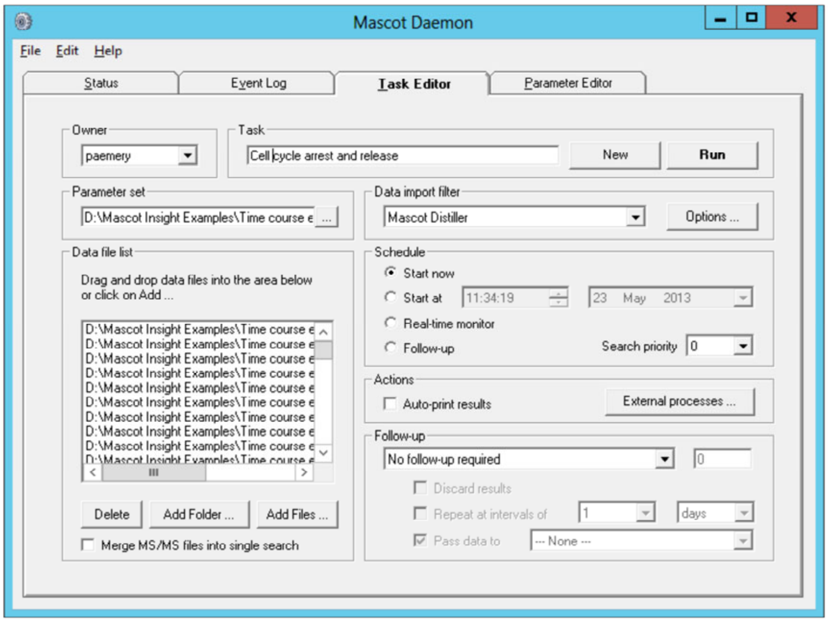
Now we set up the import assignment filter. When we set up the search tasks in Mascot Daemon, we'll run the searches as my Mascot security user, and we'll match the results on that, so this import filter will only run on my Mascot search results. We have to select the root node on the tree to associate the results with, in this case I've set up an project node for the data called 'Time course'. Now we want the system to automatically create folders and group the fractions from the same technical replicate for each time point together automatically. We can pull the required pieces of information out of the raw data filename using regular expressions, as shown above. So, under the Time course project node, we'll generate a folder named after the task name we enter into Mascot Daemon, then under that a folder for the biological sample, then the time interval and finally the technical replicate.

Although in this example, the required data for the import filter is contained in the raw data filename, it could be extracted from any other available field. For example, if the raw files were organised in a meaningful directory structure on your hard drive, the same structure could be re-created in the Explorer tree, automatically.



- ☐ Root
  - ☐ Quantitation examples
  - ☐ De novo examples
  - ☐ PRODTEST
  - ☐ Miscellaneous examples
  - ☐ ASMS2014
    - ☐ Time course
  - ☐ T61-JSC
  - ☐ Examples used in the Mascot Insight tutorial
  - ☐ YABBY
  - ☐ THINKPAD-T520
  - ☐ T440P-JSC
  - ☐ miadmin
  - ☐ daemon
  - ☐ paemery
  - ☐ jcottrell
  - ☐ dcreasy
  - ☐ Unassociated Mascot search results
  - ☐ Trash

Before we run the searches, this is what the explorer tree in Mascot Insight looks like



Owner: paemery

Task: Cell cycle arrest and release

Parameter set: D:\Mascot Insight Examples\Time course e

Data import filter: Mascot Distiller

Data file list

Drag and drop data files into the area below
or click on Add ...

