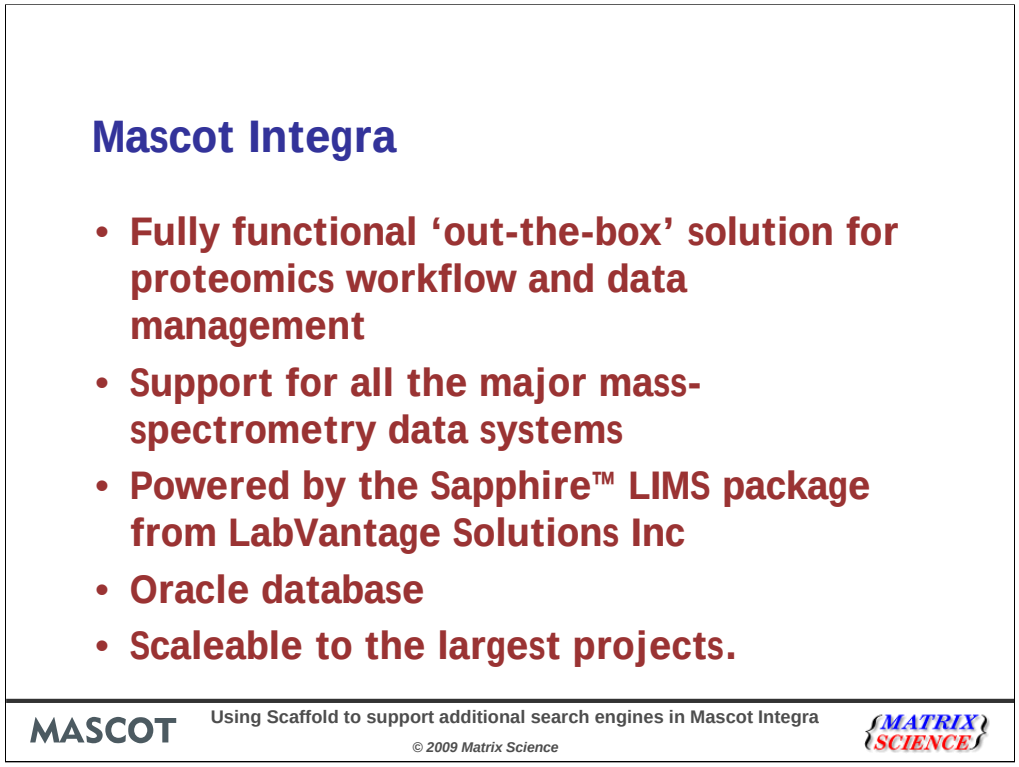
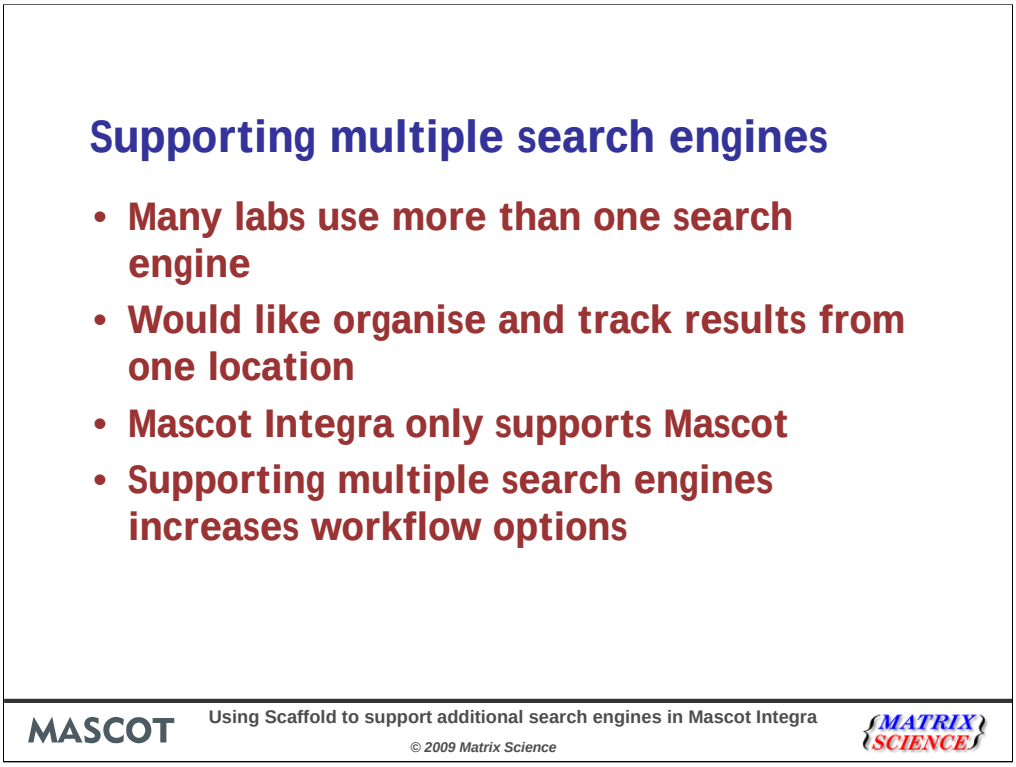
Using Scaffold to support additional search engines in Mascot Integra

MASCOT

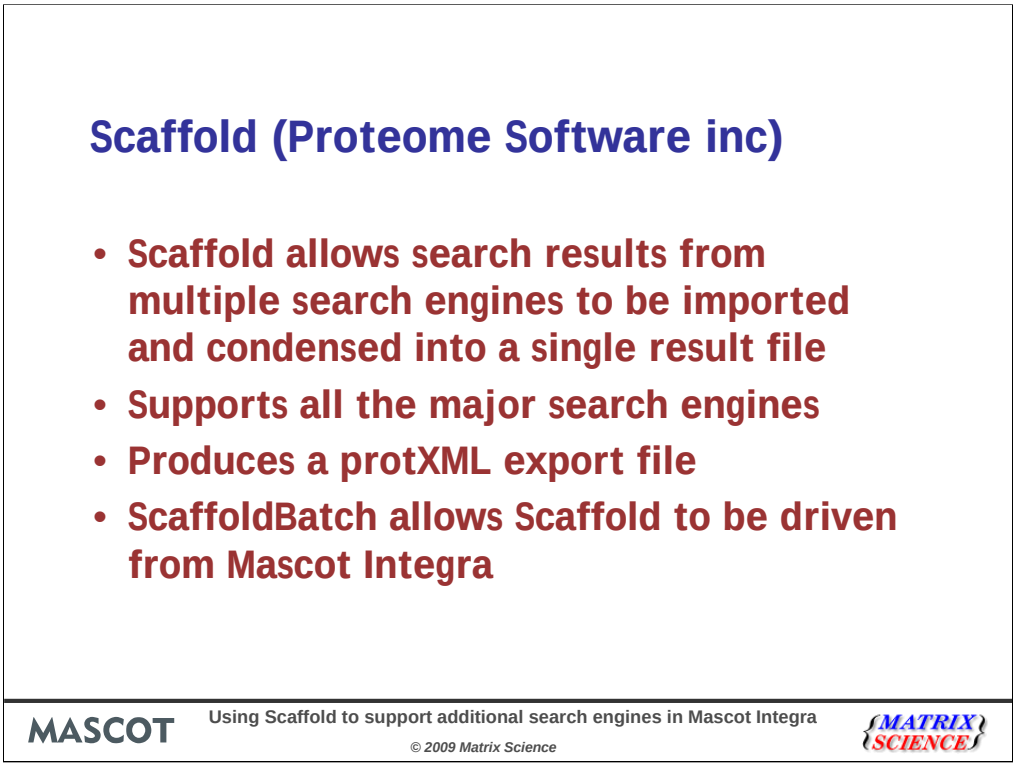
MATRIX SCIENCE



Mascot Integra is our solution for managing and tracking proteomics workflows and results. As with Mascot and Mascot distiller it supports products from all the major mass-spectrometry vendors. Rather than re-invent the wheel, we have partnered with LabVantage Solutions Inc, (www.lims.com). Their Sapphire LIMS package provides the sample tracking and workflow modelling functionality for Mascot Integra. Using the Oracle database management system enables the database to scale efficiently as your data management requirements grow

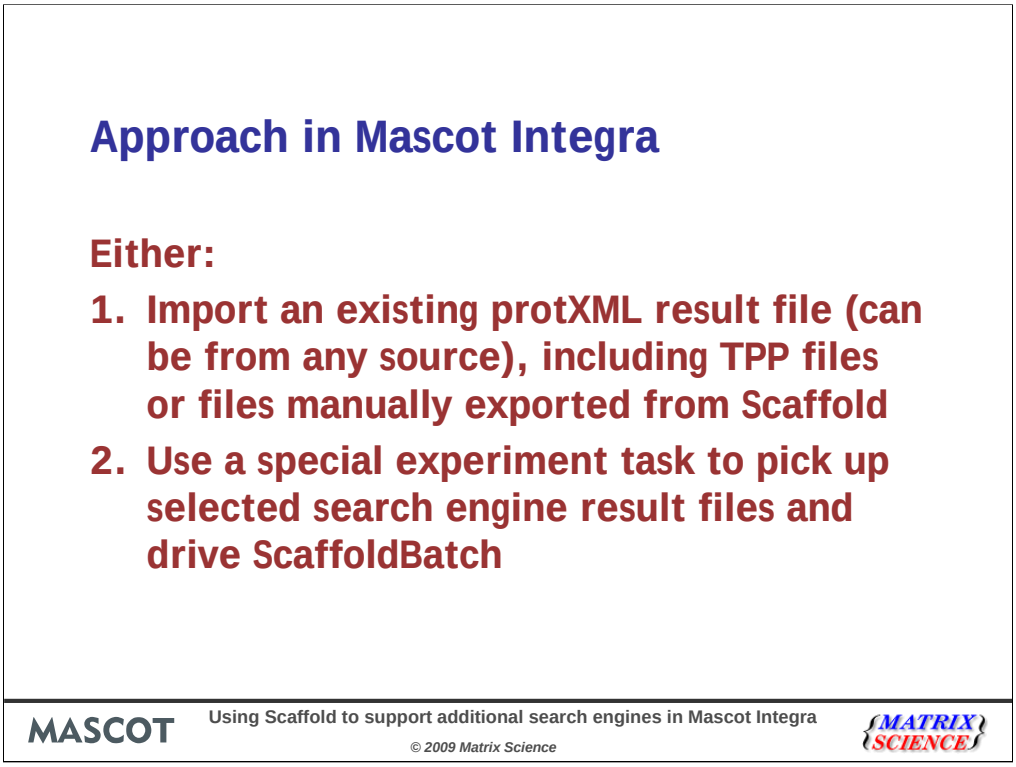


Many labs have more than one search engine (e.g. both Mascot & Sequest) and would like to be able to use a package like Mascot Integra to be able to store, organise and compare/mine data from all of their search results. The current release of Mascot Integra only supports import of Mascot search result files (.dat files), but we would like to increase support for multiple search engines as this would allow us to increase the range of workflows we can support and types of analyses that can be done.



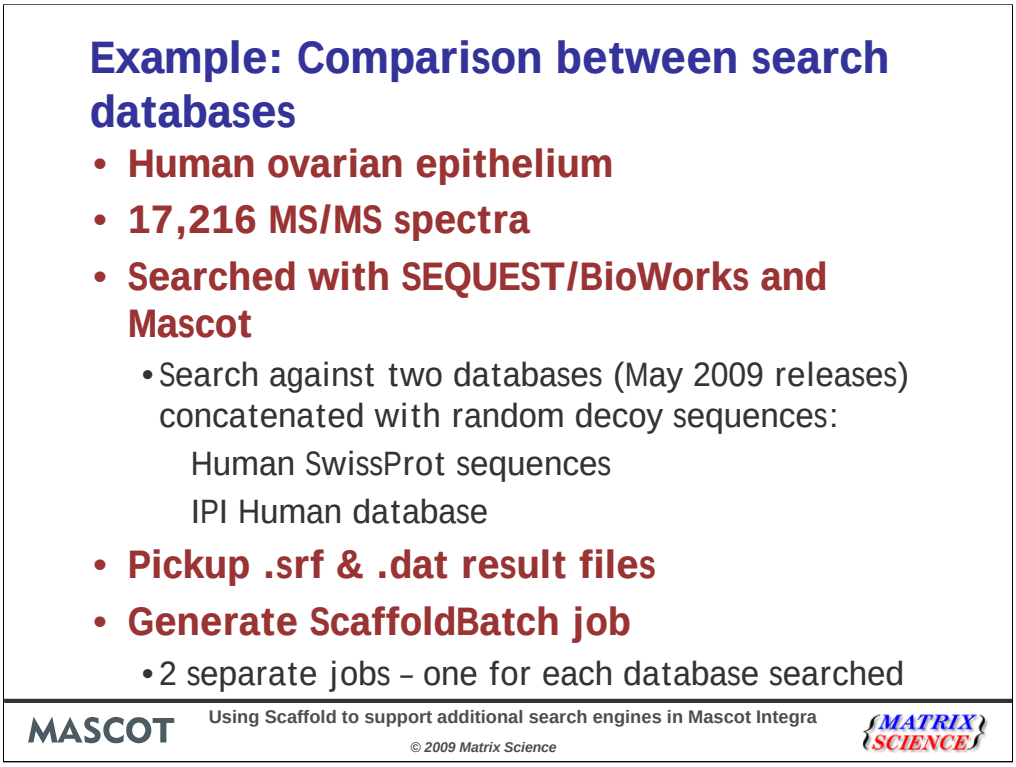
I'm sure most of you are familiar with Scaffold from Proteome Software inc, but I'll just run through some of the key aspects of the package.

Scaffold allows results from many different search engines (including Mascot, SEQUEST, X!Tandem etc) to be imported, compared and condensed into a single result file. In addition to its own result file, it can generate a protXML export of the results. Importantly from our point of view, you can license an additional component of Scaffold – ScaffoldBatch – which allows you to drive batch analysis of results files. This allows us to automatically run Scaffold analysis from Integra, pickup the protXML export file and then import the results in an action we can incorporate into our main experimental workflows.



