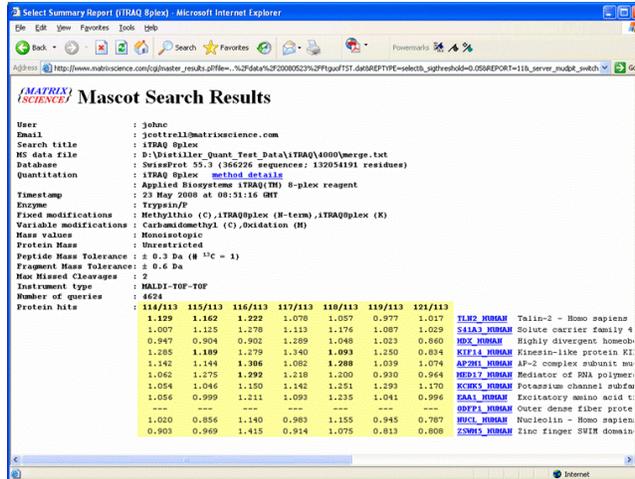


Quantitation FAQ

MASCOT

MATRIX
SCIENCE

1. What software do I need for iTRAQ or TMT?



MASCOT : Quantitation FAQ

© 2009 Matrix Science



In Mascot, the quantitation approach used for iTRAQ and TMT is called the Reporter protocol. All of the information required to perform quantitation is present in a standard, MS/MS peak list, so everything can be handled by the Mascot search engine.

Here, for example, is a Mascot search report where the tag is the Applied Biosystems 8plex iTRAQ. The data are from a 4000 series TOF-TOF.

1. What software do I need for iTRAQ or TMT?

Select Summary Report (John Rogers TMT PQD extract_msn.exe) - Microsoft Internet Explorer

Mascot Search Results

User : John Cottrell
Email : jcottrell@matrixscience.com
Search title : John Rogers TMT PQD extract_msn.exe
MS data file : \\skippy\dfs\mascot_test_data\quantitation\TMT\John Rogers - Thermo\Proteinmix2TMT_011010.ms
Database : SwissProt 55.1 (359942 sequences; 129199355 residues)
Quantitation : TMT 6plex [method details](#)
 : Proteome Sciences sixplex Tandem Mass Tag
Timestamp : 4 Apr 2008 at 12:45:45 GMT
 : Trypsin/P
Enzyme : Carbamidomethyl (C), TMT6plex (N-term), TMT6plex (K)
Mass values : Monoisotopic
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 2.5 Da
Fragment Mass Tolerance : ± 0.20 Da
Max Missed Cleavages : 2
Instrument type : ESI-TOF
Number of queries : 8131
Protein hits :

| | 127/126 | 128/126 | 129/126 | 130/126 | 131/126 | |
|-------|---------|---------|---------|---------|--|--|
| 1.800 | 2.976 | 4.305 | 1.958 | 0.679 | OVAL_CHICK Ovalbumin - Gallus gallus (Chicken) | |
| 1.877 | 2.754 | 3.884 | 1.898 | 0.512 | ALBU_BOVIN Serum albumin precursor - Bos taur | |
| 2.027 | 3.339 | 4.409 | 2.621 | 0.793 | TRYI_BOVIN Cationic trypsin precursor - Bos ta | |
| 2.227 | 3.030 | 4.338 | 2.191 | 0.619 | LYSC_CHICK Lysozyme C precursor - Gallus gallu | |
| 2.082 | 2.963 | 4.078 | 2.032 | 0.735 | MYG_BOVIN Myoglobin - Equus caballus (Horse) | |
| 2.213 | 3.186 | 4.211 | 2.276 | 0.779 | CARB_BOVIN Carbonic anhydrase 2 - Bos taurus (| |
| 1.708 | 2.693 | 3.259 | 2.481 | 0.626 | G3P_FIG Glycerinaldehyde-3-phosphate dehydrog | |

Select Summary Report

6plex TMT data from a collaborative demonstration by Proteome Sciences and Thermo Scientific, (PQD on LTQ XL).