D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e
D:\Mascot Insight Examples\Time course e

☐ Merge MS/MS files into single search

Schedule
● Start now
○ Start at 11:34:19 23 May 2013
○ Real-time monitor
○ Follow-up
Search priority 0

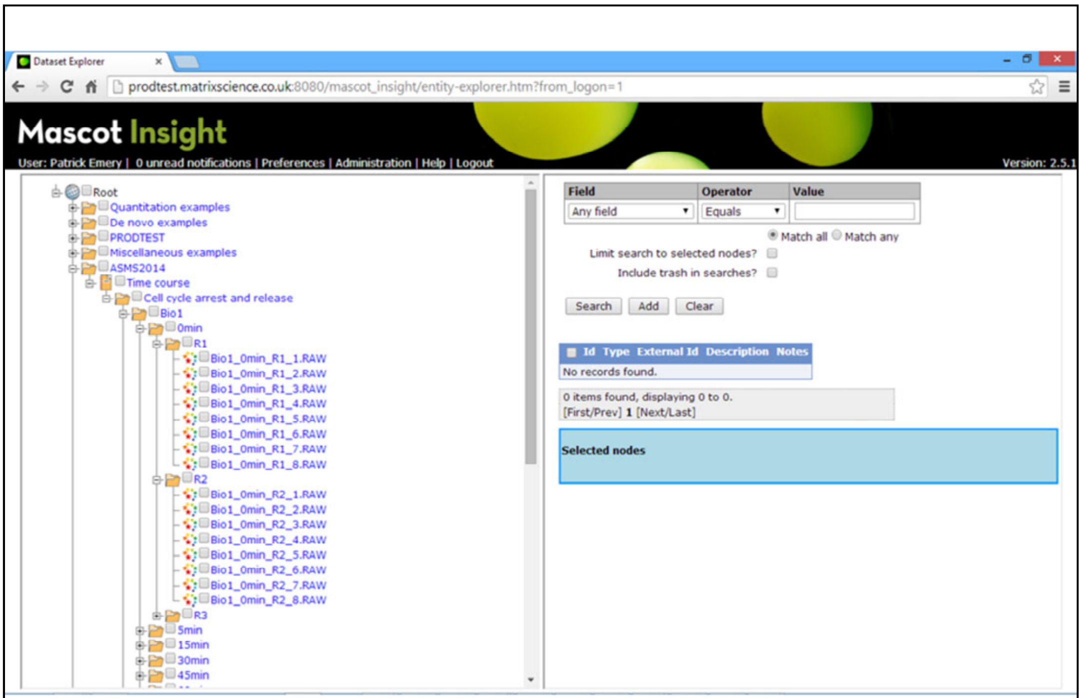
Actions
☐ Auto-print results

Follow-up
No follow-up required 0
☐ Discard results
☐ Repeat at intervals of 1 days
☑ Pass data to --- None ---

MASCOT : *Mascot Insight*

MATRIX SCIENCE

We set up the task in Mascot Daemon as normal, making sure we choose Mascot Distiller as the import filter and, since this is an N15 quantitation dataset, we'll set it to automatically carry out quantitation using Mascot Distiller as well. When a search is completed, the result will automatically be imported into Mascot Insight, and the Import assignment filter we set up will direct it to the desired location on the explorer tree. When Daemon completes quantitation for a dataset, the quantitation results are also automatically uploaded and imported.



- Root
  - Quantitation examples
  - De novo examples
  - PRODTEST
  - Miscellaneous examples
  - ASMS2014
    - Time course
      - Cell cycle arrest and release
        - Bio1
          - 0min
            - R1
              - Bio1_0min_R1_1.RAW
              - Bio1_0min_R1_2.RAW
              - Bio1_0min_R1_3.RAW
              - Bio1_0min_R1_4.RAW
              - Bio1_0min_R1_5.RAW
              - Bio1_0min_R1_6.RAW
              - Bio1_0min_R1_7.RAW
              - Bio1_0min_R1_8.RAW
            - R2
              - Bio1_0min_R2_1.RAW
              - Bio1_0min_R2_2.RAW
              - Bio1_0min_R2_3.RAW
              - Bio1_0min_R2_4.RAW
              - Bio1_0min_R2_5.RAW
              - Bio1_0min_R2_6.RAW
              - Bio1_0min_R2_7.RAW
              - Bio1_0min_R2_8.RAW
            - R3
          - 5min
          - 15min
          - 30min
          - 45min