If you want to import protXML result files into Mascot Integra, you will have two options. You can either pickup and import an existing protXML file, generated from any source (including the TPP if you're running that, or manually exported from Scaffold). Note that Mascot Integra does not directly support the Trans proteomic pipeline and does not support import of pepXML results.

Or you can use a new experimental task from within an Integra experiment to pick up the search result files from whatever search engine you want to use, and then have Integra drive ScaffoldBatch to automatically generate the protXML result file and import it. For most of the rest of this talk we'll be concentrating on approach 2 – using Mascot Integra to drive ScaffoldBatch to automatically run Scaffold, generate the protXML results and then parse and import those results into the Mascot Integra database.



I'll be using an example dataset derived from human ovarian epithelium. This is a large dataset but we're only going to look at a subset it; a single fraction which has 17,216 MS/MS spectra.

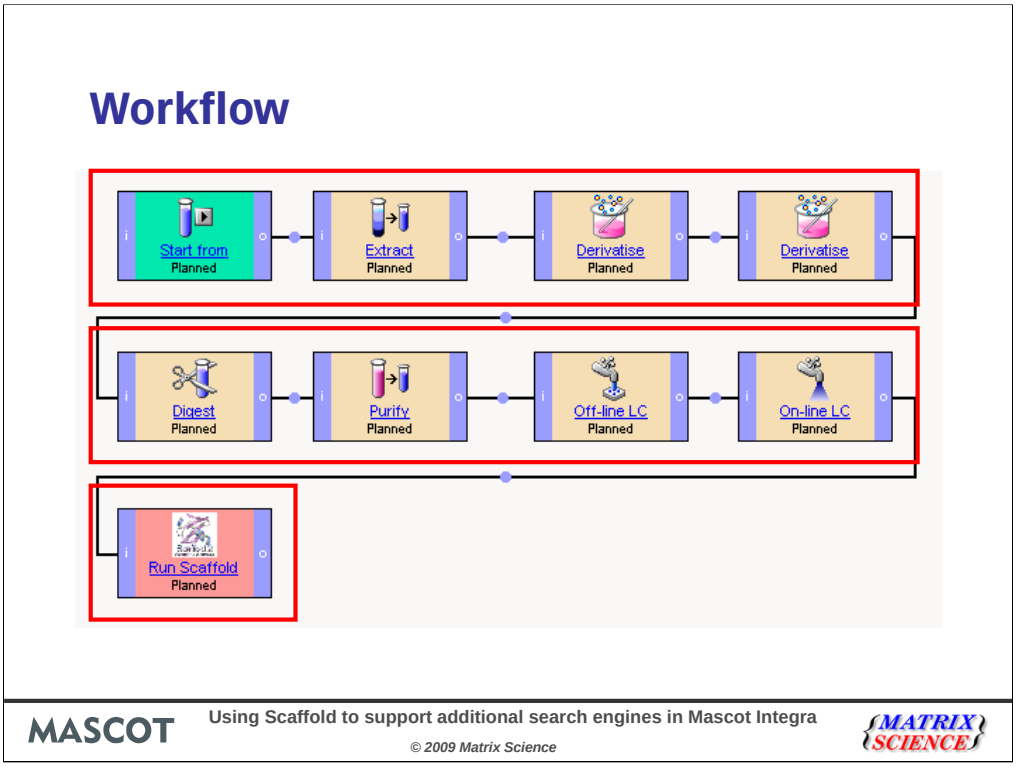
The dataset was searched using SEQUEST from the BioWorks package and with Mascot 2.2 against two different search databases 1: The human sequences from SwissProt and 2: the IPI Human database

Both databases were concatenated with random sequences to give an estimate of the False Positive Rates.

We're going to do a comparison of searching between the two databases. The IPI Human database now contains just over 80,000 protein sequences compared with just over 20,000 human protein sequences in SwissProt – does the increased coverage in IPI Human result in increased coverage in our search results?

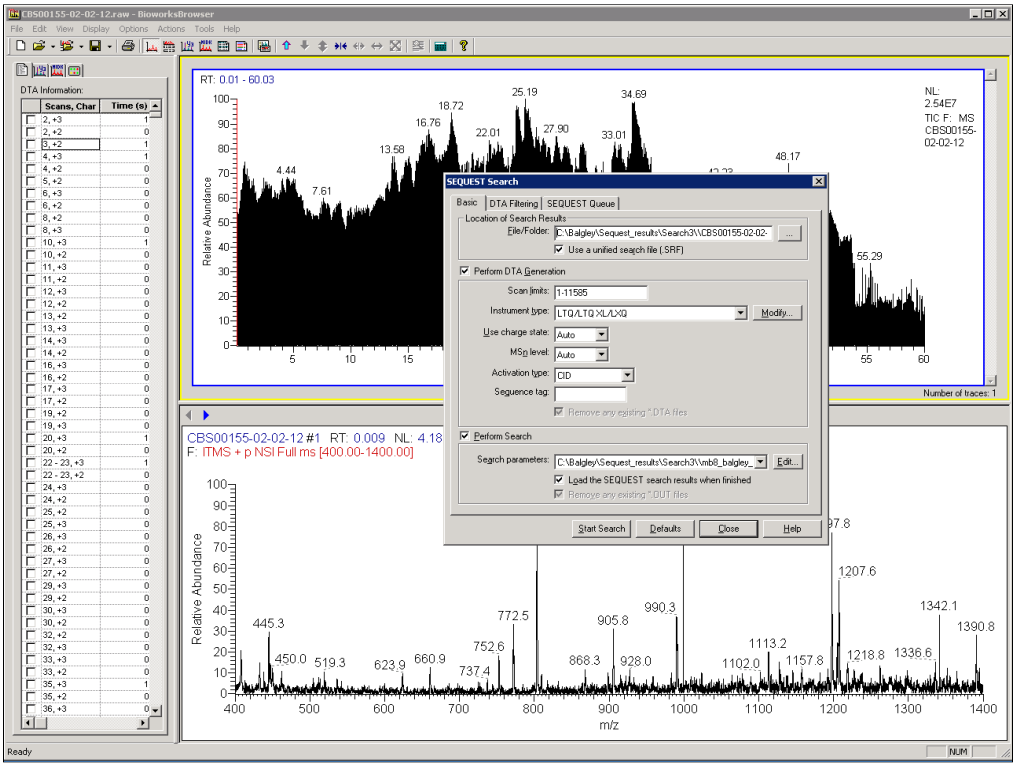
Unlike for Mascot Daemon searches, Integra doesn't try and run the BioWorks SEQUEST job – the end user needs to run the searches themselves and have Mascot Integra pick up the results files.

Once the searches have been completed, Integra can be pointed towards the result files (in this case BioWorks srf file (and also the Mascot .dat results)) and it will then generate the ScaffoldBatch job. Again, here we're going to process each search set differently, so we'll get 2 different Scaffold search results.

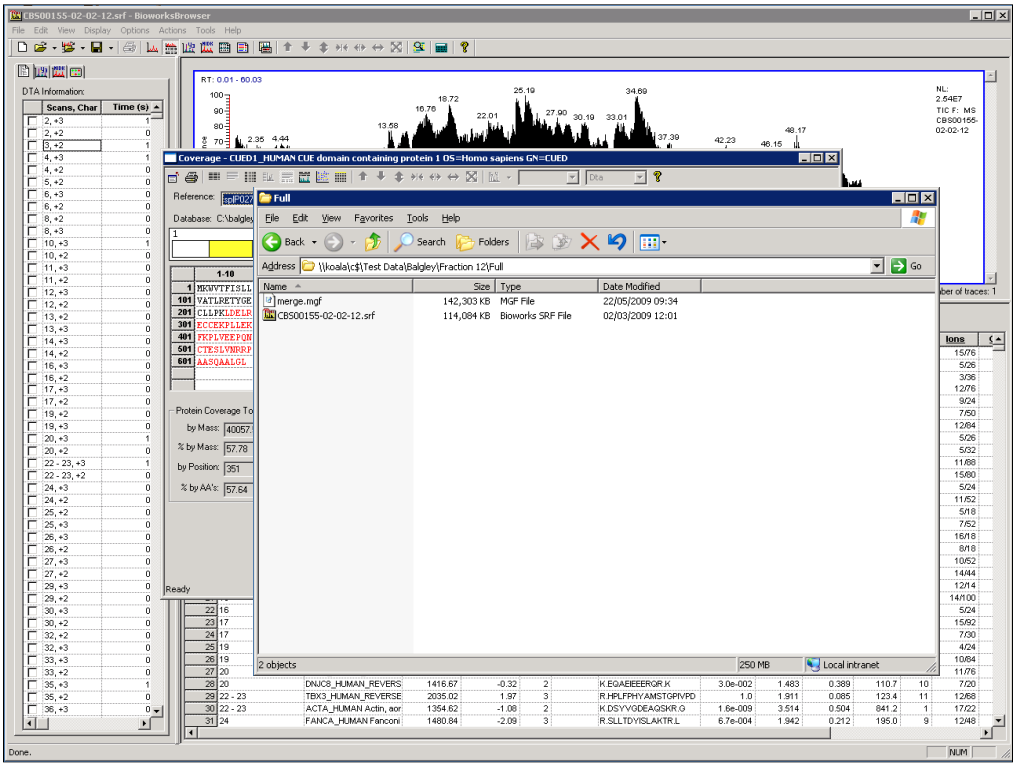


A: Upstream workflow – this tracks the main experimental steps involved in generating the MS/MS samples, so we can track our protein extraction, reduction & alkylation steps, digestion and purification conditions and chromatography steps.

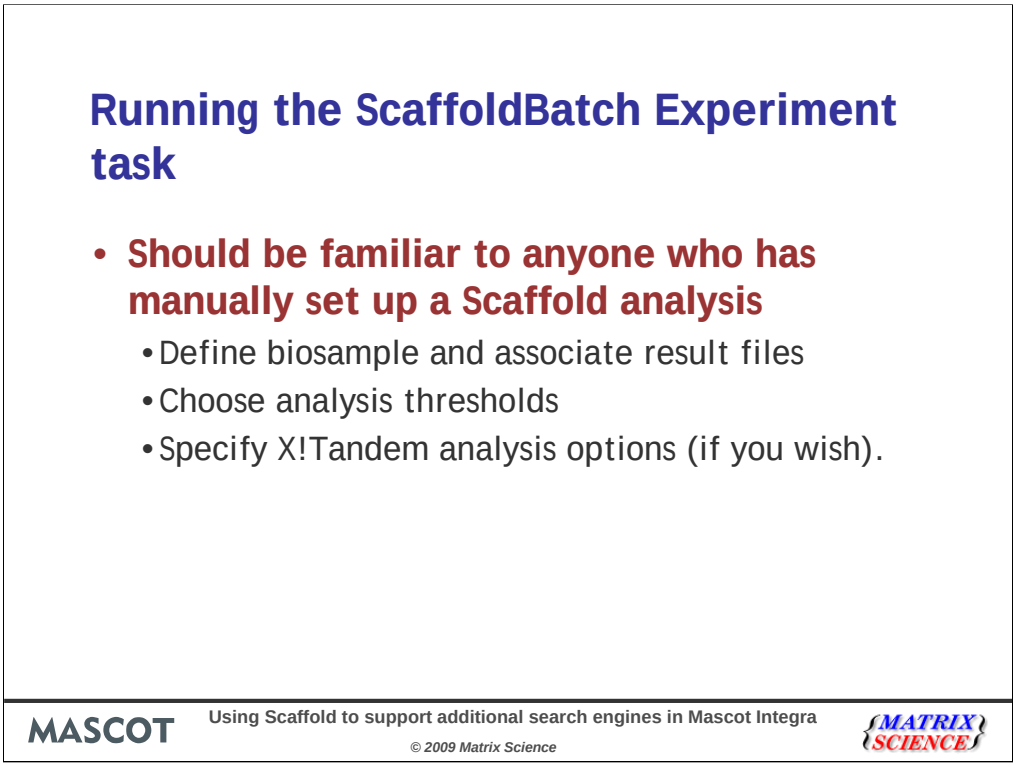
B: Unlike with Mascot searches, Mascot Integra does not try and run the additional search engines – you run the searches as you would normally (e.g. through BioWorks), then point Mascot Integra towards the generated result files.



Here we are in BioWorks with some of the raw data loaded and ready for searching. You just have to submit the search to SEQUEST in the normal way – note that here we are going to produce .SRF unified search result files. Scaffold can handle other SEQUEST result files such as .out files etc, but if you're using BioWorks then the .SRF unified files are much easier to handle.



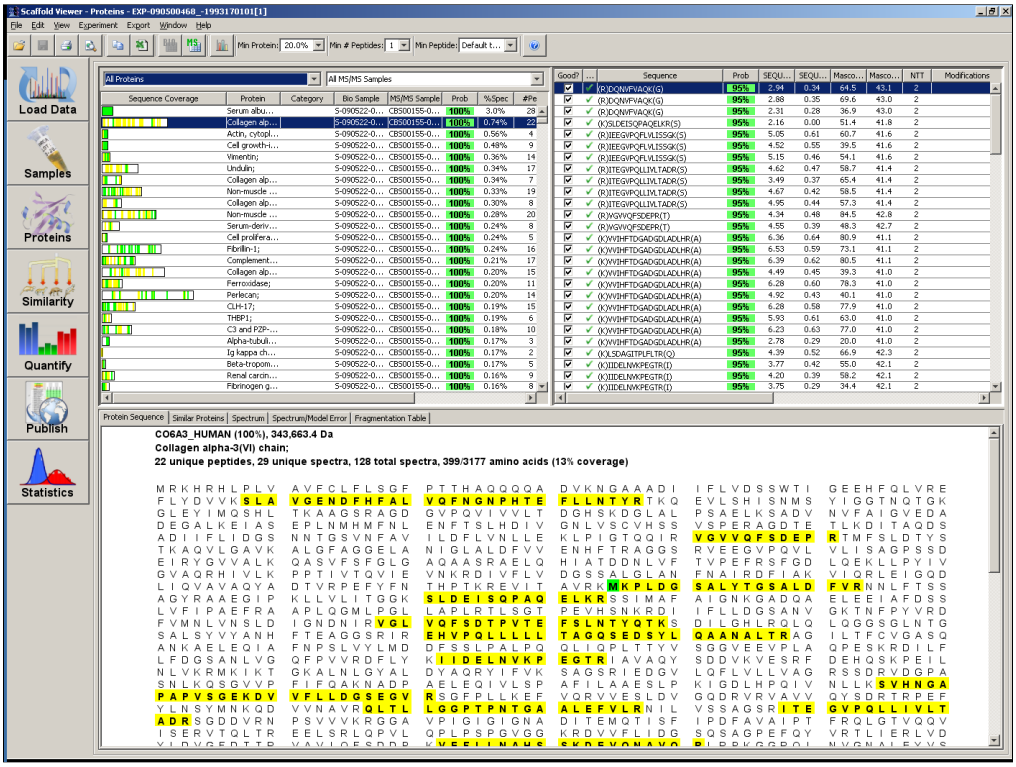
Once the search is completed, you'll have a set of .SRF files – the next step is to point Mascot Integra towards these and get it to setup and run ScaffoldBatch.