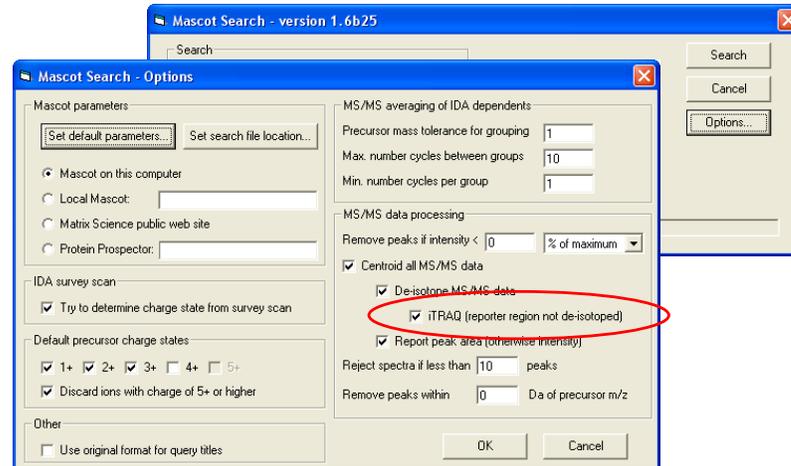
MASCOT : Quantitation FAQ

© 2009 Matrix Science



While this shows the 6plex TMT tag, courtesy of Proteome Sciences and Thermo Scientific. These data are from an LTQ XL using Pulsed Q Dissociation

1. What software do I need for iTRAQ or TMT?



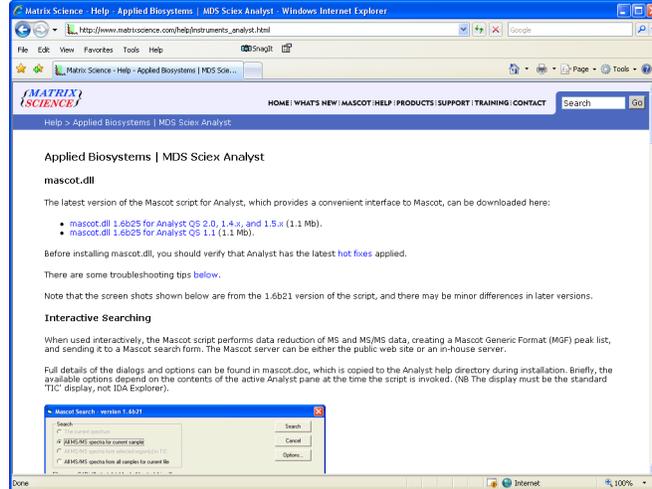
MASCOT : Quantitation FAQ

© 2009 Matrix Science



You can use Distiller to create the peak list, but this isn't essential. Any peak picking utility can be used provided it doesn't try to de-isotope the reporter ion peaks. This is essential because reporter ion peaks are not peptide ions; they are synthetic molecules with multiple heavy isotopes. If you try to de-isotope them using the simple rules, this will distort the true intensities. Recent versions of mascot.dll for Sciex Analyst are "reporter ion aware". If you are using Analyst and you don't see a checkbox for iTRAQ then you should update to the latest mascot.dll.