Field	Operator	Value
Any field	Equals	

● Match all ○ Match any

Limit search to selected nodes? ☐

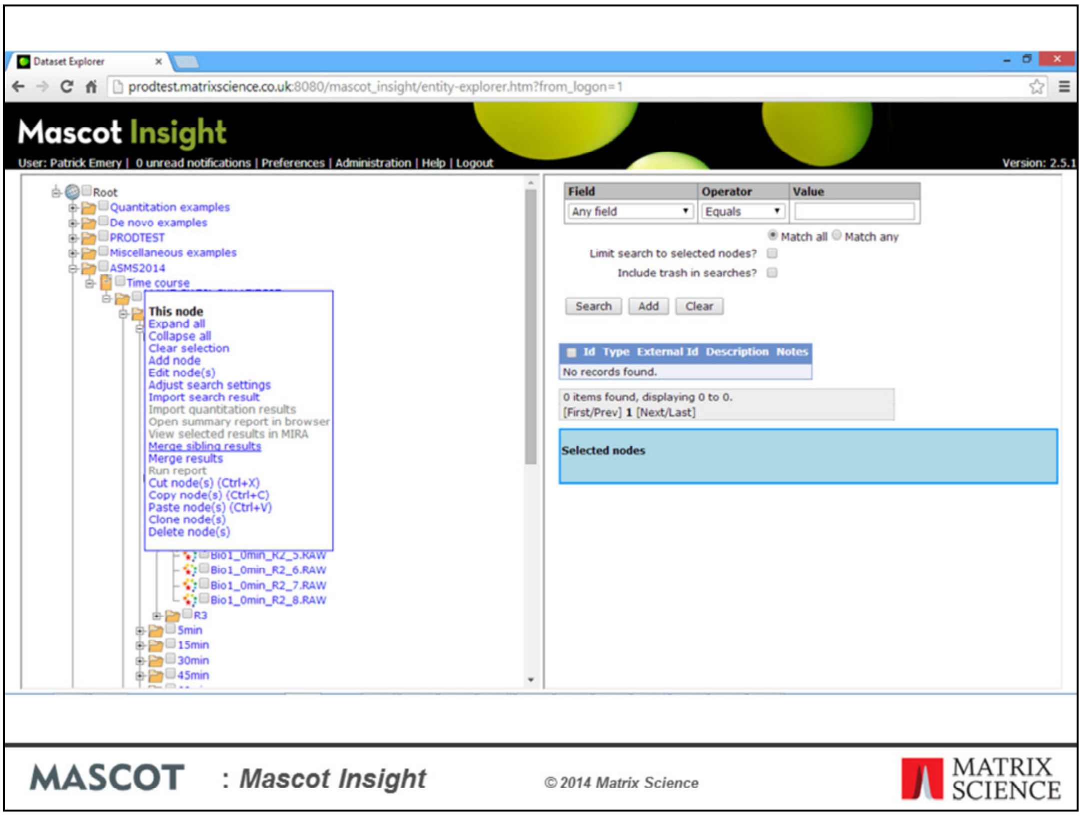
Include trash in searches? ☐

☐	Id	Type	External Id	Description	Notes
No records found.					

0 items found, displaying 0 to 0.
[First/Prev] 1 [Next/Last]