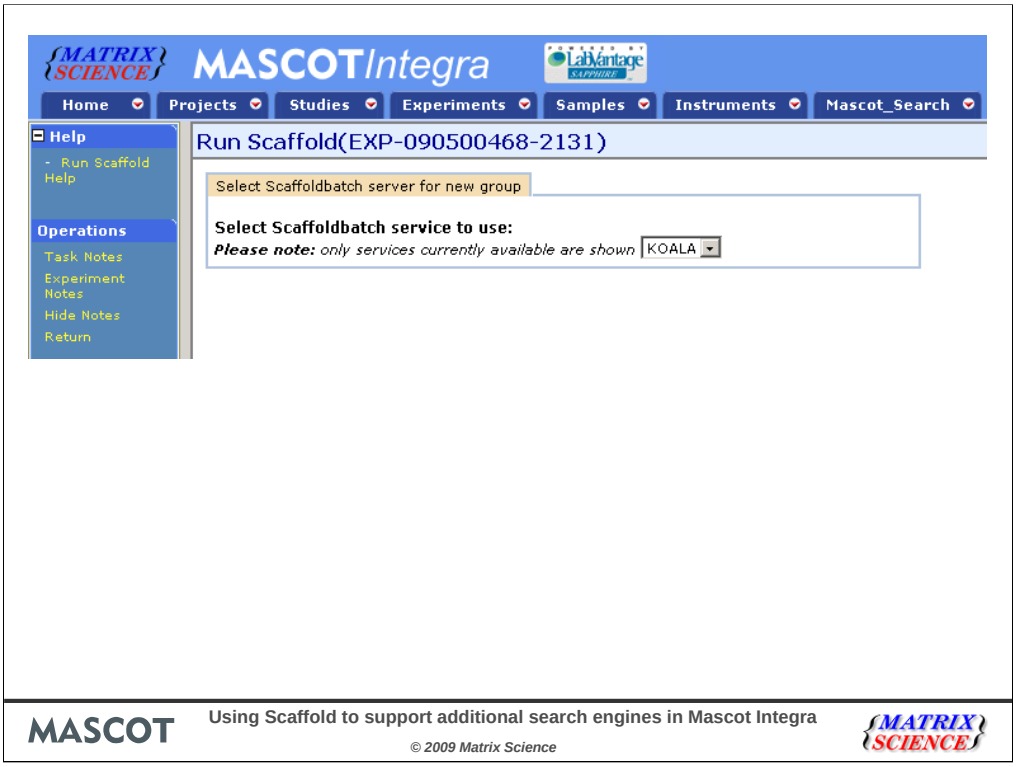
Now we'll run through how you set up and run ScaffoldBatch from within a Mascot Integra experiment. The steps involved should look very familiar to anyone who has run an analysis through the main Scaffold interface.

The three main steps are:

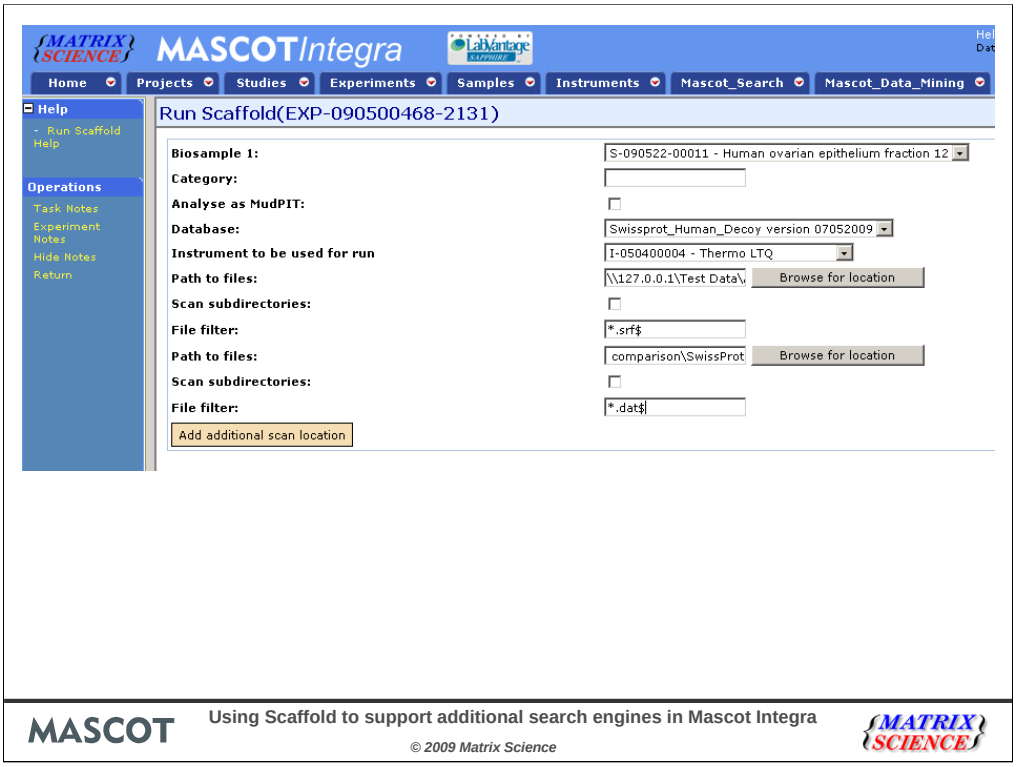
1. Pick your sample (biological sample) and associated the search result files with it
2. Choose the analysis thresholds
3. Specify any integrated X!Tandem analysis options if you want Scaffold to run an X!Tandem search for you



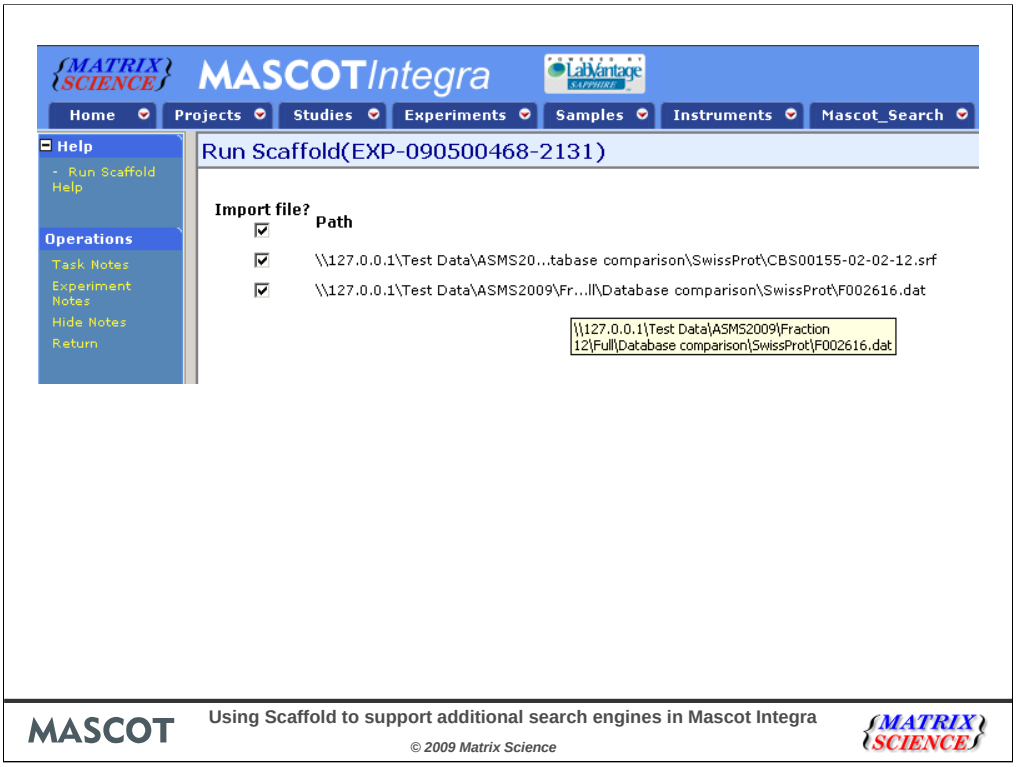
Running the ScaffoldBatch task will produce both a standard Scaffold .sfd result file and a protXML export file. The protXML file will be automatically parsed and imported into the database, and the .sfd file will be attached to the parsed results in the database. The .sfd file will then be available from the database for any user who has access to the parsed results to download and open. Proteome Software provide the free Scaffold viewer for use on client PCs which do not have a Scaffold license allowing users to look at, but not to modify or analyse, the .sfd file. So these are are combined results from SEQUEST and Mascot shown in the Scaffold viewer.



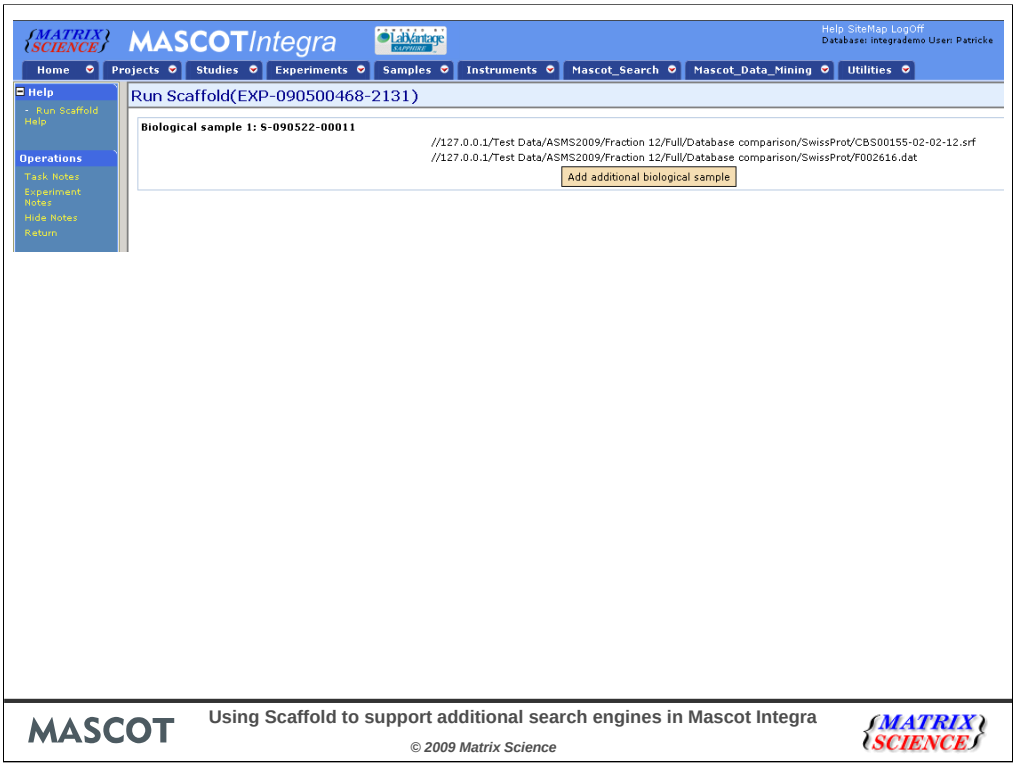
First, you select the system your copy of Scaffold and ScaffoldBatch is installed and licensed on.



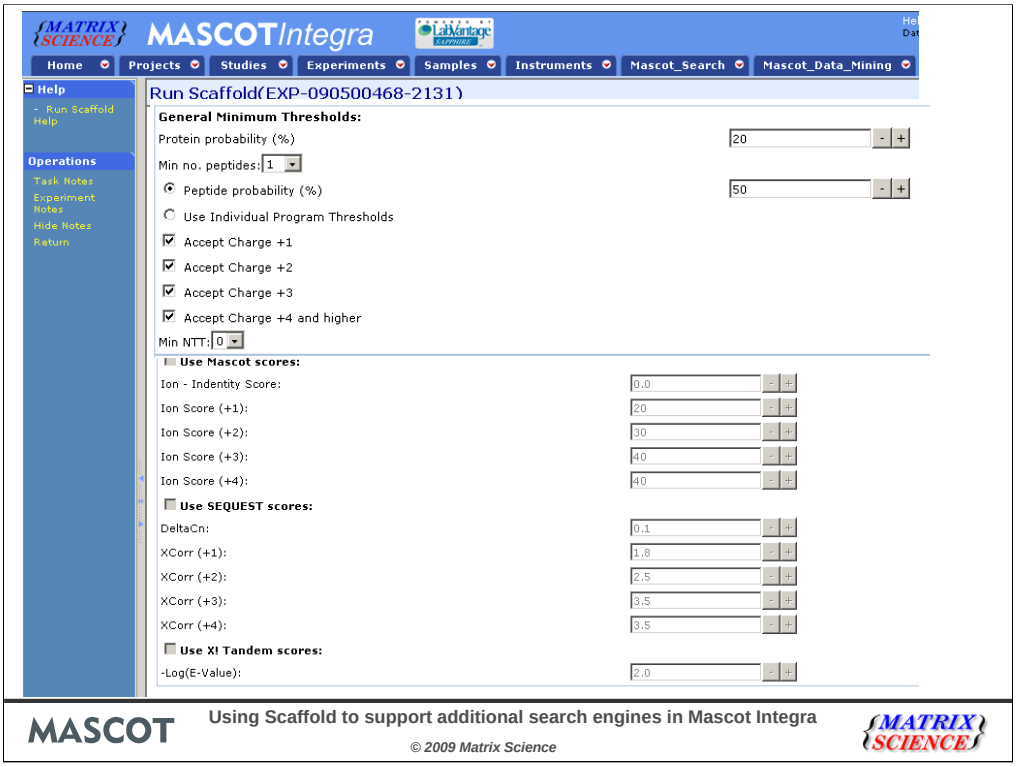
Next you define the Scaffold 'Biosample' – including the FASTA database the search was carried out against, the instrument the MS/MS run was done on and also point the system towards the location the result files are in and specify a filter to pick out the files. You can easily specify multiple scan locations and filters for a single biosample. In this instance, this allows us to pick up both the SEQUEST .SRF and the Mascot .dat result files.