1. What software do I need for iTRAQ or TMT?



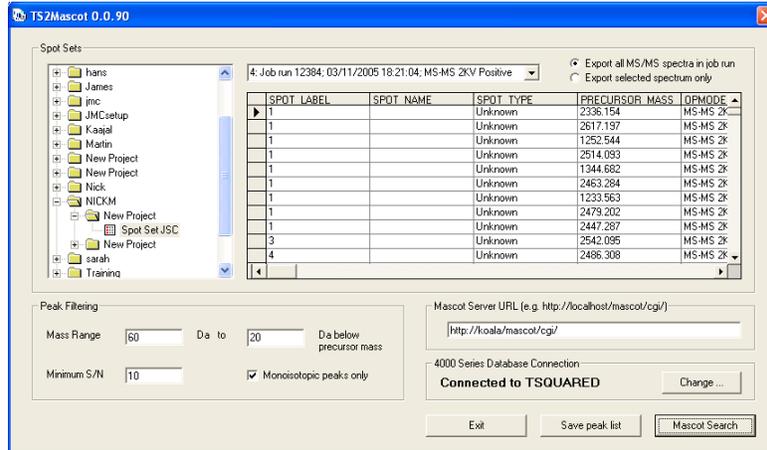
MASCOT : Quantitation FAQ

© 2009 Matrix Science



This can be downloaded from our web site. If you use mascot.dll as an import filter in Mascot Daemon, you won't see the checkbox in the options dialog in Daemon. This is because it was added after Daemon 2.2 was released. This is fixed in Mascot 2.3. For Mascot 2.2 and earlier, controls that are missing from the dialog take the defaults set when mascot.dll called as an Analyst script, so the workaround is to process a file using the correct settings in Analyst before processing files from Daemon.