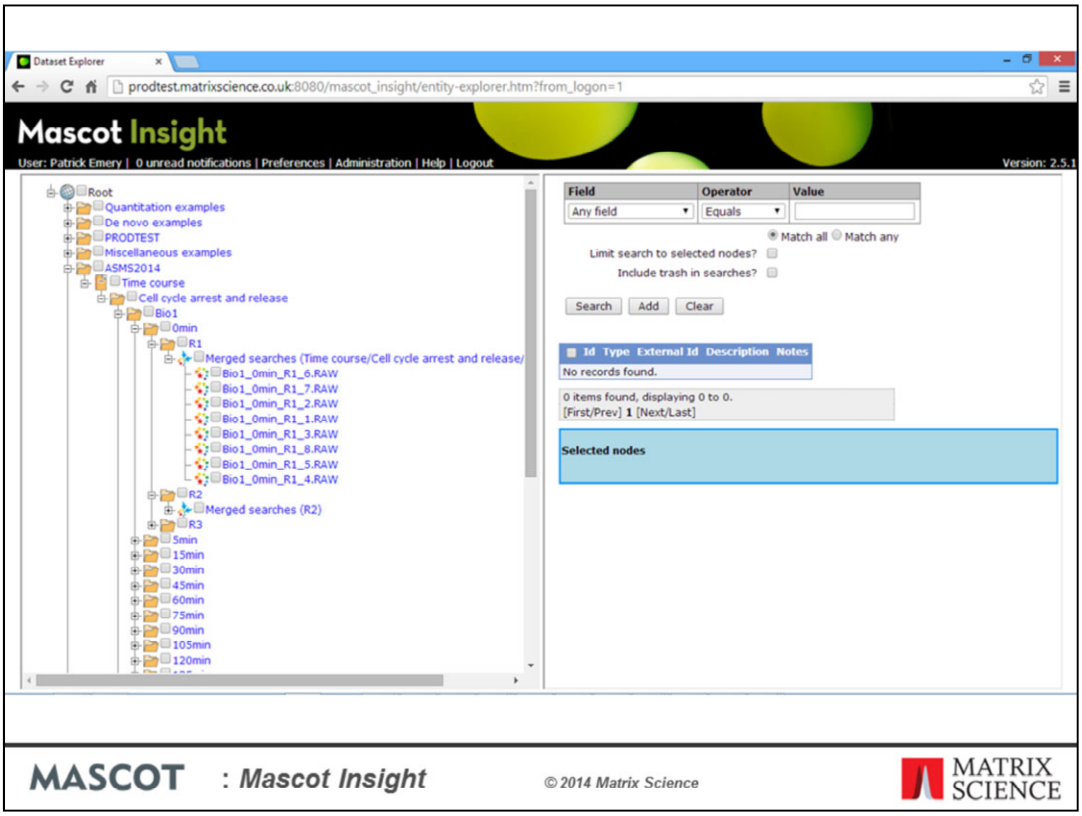
Selected nodes

So that when the searches have completed, the explorer tree now looks like this - with a total of 181 folders automatically created under the 'Time course' project node, and under each of those folders the eight fractions for each technical replicate at each time interval, all without any manual intervention after the import assignment filter had been set up.



Now we need to merge each of those groups of eight fractions into merged search results. Because the results are grouped together on the tree in the way we want to merge them, to do this we just need to right click over the 'Time course' project node (the common parent node) and select the 'Merge sibling results' option from the context menu.

The system will then run through the dataset, merging together search results at the same level of the tree so, all of the fractions for each technical replicate will be merged together. More advanced, manual merge options are available if you need to do some additional processing when carrying out the merge – for example, if you have rotated the samples around iTRAQ report ions channels between replicates