The system then scans the specified location and picks out matching files, which you can then check and exclude and result files you're not interested in submitting to Scaffold.

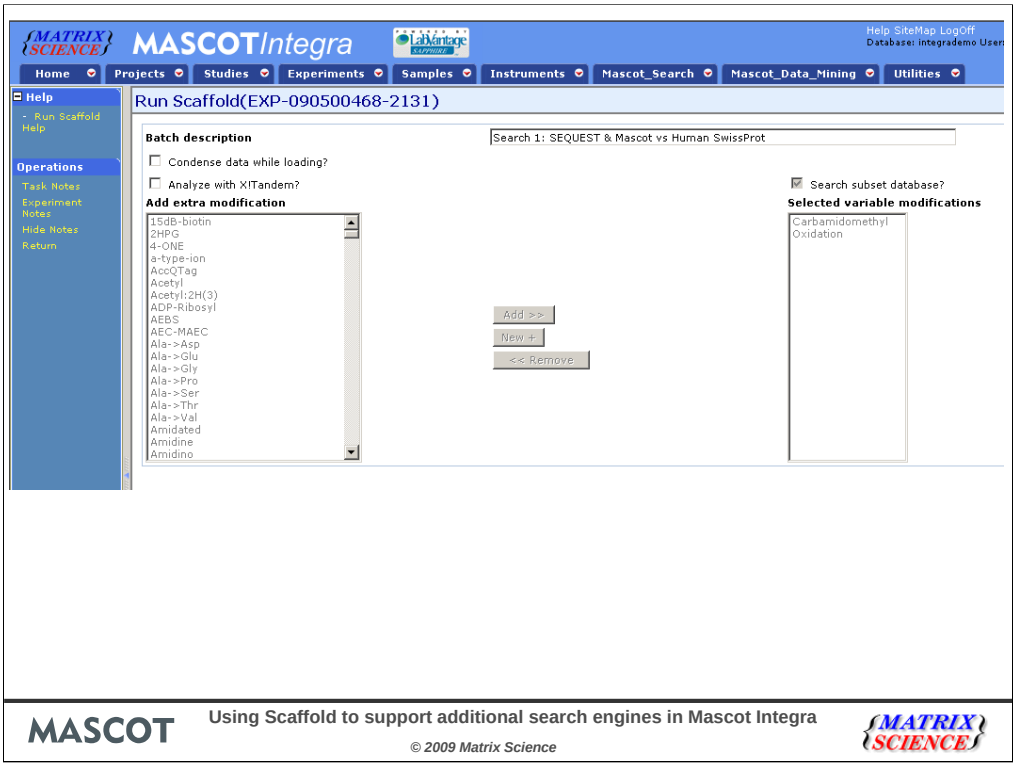


You can then go back and setup another Biosample if you wish.

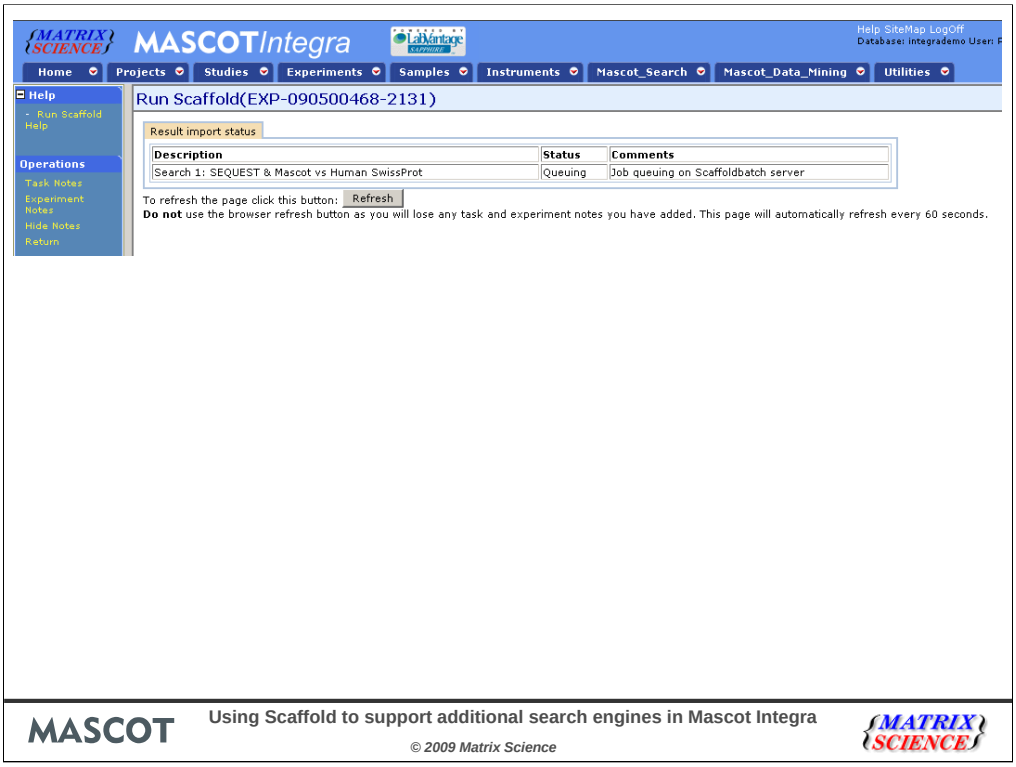


The next step is to specify the export thresholds. The thresholds you specify here will determine which protein and peptide hits are going to be exported into the protXML result file by ScaffoldBatch. Since we can do additional filtering and comparisons once the data is in the Mascot Integra database, in this instance we want to import more than just the high confidence hits, so we're going to change the thresholds.

Now we're going to include all protein hits including one hit wonders, with a peptide probability threshold of 50% and the protein probability threshold of 20 percent. Just like when you're running Scaffold interactively, you can alternatively specify search engine specific thresholds instead. One thing to note if you're running Scaffold on Mascot search results - I've tended to find that Scaffold is quite conservative with Mascot hits. Scaffold is taking its thresholds for Mascot from the identity threshold which we tend to find is quite conservative (if you've attended many of our past user group meetings you'll have seen many talks where we've advocated using the less conservative homology threshold) – for example, for this dataset I'd be quite happy to drop the peptide probability threshold from Scaffold for Mascot matches to 0.8 as this actually gives similar results to using the homology threshold in Mascot and with a low FDR.



The next step is to define any search conditions for use with X!Tandem. Scaffold comes with X!Tandem incorporated into it and you can (optionally) carry out an X!Tandem search for incorporation into the Scaffold results; either against the whole database or just against those protein hits identified in the source results files. You can also specify any additional variable modifications, or define a custom variable modification 'on the fly'.

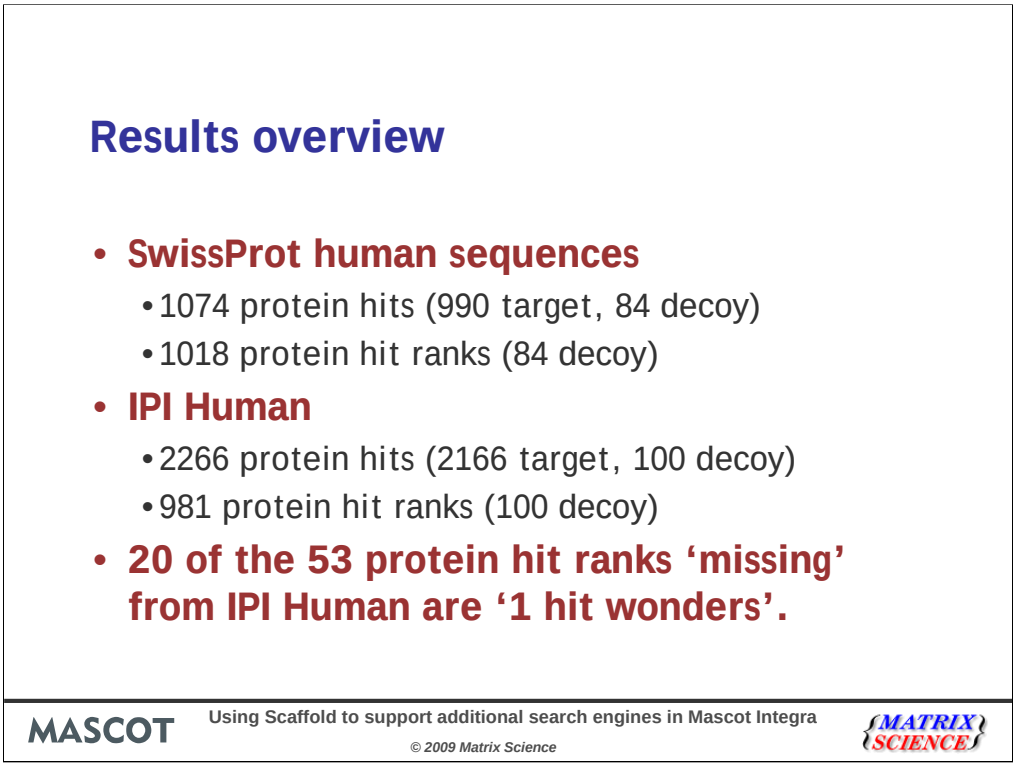


Now, just as with carrying out Mascot searches from Mascot Integra, you can sit back and leave the system to run everything in the background – you don't need to stay logged into the system. It'll run ScaffoldBatch automatically, pick up the resulting files, parse and import the protXML results and attach the Scaffold .sfd result file to the dataset in the database (so any user with access to the results in Integra can easily also access the original Scaffold results).

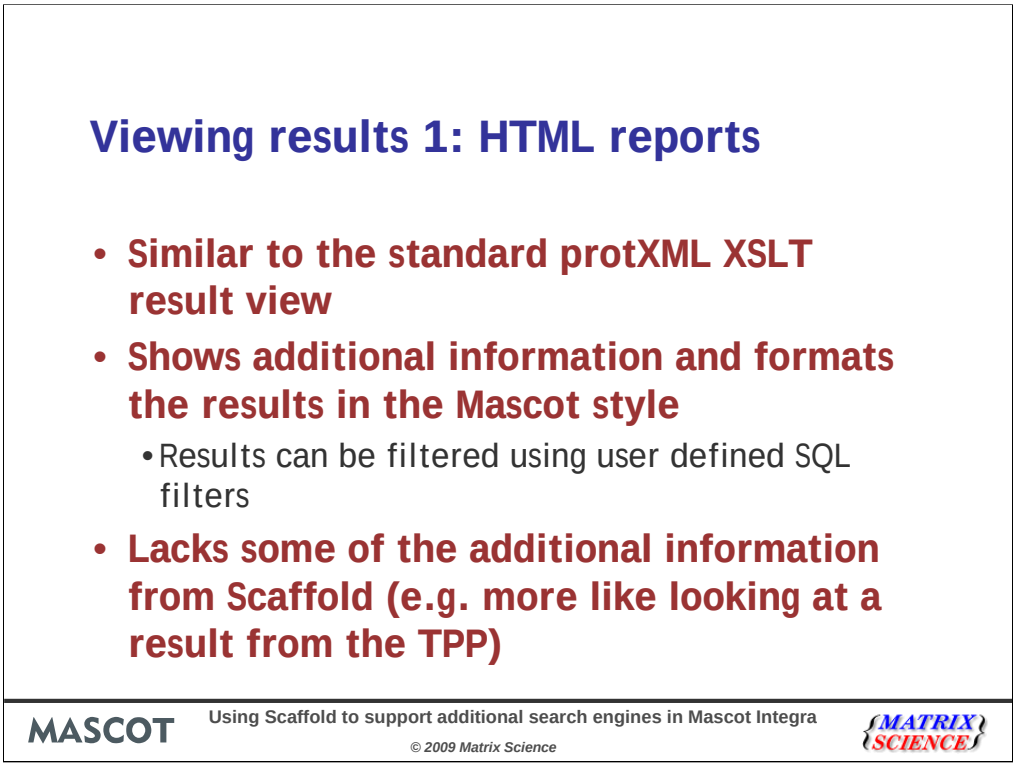
Result Import

- **Additional values not present in protXML are calculated during result parsing/import including:**
  - Protein mass
  - Protein pI
  - Peptide start/end positions and flanking residues
  - Missed enzyme cleavage points
- **Results then appear in lists alongside Mascot results**

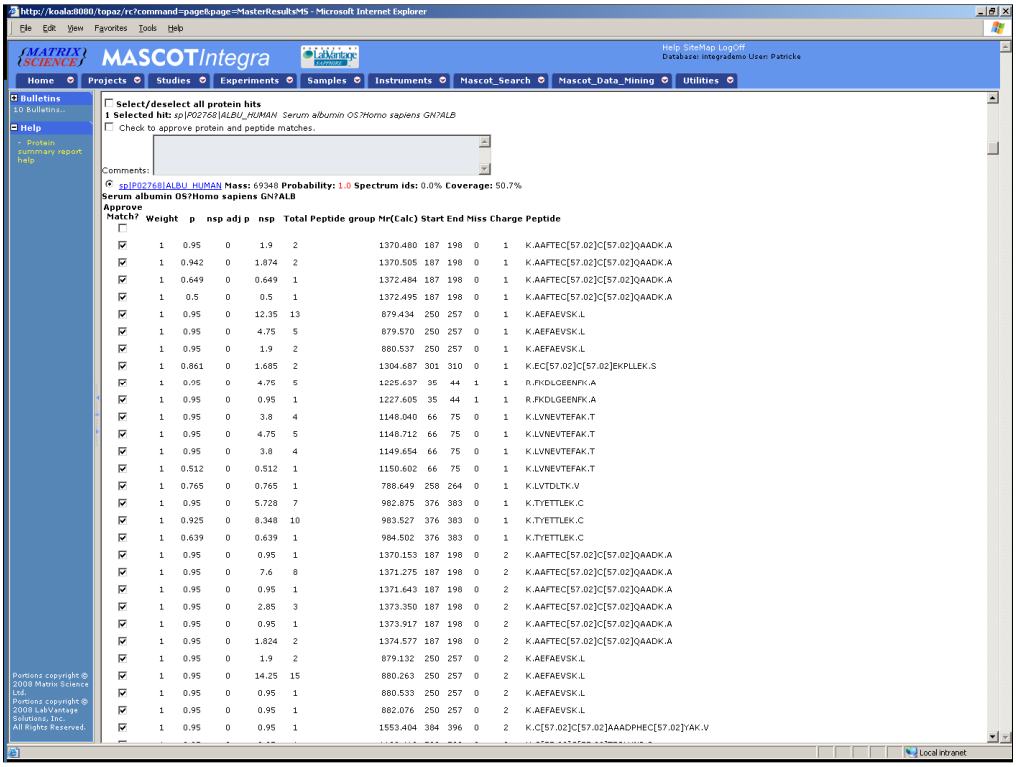
During the parsing phase, Mascot Integra calculates and stores some additional information about the protein hits not included in the protXML file, including the protein mass and pI, peptide start/end & flanking residues, missed enzyme cleavage points in the peptide. Once the results have been imported, then they are accessed in Mascot Integra in the same way as imported Mascot search results.