1. What software do I need for iTRAQ or TMT?



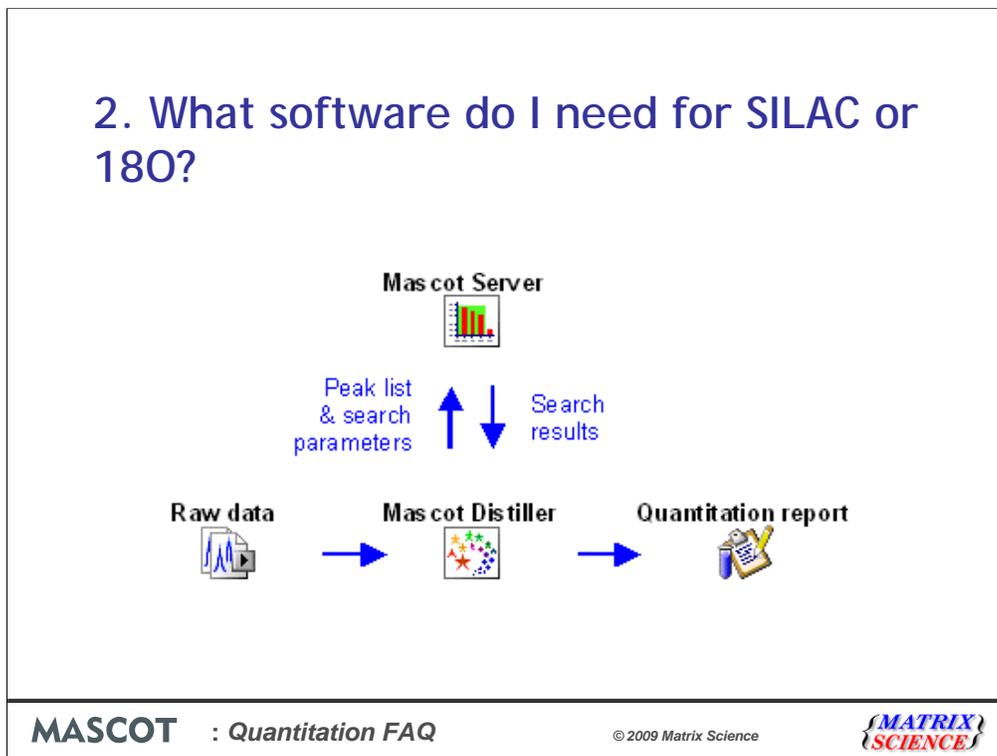
MASCOT : Quantitation FAQ

© 2009 Matrix Science



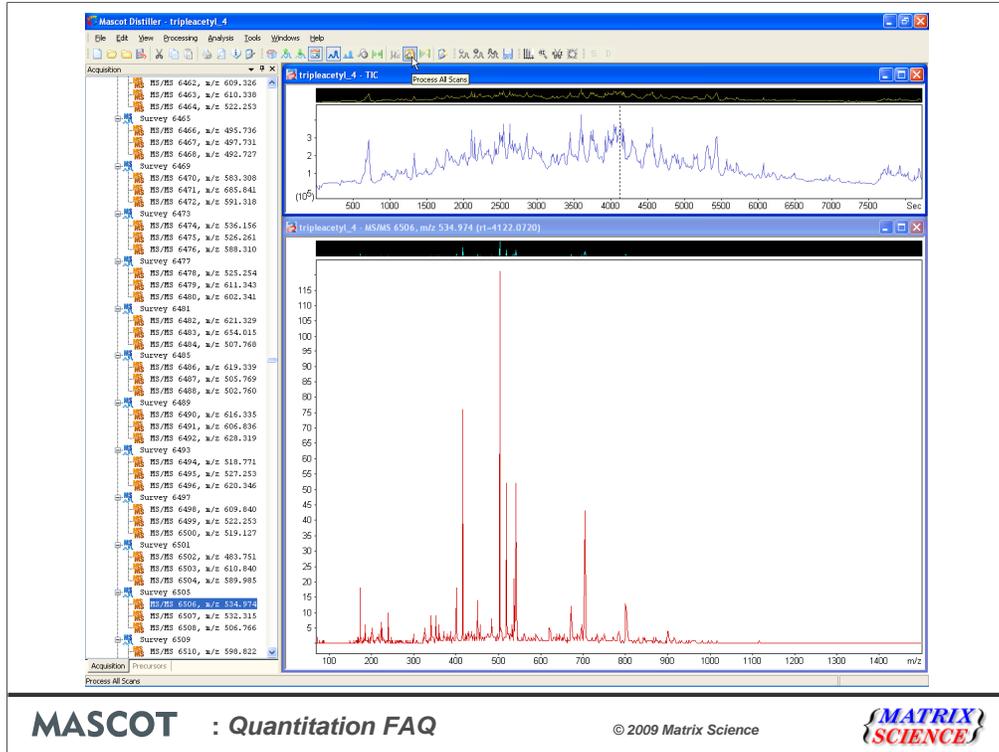
In the case of the Applied Biosystems TOF-TOF, the peak list submitted to Mascot by GPS Explorer or 4000 Series Explorer contains de-isotoped reporter ions and should not be used for iTRAQ quantitation. There is a freeware application called TS2Mascot that can be downloaded from our web site and used to submit a search with the correct peak areas.

2. What software do I need for SILAC or 18O?



Quantitation based on the relative intensities of extracted ion chromatograms (XICs) for precursors is known in Mascot as Precursor protocol and by far the most widely used approach, applicable to any chemistry that creates a precursor mass shift. For example, 18O, AQUA, ICAT, ICPL, Metabolic, SILAC, etc., etc.

For this approach, the peak list alone is not sufficient; additional information is needed from the raw data file. If you go the Mascot route, the workflow looks like this. The raw data file is processed in Distiller and the search submitted to Mascot. When the search is complete, the results are returned to Distiller. The quantitation report can then be generated in Mascot Distiller, which has access to both the Mascot search results and the raw data