Field: Any field ▾
Operator: Equals ▾

● Match all ○ Match any

Limit search to selected nodes? ☐

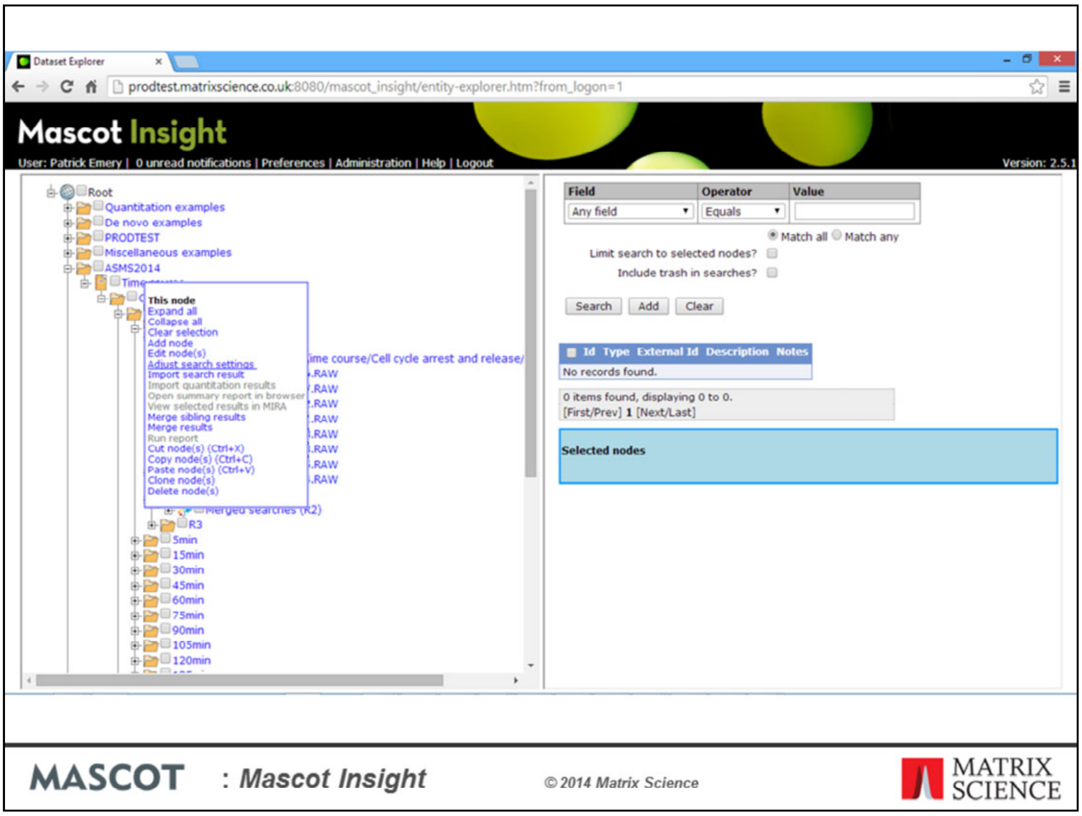
Include trash in searches? ☐

No records found.

0 items found, displaying 0 to 0.

[First/Prev] 1 [Next/Last]

When the merge is completed, the organisation of the search results on the tree will have changed, so that each of the eight fraction search results for each technical replicate of each time interval is now a child node of a new ‘Merged searches’ result node. Now that we have all of the merged search nodes generated, we can use them just like any ‘normal’ search result node and view them in MIRA, or generate reports from them.



prodtest.matrixscience.co.uk:8080/mascot_insight/entity-explorer.htm?from_logon=1