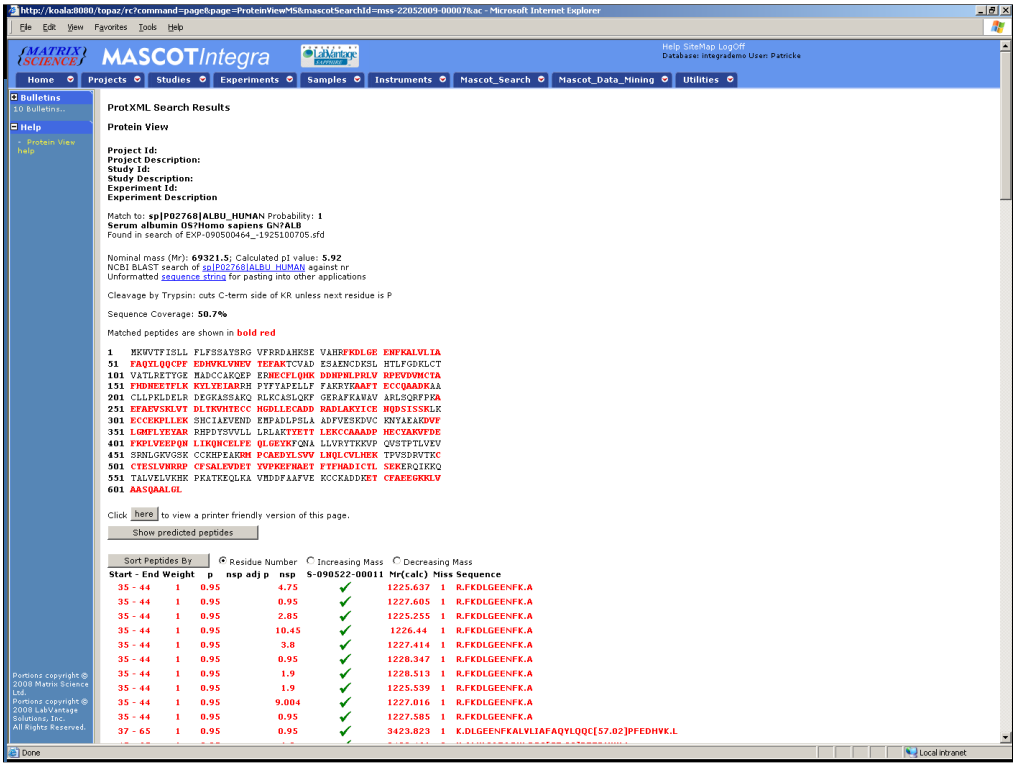
After processing the results with Scaffold we can see that we have many more protein hits from IPI Human with only a slight increase in the number of decoy hits. However, the number of individual groups of proteins has actually slightly decreased (there are 53 fewer target protein hit ranks). This suggests that many of the additional protein hits from IPI Human could be as a result of increased redundancy in IPI Human compared with SwissProt human.



This means that once the protXML results are imported into Mascot Integra, you can use them in much the same way you would any Mascot results imported into the system, including being able to manually validate and approve protein and peptide matches. There are four main ways to view results from within Mascot Integra. Option 1 is to view relatively static HTML based reports. For those of you familiar with the Trans Proteomic Pipeline, these reports show a similar level of information to the standard XSLT reports generated from that. It does include some additional information so that the view is also a bit like looking at a standard Mascot HTML report (master_results report). However, Integra adds some additional functionality such as being able to filter the reports with user definable filters, approve protein hits and view different chunks of the report rather than the whole report in one go.



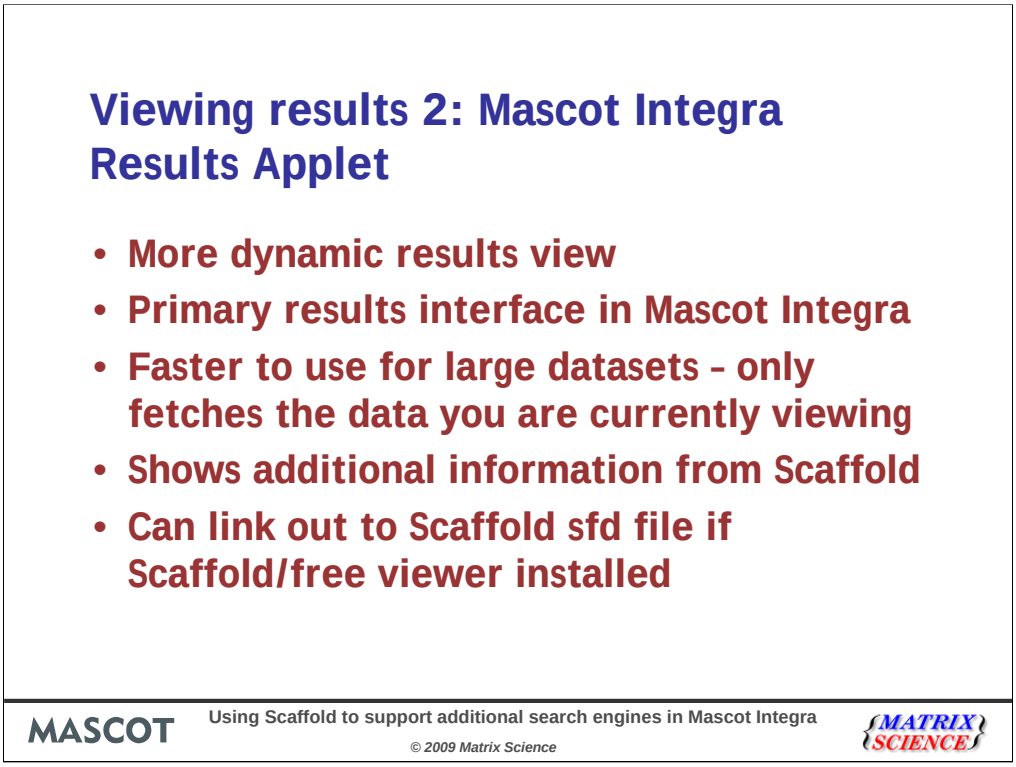
Here we have an example of an HTML report generated from within Integra. As you can see it looks fairly similar to a standard Mascot result report, but we have the additional protXML associated columns, such as Weight, nsp. We also have checkboxes available for manual validation of the protein and peptide hits.



In addition to the main protein hit list, we can get a protein view for any of the hits, showing protein coverage, peptide match positions etc.

What we can't generate though is a peptide fragmentation view though as protXML doesn't include enough information to generate this.

Doesn't contain peptide level information. If you want to see the fragmentation details you'll need to open up the original Scaffold sfd file. Since the ScaffoldBatch task automatically attached/associated the sfd, that file is available for downloading from the Browser and if you have the free Scaffold viewer installed on your client PC, you can open up and view the Scaffold file.

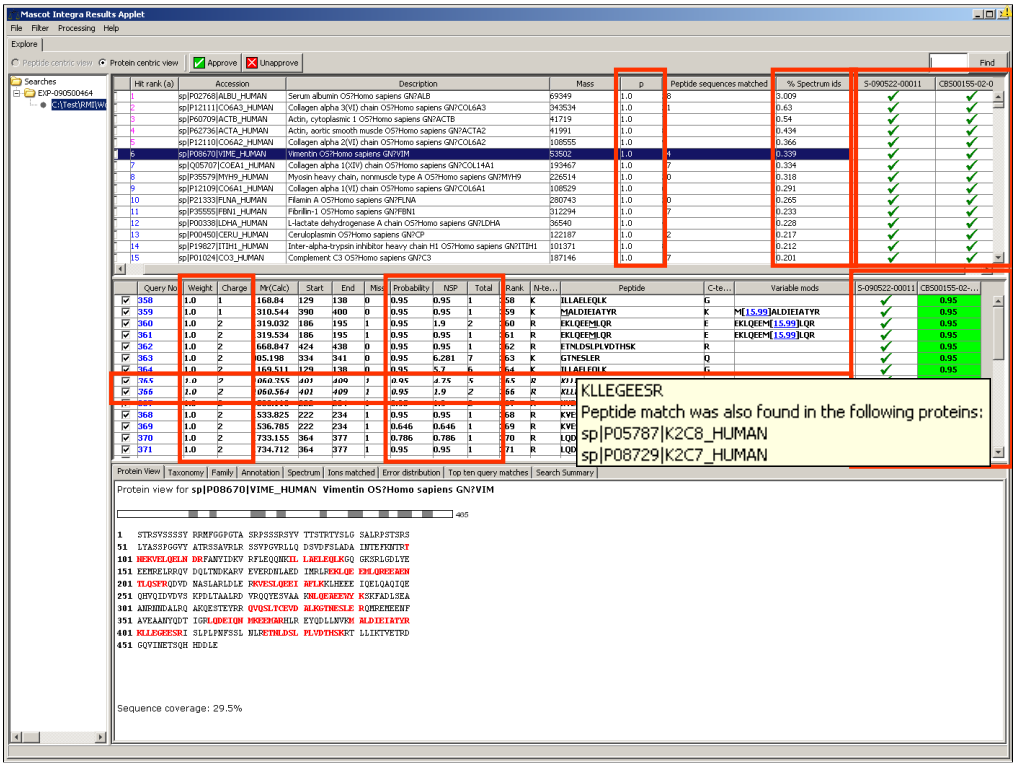


Our recommended result interface in Mascot Integra is the Java results applet which offers a more dynamic result interface and is more like using the main Scaffold viewer than the static HTML reports.

The applet is much faster to use when viewing large datasets as it only fetches the current slice of results you are looking at, making examining large datasets a far more practical prospect. It also offers more advanced and dynamic filtering options (at both a protein and peptide level).

More of the additional information from Scaffold which has been included in the protXML export is shown.

As with the HTML reports, we can't generate a peptide fragmentation view. You can open up the original Scaffold sfd file easily though if you have either Scaffold or the free Scaffold viewer installed on your client PC.

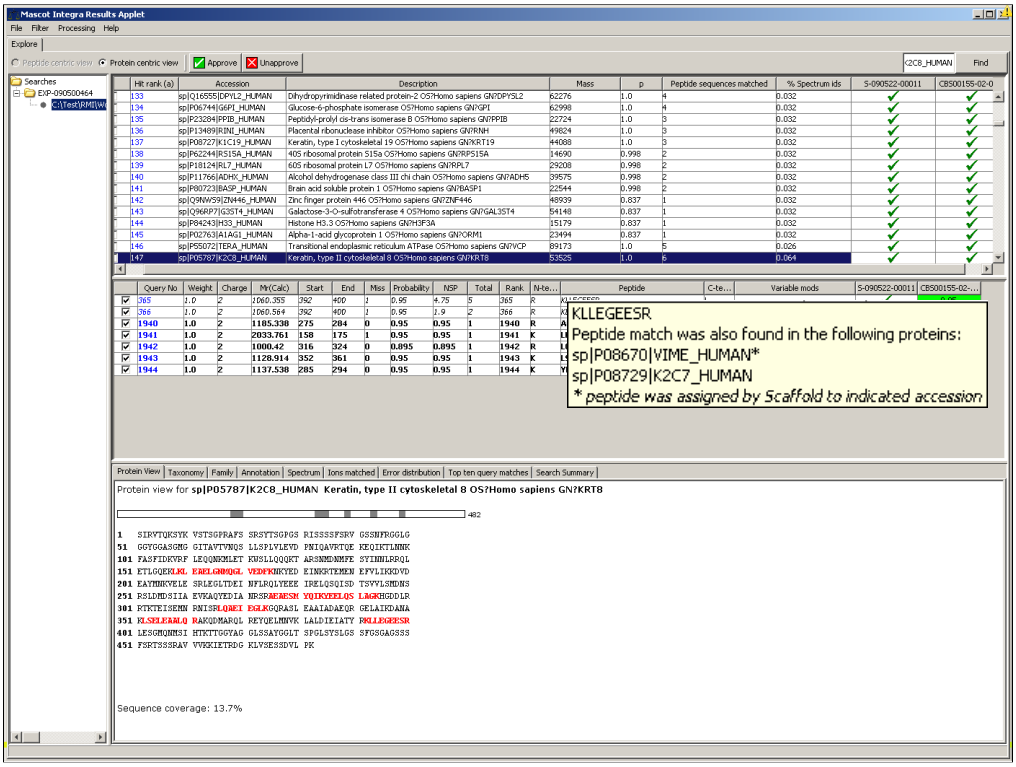


This is one of our Scaffold/protXML results viewed in the Results Applet. Here we can see some of the additional columns available. These include:

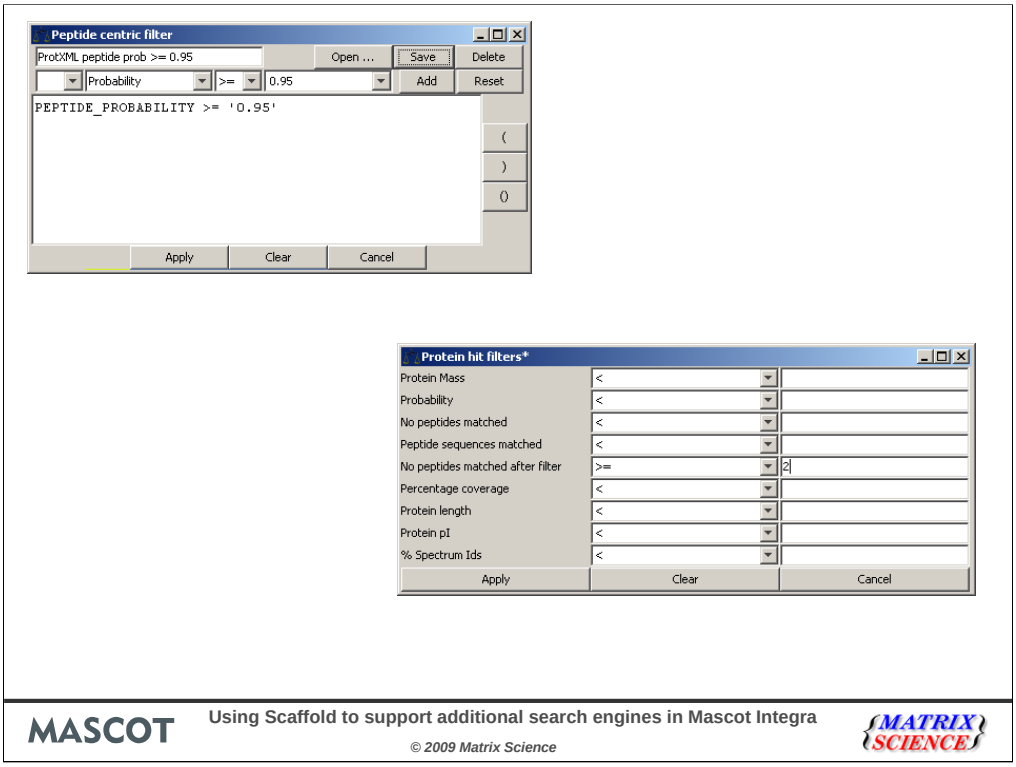
For protein hits – the assigned probability value and the percentage of spectra assigned to the protein hit. For peptide hits – the weighting associated with the peptide and the assigned probability. The other common protXML values such as the number of sibling peptide (NSP) etc are also available – the applet lets you dynamically add an remove columns from the report.

The applet also allows you to view the additional information that comes back from Scaffold as an extension to the standard protXML export. This includes which biological sample each protein or peptide was found associated with, and also details about the presence of the protein or peptide match within each of the source peaklist files the Scaffold results were generated from.

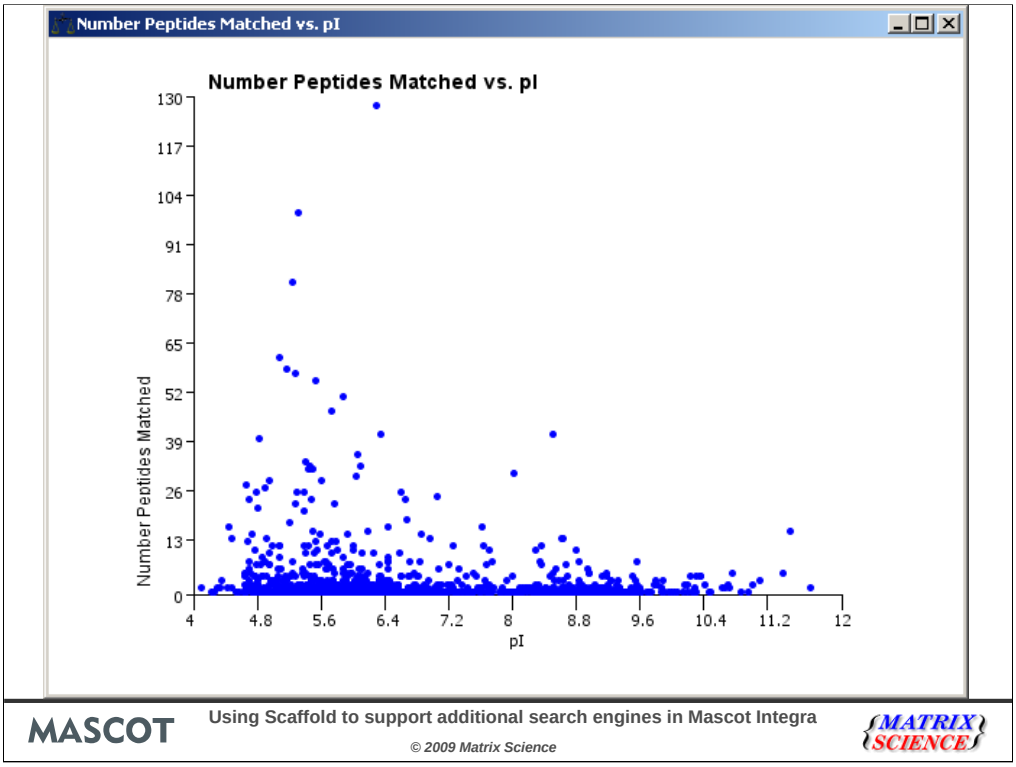
Some additional features – Scaffold tries to unambiguously assign peptide matches to a specific protein hit. If Scaffold has assigned a peptide to the current hit (which means it has been used in calculating the protein hit details such as the protein probability, % spectrum ids etc), then it is shown in bold face as we can see here. However, if there is some ambiguity about the assignment the the applet flags this up by showing the hit in italic. Hovering over the match will bring up a tool tip which includes the information about which additional proteins the peptide could have been associated with. Therefore in this case, these peptides were assigned to VIME_HUMAN by Scaffold, but the italic face is telling us that there are alternative assignments. If we now take a look at one of those possible matches – K2C8_HUMAN



We can see those matches again, but they're not being shown in bold which means Scaffold did not assign the match to this protein. If we take a look at the tooltip, were now getting the additional information that Scaffold in fact assigned these peptides to VIME_HUMAN.



The applet also allows extensive filtering of the results at both the peptide level (via user definable SQL filters) and the protein level



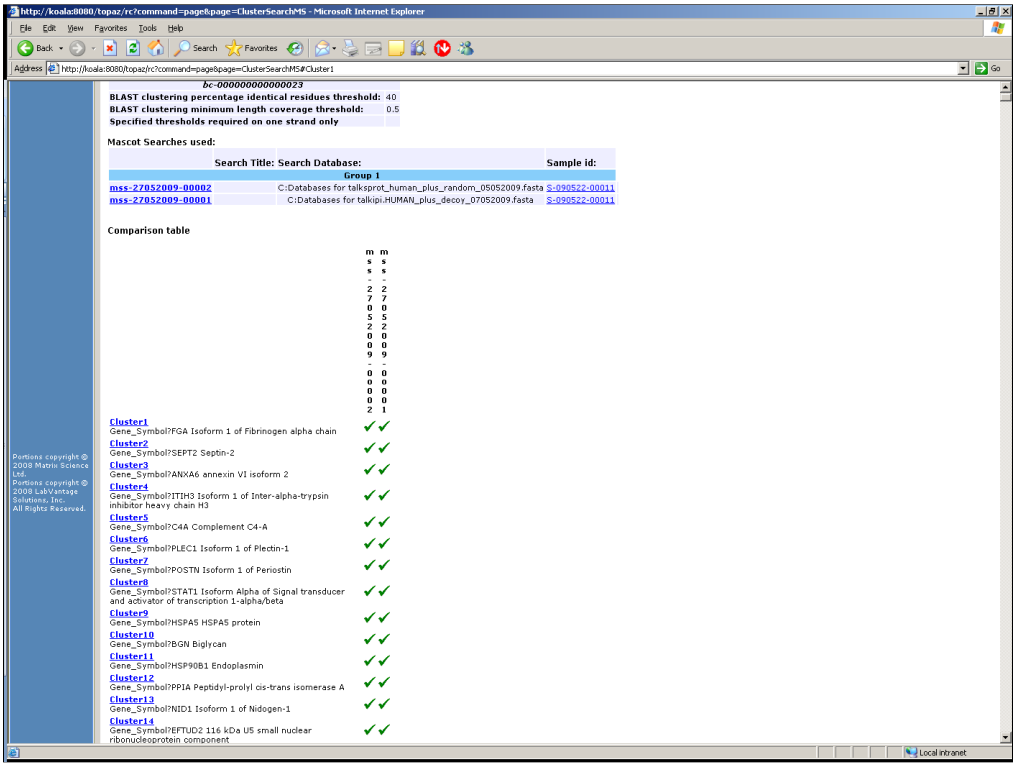
Within the Applet you can also plot graphs from either the protein or peptide level information – here we're plotting the number of peptides matched against the calculated protein isoelectric point; we can see that there is a cluster of proteins with a larger number of peptide matches around the pI range of 4.5-6.4.

Viewing results 3: BlastCluster

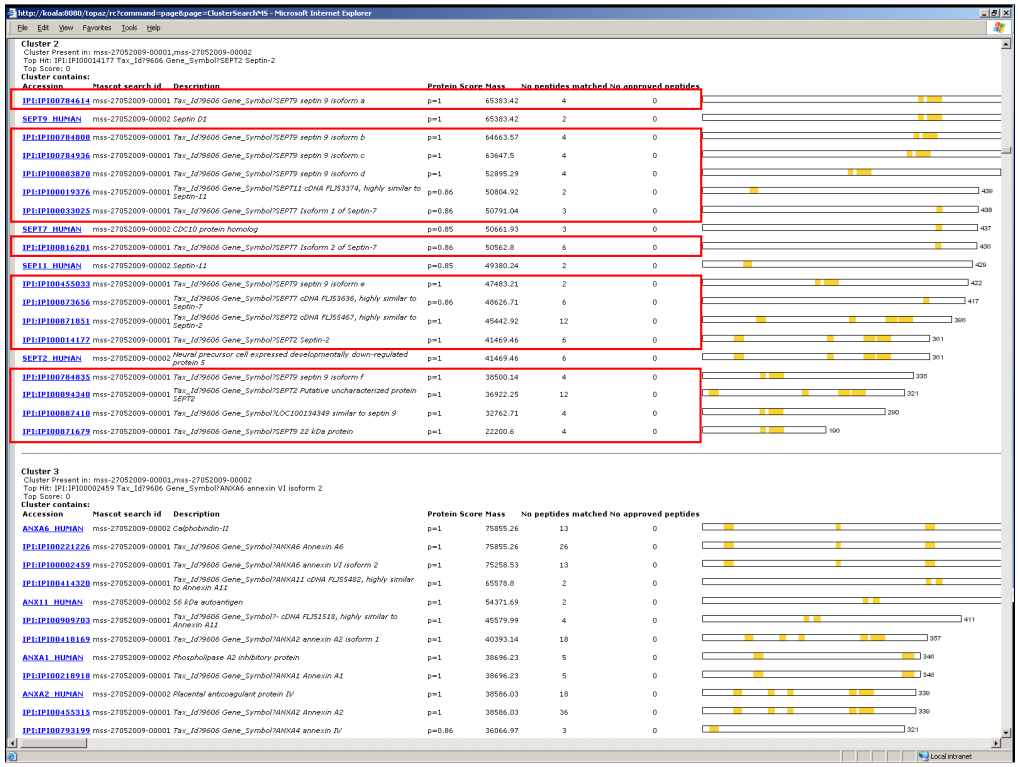
- Uses NCBI BLAST to cluster protein hits together based on sequence homology/identity
- Can use to compare between imported Mascot and protXML results
- Can group together searches based on shared protein clusters

Mascot Integra can produce a cluster report using NCBI BLAST Cluster to group together protein hits based on the primary protein sequence and not on shared peptide matches.

This allows us to generate a report which directly compares between Mascot and protXML searches which have been imported into the system. The source searches can be grouped together based on their shared protein clusters.



Here we've clustered together the results from the two different searches – as we can see the top ranked clusters are present in both searches.

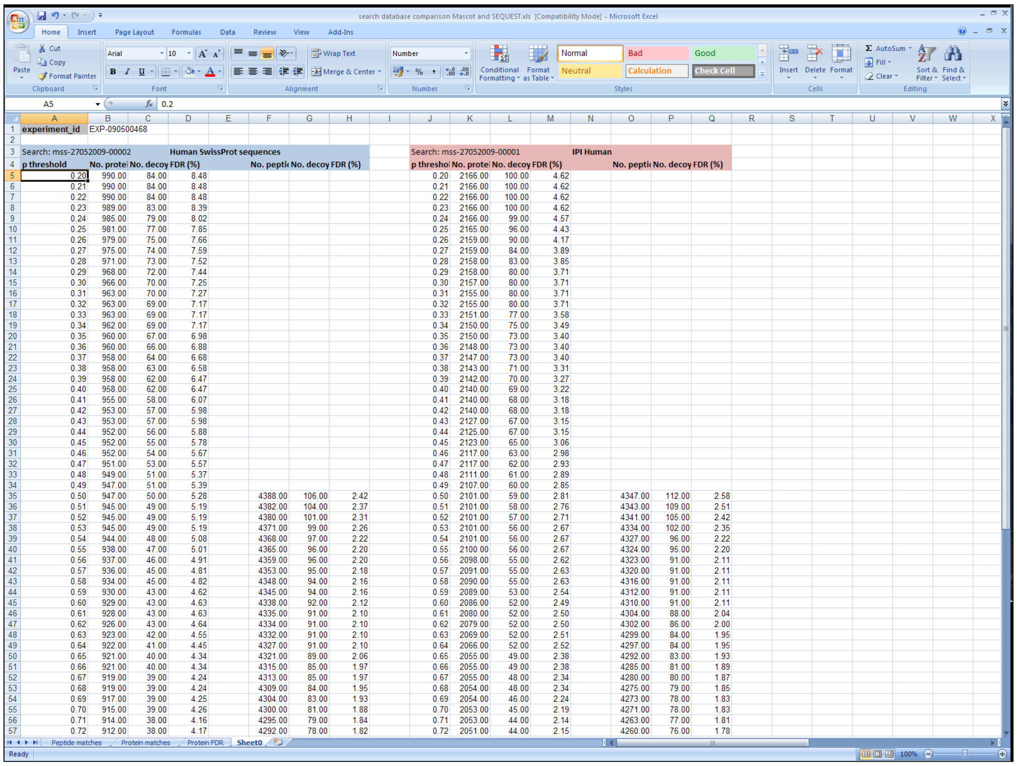


If we take a closer look at cluster 2, we can see that there are far more protein hits from IPI Human in the cluster than there are from the SwissProt human search. The SwissProt search results has matches to Septin 2, 7, 9 and 11. The IPI Human search results have the members of the same four Septin family members, but has additional members of each Septin subgroup present with 6 different isoforms of Septin 9 for example.

Viewing results 4: Custom reports

- Full access to the results schema
- Can generate report based on Excel template or from a Java custom report class
- Allows for generation of more complex, customised reports

The final reporting option is to generate your own custom reports. The Mascot Integra results schema is fully described allowing you to put together your own SQL based reports. These can be generated from either an Excel template or from a custom Java class. Generating the report from Java is the most powerful option, but either method allows you to generate your own, complex and customised reports.



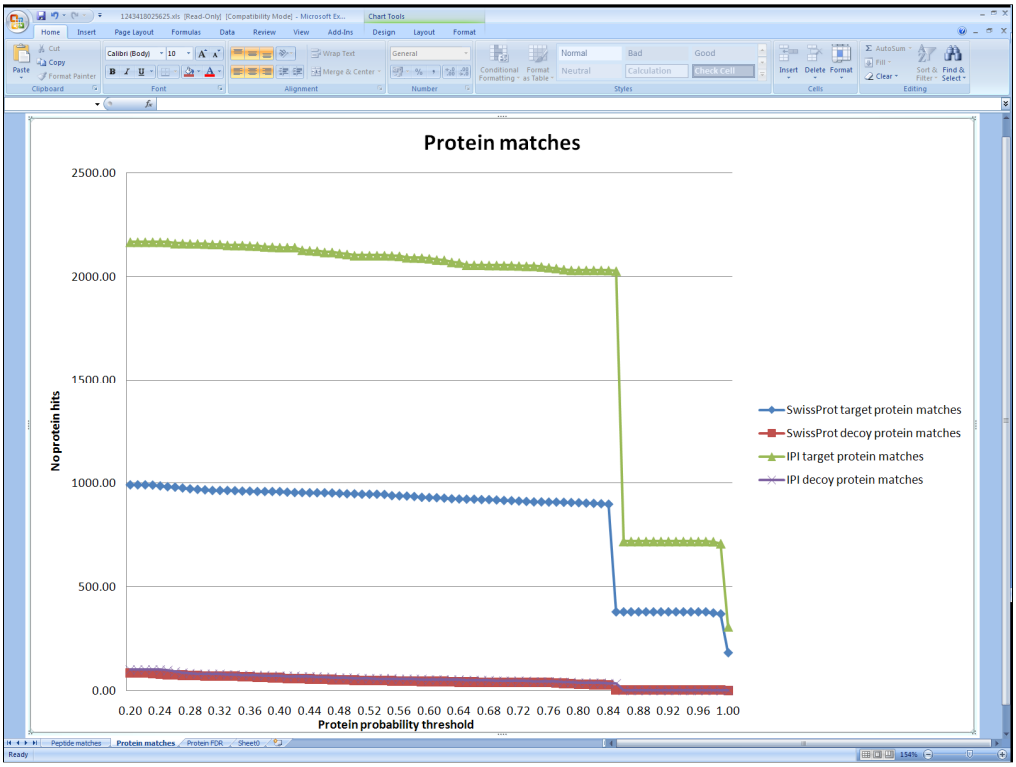
experiment_id	EXP-090500468					
Search: mss-27052009-00002		Human SwissProt sequences				
p threshold	No. protei	No. decoy	FDR (%)	No. pepti	No. decoy	FDR (%)
0.20	990.00	84.00	8.48			
0.21	990.00	84.00	8.48			
0.22	990.00	84.00	8.48			
0.23	989.00	83.00	8.39			
0.24	985.00	79.00	8.02			
0.25	981.00	77.00	7.85			
0.26	979.00	75.00	7.66			
0.27	975.00	74.00	7.59			
0.28	971.00	73.00	7.52			
0.29	968.00	72.00	7.44			
0.30	966.00	70.00	7.25			
0.31	963.00	70.00	7.27			
0.32	963.00	69.00	7.17			
0.33	963.00	69.00	7.17			
0.34	962.00	69.00	7.17			
0.35	960.00	67.00	6.98			
0.36	960.00	66.00	6.88			
0.37	958.00	64.00	6.68			
0.38	958.00	63.00	6.58			
0.39	958.00	62.00	6.47			
0.40	958.00	62.00	6.47			
0.41	955.00	58.00	6.07			
0.42	953.00	57.00	5.98			
0.43	953.00	57.00	5.98			
0.44	952.00	56.00	5.88			
0.45	952.00	55.00	5.78			
0.46	952.00	54.00	5.67			
0.47	951.00	53.00	5.57			
0.48	949.00	51.00	5.37			
0.49	947.00	51.00	5.39			
0.50	947.00	50.00	5.28	4388.00	106.00	2.42
0.51	945.00	49.00	5.19	4382.00	104.00	2.37
0.52	945.00	49.00	5.19	4380.00	101.00	2.31
0.53	945.00	49.00	5.19	4371.00	99.00	2.26
0.54	944.00	48.00	5.08	4368.00	97.00	2.22
0.55	938.00	47.00	5.01	4365.00	96.00	2.20
0.56	937.00	46.00	4.91	4359.00	96.00	2.20
0.57	936.00	45.00	4.81	4353.00	95.00	2.18
0.58	934.00	45.00	4.82	4348.00	94.00	2.16
0.59	930.00	43.00	4.62	4345.00	94.00	2.16
0.60	929.00	43.00	4.63	4338.00	92.00	2.12
0.61	928.00	43.00	4.63	4335.00	91.00	2.10
0.62	926.00	43.00	4.64	4334.00	91.00	2.10
0.63	923.00	42.00	4.55	4332.00	91.00	2.10
0.64	922.00	41.00	4.45	4327.00	91.00	2.10
0.65	921.00	40.00	4.34	4321.00	89.00	2.06
0.66	921.00	40.00	4.34	4315.00	85.00	1.97
0.67	919.00	39.00	4.24	4313.00	85.00	1.97
0.68	919.00	39.00	4.24	4309.00	84.00	1.95
0.69	917.00	39.00	4.25	4304.00	83.00	1.93
0.70	915.00	39.00	4.26	4300.00	81.00	1.88
0.71	914.00	38.00	4.16	4295.00	79.00	1.84
0.72	912.00	38.00	4.17	4292.00	78.00	1.82

Search: mss-27052009-00001			IPI Human			
p threshol	No. protei	No. decoy	FDR (%)	No. pepti	No. decoy	FDR (%)
0.20	2166.00	100.00	4.62			
0.21	2166.00	100.00	4.62			
0.22	2166.00	100.00	4.62			
0.23	2166.00	100.00	4.62			
0.24	2166.00	99.00	4.57			
0.25	2165.00	96.00	4.43			
0.26	2159.00	90.00	4.17			
0.27	2159.00	84.00	3.89			
0.28	2158.00	83.00	3.85			
0.29	2158.00	80.00	3.71			
0.30	2157.00	80.00	3.71			
0.31	2155.00	80.00	3.71			
0.32	2155.00	80.00	3.71			
0.33	2151.00	77.00	3.58			
0.34	2150.00	75.00	3.49			
0.35	2150.00	73.00	3.40			
0.36	2148.00	73.00	3.40			
0.37	2147.00	73.00	3.40			
0.38	2143.00	71.00	3.31			
0.39	2142.00	70.00	3.27			
0.40	2140.00	69.00	3.22			
0.41	2140.00	68.00	3.18			
0.42	2140.00	68.00	3.18			
0.43	2127.00	67.00	3.15			
0.44	2125.00	67.00	3.15			
0.45	2123.00	65.00	3.06			
0.46	2117.00	63.00	2.98			
0.47	2117.00	62.00	2.93			
0.48	2111.00	61.00	2.89			
0.49	2107.00	60.00	2.85			
0.50	2101.00	59.00	2.81	4347.00	112.00	2.58
0.51	2101.00	58.00	2.76	4343.00	109.00	2.51
0.52	2101.00	57.00	2.71	4341.00	105.00	2.42
0.53	2101.00	56.00	2.67	4334.00	102.00	2.35
0.54	2101.00	56.00	2.67	4327.00	96.00	2.22
0.55	2100.00	56.00	2.67	4324.00	95.00	2.20
0.56	2098.00	55.00	2.62	4323.00	91.00	2.11
0.57	2091.00	55.00	2.63	4320.00	91.00	2.11
0.58	2090.00	55.00	2.63	4316.00	91.00	2.11
0.59	2089.00	53.00	2.54	4312.00	91.00	2.11
0.60	2086.00	52.00	2.49	4310.00	91.00	2.11
0.61	2080.00	52.00	2.50	4304.00	88.00	2.04
0.62	2079.00	52.00	2.50	4302.00	86.00	2.00
0.63	2069.00	52.00	2.51	4299.00	84.00	1.95
0.64	2066.00	52.00	2.52	4297.00	84.00	1.95
0.65	2055.00	49.00	2.38	4292.00	83.00	1.93
0.66	2055.00	49.00	2.38	4285.00	81.00	1.89
0.67	2055.00	48.00	2.34	4280.00	80.00	1.87
0.68	2054.00	48.00	2.34	4275.00	79.00	1.85
0.69	2054.00	46.00	2.24	4273.00	78.00	1.83
0.70	2053.00	45.00	2.19	4271.00	78.00	1.83
0.71	2053.00	44.00	2.14	4263.00	77.00	1.81
0.72	2051.00	44.00	2.15	4260.00	76.00	1.78

In this example, I've written a custom Java class to generate an Excel sheet which pulls out the target and decoy protein and peptide matches for each search at different probability thresholds. Because the data has been exported from Scaffold, unassigned peptide matches are not included – only peptide matches assigned to a protein hit are present. Scaffold does try and assign a spectrum unambiguously to a protein hit – where ambiguity was reported (as we saw in the results earlier), then this will be taken into account (if a peptide is assigned to both a target and decoy sequence for example).

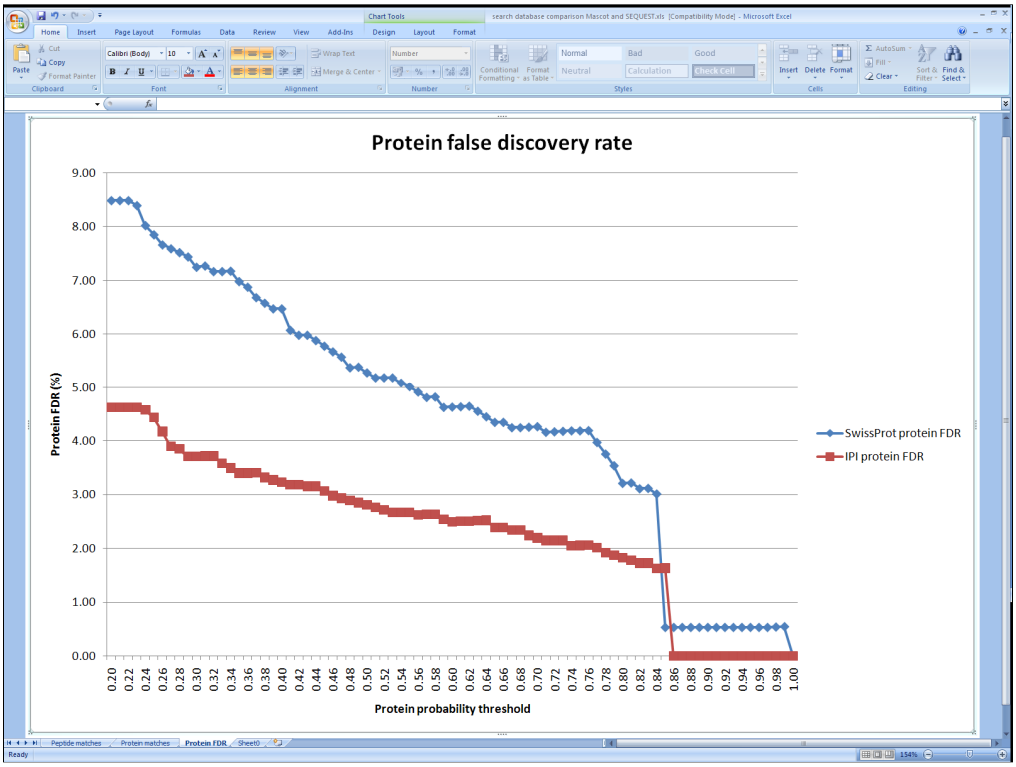
To remind you of the two sets of search conditions:

1. Search against the human protein sequences in SwissProt, approx 20,000 sequences
2. Search against the IPI Human database, approx 80,000 sequences



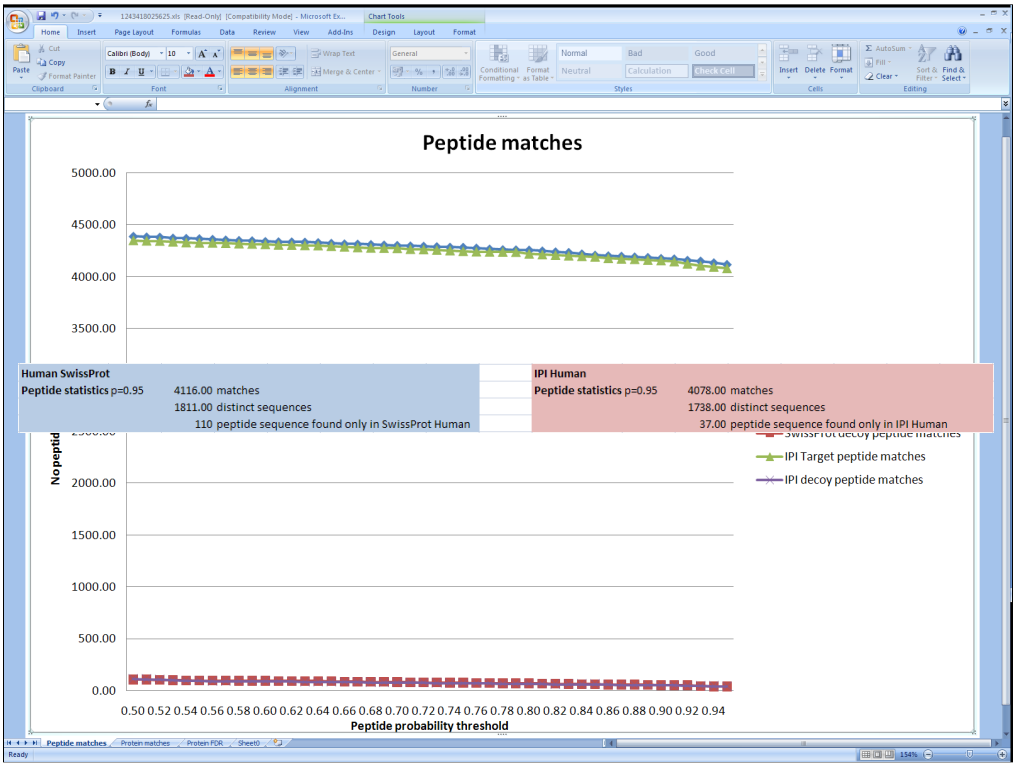
Please note that this is not a ROC plot.