Any field ▾ | Equals ▾ |

● Match all ○ Match any

Limit search to selected nodes? ☐

Include trash in searches? ☐

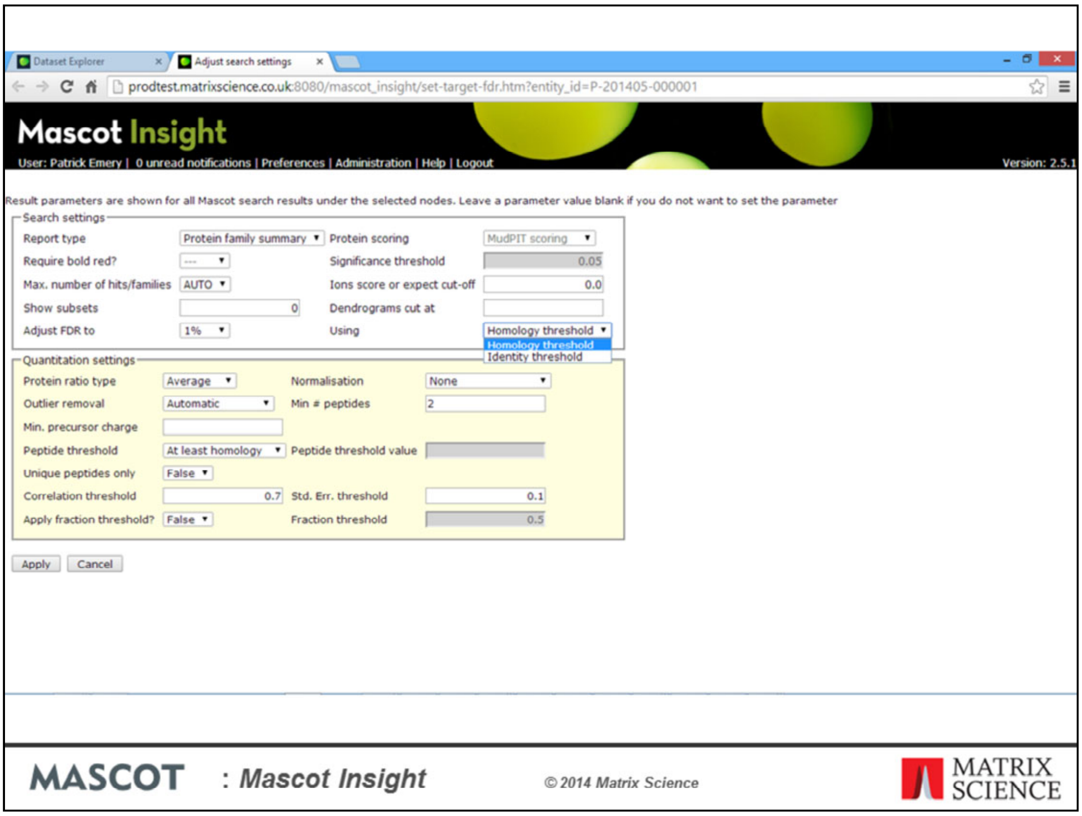
No records found.

0 items found, displaying 0 to 0.
[First/Prev] **1** [Next/Last]

MASCOT : *Mascot Insight*

MATRIX SCIENCE

Now that we’ve merged the results, we want to set the peptide false discovery rate to 1% for all of the searches. Again, we can do this for all the searches by right clicking over the ‘Time course’ project node and selecting ‘Adjust search settings’



prodtest.matrixscience.co.uk:8080/mascot_insight/set-target-fdr.htm?entity_id=P-201405-000001

Result parameters are shown for all Mascot search results under the selected nodes. Leave a parameter value blank if you do not want to set the parameter

Search settings

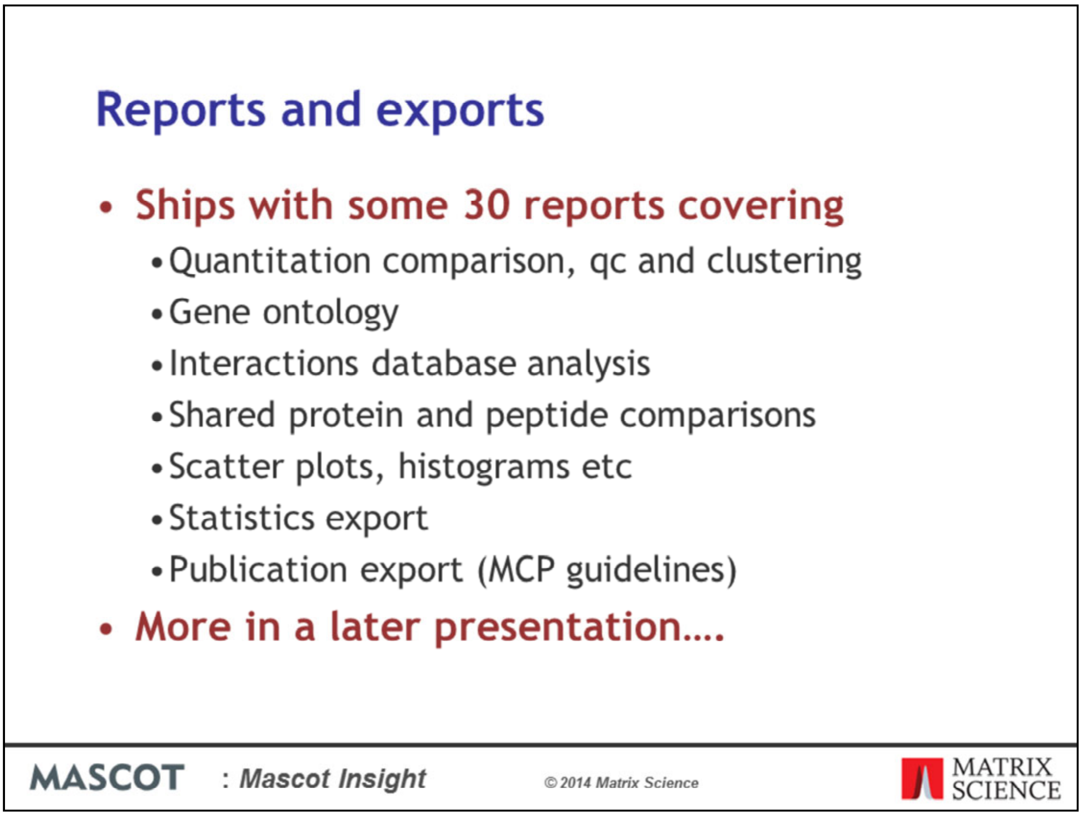
Report type	Protein family summary	Protein scoring	MudPIT scoring
Require bold red?	---	Significance threshold	0.05
Max. number of hits/families	AUTO	Ions score or expect cut-off	0.0
Show subsets	0	Dendrograms cut at	
Adjust FDR to	1%	Using	Homology threshold Homology threshold Identity threshold

Quantitation settings

Protein ratio type	Average	Normalisation	None
Outlier removal	Automatic	Min # peptides	2
Min. precursor charge			
Peptide threshold	At least homology	Peptide threshold value	
Unique peptides only	False		
Correlation threshold	0.7	Std. Err. threshold	0.1
Apply fraction threshold?	False	Fraction threshold	0.5

On the form that opens we can set the target FDR and any other search and quantitation settings that we wish. I'm sure you'll appreciate how powerful it is to be able to change settings like these across tens or hundreds of files from a single form - the changes you make here are saved in the database, so you don't need to keep adjusting them every time you want view the dataset.

That was a fairly complex example of how you can use an import assignment filter, and often the desired folder structure would be much simpler. However, it does illustrate that you can easily and automatically create a complex, organised folder structure in Mascot Insight.