This graph shows the numbers of target and decoy protein hits against each search database. The number of protein hits for both search conditions (for both target and decoy) show a rapid fall in the number of protein hits at a probability of around 0.84, while an increase in probability threshold from 0.99 to 1 results in a large decrease in the number of target hits without a significant decrease in the number of decoy hits. Looking at the results at a probability threshold of 0.9, we can see that IPI human is giving us around twice as many protein hits as searching SwissProt.



Please note that this is not a ROC plot.

Which is more clearly shown on this graph which shows the protein hit FDR for each search. As you can see, the protein false discovery rate shows the same pattern from searching either database. However, the protein FDR from searching IPI Human does appear to be consistently lower than from searching human SwissProt . This probably reflects the fact that the decoy segment of the databases don't have the redundancy present in the target portion since the decoy sequences are randomised sequences; hence with the larger database the increase in matches to the random sequences in the decoy section of the database isn't as great as the increase in matches to redundant protein sequences in the target database. This shows one of the difficulties in trying to obtain reliable protein false discovery rates.



Here we have the absolute number of target and decoy peptide matches from both databases. Despite the fact that IPI Human is giving us twice as many protein hits, we're getting pretty much identical numbers of peptide matches from both databases for both target and decoy. In fact the we have very slightly lower numbers from IPI Human. This probably reflects the larger database size tightening up the statistics – for example the Mascot identity thresholds would be higher which could remove some more borderline matches. Therefore, the peptide discovery and false discovery rates are in fact approximately equal for the two databases. Taking a closer look at the peptide dataset, we've pulled out some additional information for each search. Taking the probability threshold at p=0.95, the SwissProt Human search has given us 4116 high confidence target matches, representing 1811 distinct peptide sequences, while IPI Human has 4078 high confidence matches representing 1738 distinct sequences – so the numbers are very similar. The SwissProt dataset has 110 peptide sequences unique to it, while the IPI Human dataset has 37 sequences unique to it, so searching both databases has given us a total of 147 peptide sequences which might otherwise have been missed.

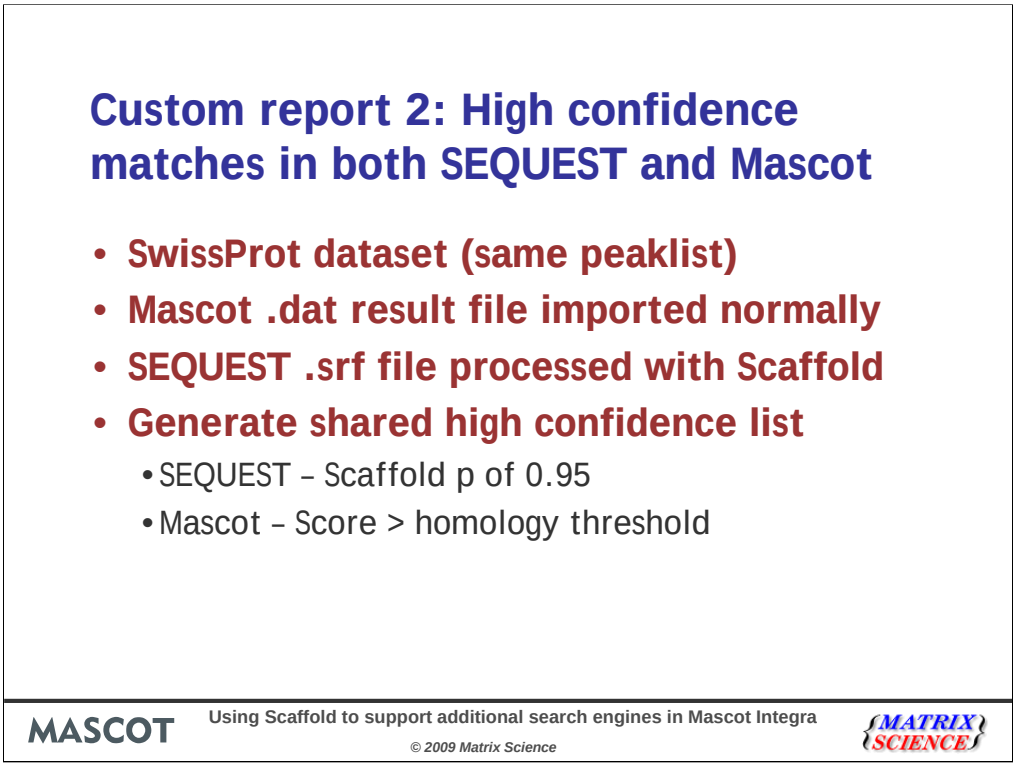
Database comparison summary

- **No significant change in the number of peptides identified**
- **More protein hits identified in IPI Human**
  - Increased redundancy
  - Protein threshold p=0.9

    Swissprot: 357 target protein hit ranks
    378 protein hits

    IPI Human: 349 target protein hit ranks
    718 protein hits

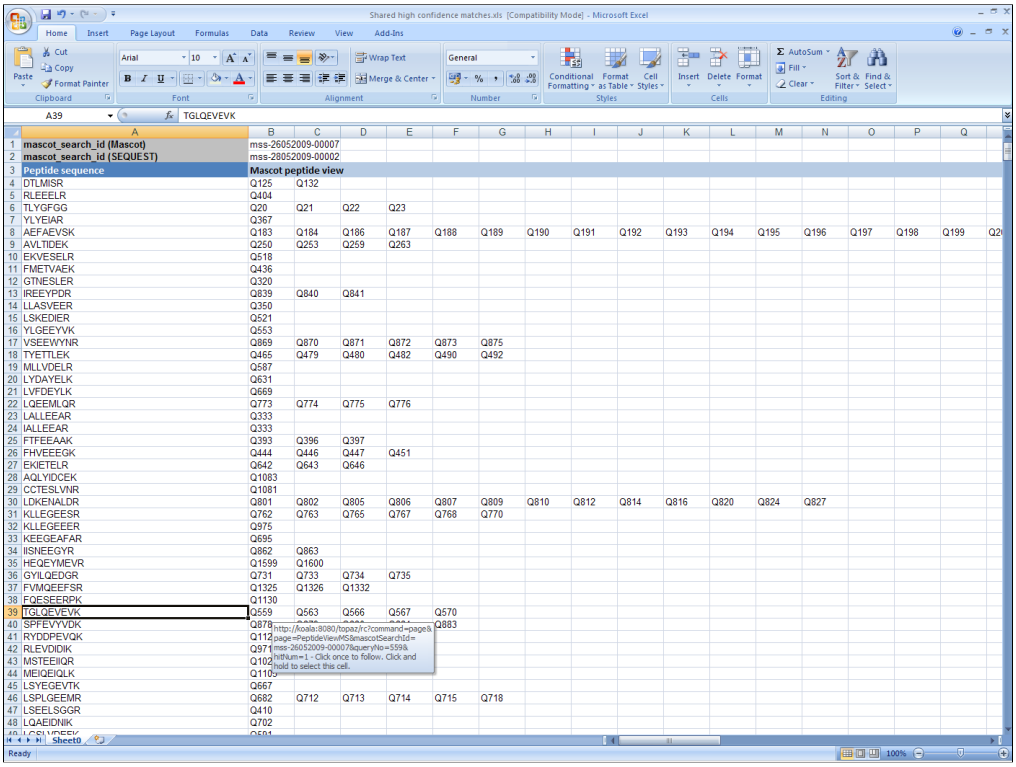
So to summarise; The peptide set identified by searching both databases was very similar, despite the much larger size of the IPI Human database. Searching IPI Human did increase the total possible protein hit list though. This was largely due to an increase in redundancy of identified protein hits (such as the increase in the number of Septins matched). Depending on what you are looking for, searching SwissProt may well have been sufficient for this dataset.



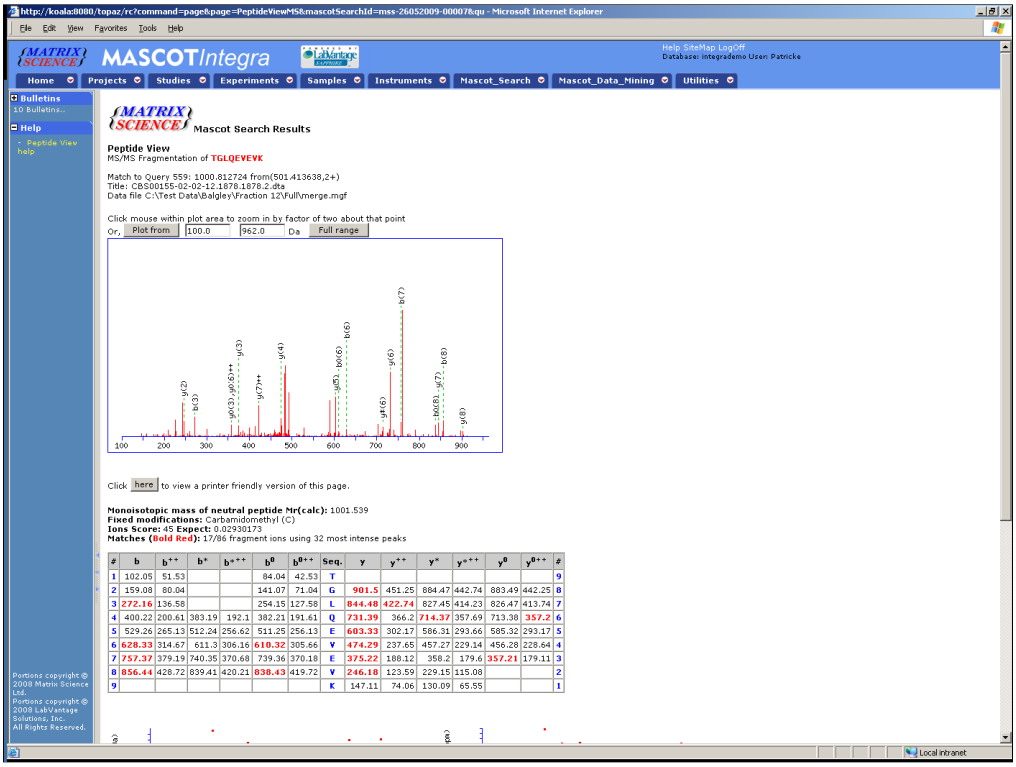
Something else you would want to be able to do easily is to compare results between search engines and runs from within the Integra database, allowing for large scale comparisons.

Using the same search results (just the SwissProt results), I've imported the Mascot result files into Integra (so they're present separately as standard Mascot results), and processed and imported the SEQUEST results with Scaffold separately.

I've then put together a report which pulls out the peptide matches which have a high confidence in both search engines (p 0.95 for SEQUEST/Scaffold and a score above the homology threshold for Mascot results).



Here we have our list of shared high confidence peptide matches from SEQUEST and Mascot. Because we've looked at the source Mascot results and compared those with the SEQUEST/Scaffold protXML export we have access to the full peptide information from the Mascot search. Next to each peptide sequence I've incorporated a reference to the source Mascot query match – which is actually a link to the peptide view from Integra. Click on the cell to follow the link (here we're going to look at query 559) and the peptide view will open in a browser window



#	b	b^{++}	b^{*}	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y^{*}	y^{*++}	y^{0}	y^{0++}	#
1	102.05	51.53			84.04	42.53	T							9
2	159.08	80.04			141.07	71.04	G	**901.5**	451.25	884.47	442.74	883.49	442.25	8
3	**272.16**	136.58			254.15	127.58	L	**844.48**	**422.74**	827.45	414.23	826.47	413.74	7
4	400.22	200.61	383.19	192.1	382.21	191.61	Q	**731.39**	366.2	**714.37**	357.69	713.38	**357.2**	6
5	529.26	265.13	512.24	256.62	511.25	256.13	E	**603.33**	302.17	586.31	293.66	585.32	293.17	5
6	**628.33**	314.67	611.3	306.16	**610.32**	305.66	V	**474.29**	237.65	457.27	229.14	456.28	228.64	4
7	**757.37**	379.19	740.35	370.68	739.36	370.18	E	**375.22**	188.12	358.2	179.6	**357.21**	179.11	3
8	**856.44**	428.72	839.41	420.21	**838.43**	419.72	V	**246.18**	123.59	229.15	115.08			2
9							K	147.11	74.06	130.09	65.55			1

And here we have our match from Mascot.

Acknowledgements

Brian Searle (Proteome Software inc)
LabVantage Solutions Inc