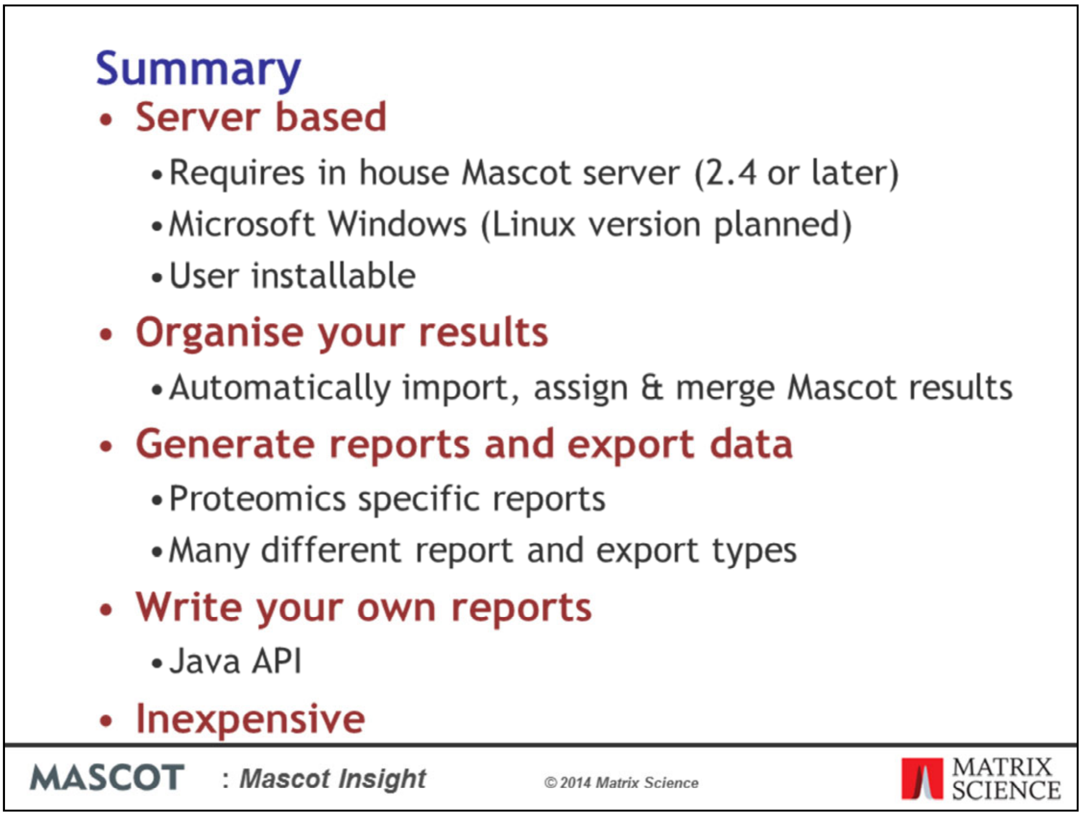


Reports and exports

- Ships with some 30 reports covering
  - Quantitation comparison, qc and clustering
  - Gene ontology
  - Interactions database analysis
  - Shared protein and peptide comparisons
  - Scatter plots, histograms etc
  - Statistics export
  - Publication export (MCP guidelines)
- More in a later presentation….

MASCOT : *Mascot Insight* © 2014 Matrix Science MATRIX SCIENCE

The system ships with some 30 reports designed specifically for proteomics data. These cover a wide range of areas such as result comparison, quantitation and quantitation clustering/grouping, Gene Ontology analysis, Interactions database analysis and general graphing reports such as scatter plots. You can also write your own reports. In a later presentation, Richard will be showing you some examples of the available reports.



Summary

- Server based
  - Requires in house Mascot server (2.4 or later)
  - Microsoft Windows (Linux version planned)
  - User installable
- Organise your results
  - Automatically import, assign & merge Mascot results
- Generate reports and export data
  - Proteomics specific reports
  - Many different report and export types
- Write your own reports
  - Java API
- Inexpensive

MASCOT : *Mascot Insight* © 2014 Matrix Science MATRIX SCIENCE

So, in summary, Mascot Insight is a new server based application – It can be installed on the same server as your Mascot server, but currently only if you're running Mascot on a Windows system (we are planning to release a Linux version). It allows you to easily organise your search results, and allows automatic import and assignment of Mascot search results. It allows you to merge Mascot search results, including all types of quantitation dataset.

The system ships with a large number of reports built into it, which cover areas such as protein and peptide level comparisons between datasets, various quantitation based reports including reports for comparison, ratio clustering and quality control. You can carry out Gene Onotology analysis of your data, including looking for enriched and depleted GO terms compared with a reference proteome and generate reports based on molecular interactions databases in the PSI-MITAB format.

The results of these reports are easily exported from the system, and there are various other export options.

And, if you have a specific report you want to be able to generate, you can write your own reports in Java using a common API across different result formats. Richard will be giving an example of this in his presentation later on.

Finally, Mascot Insight is not expensive to license – licensing cost is related to the number of CPUs licensed for the associated Mascot server, not 'per seat'.