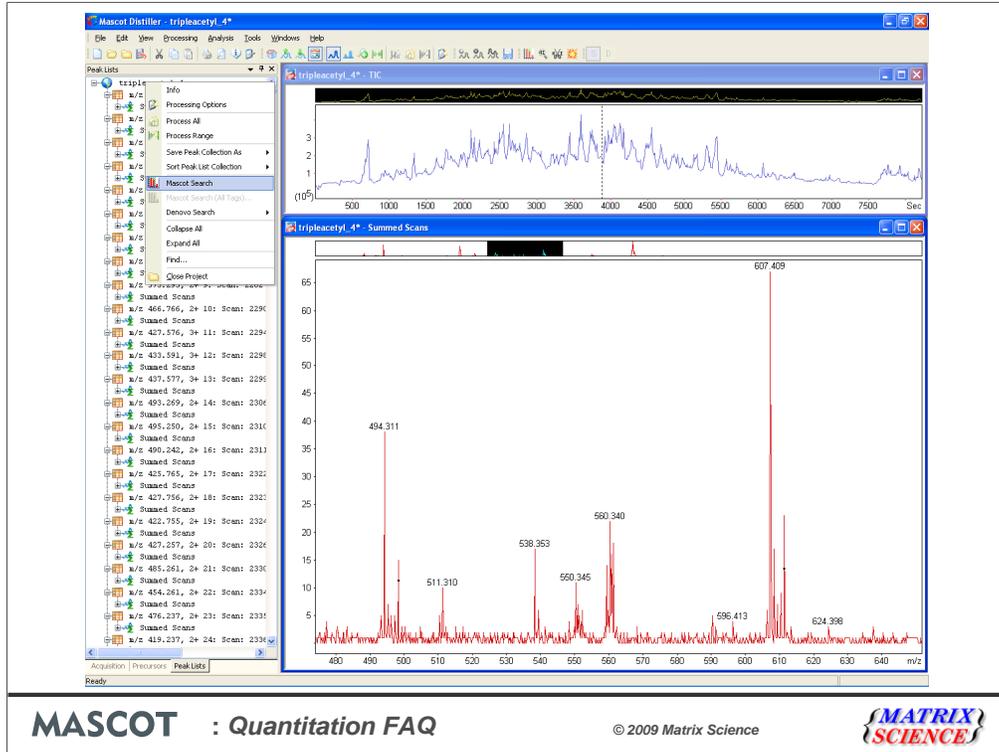


MASCOT : Quantitation FAQ

© 2009 Matrix Science



The next few slides illustrate this workflow using one of the QStar sample data sets that can be downloaded from the MSQuant home page. This is a SILAC sample with three components: unlabelled, Arg labelled with $^{13}\text{C}(6)$, and Arg labelled with $^{13}\text{C}(6)^{15}\text{N}(4)$. First, we open the Wiff file in Mascot Distiller, and process it into peak lists



The peak lists are submitted to a Mascot Server to be searched

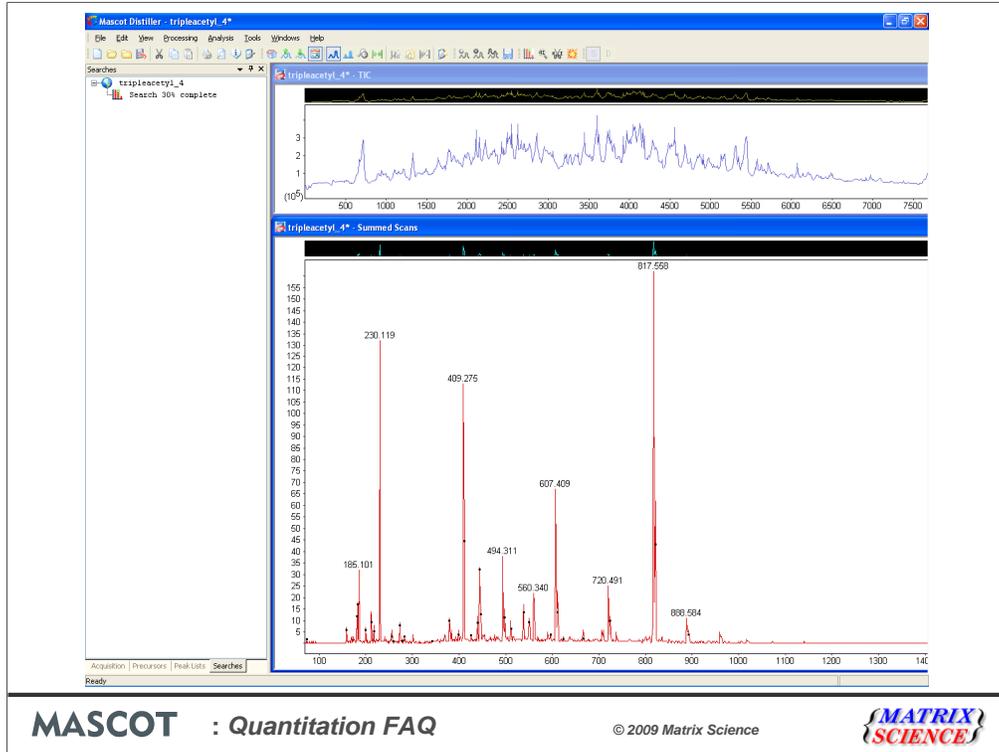
The screenshot displays the Mascot Distiller interface with a MASCOT MS/MS Ions Search dialog box open. The dialog box contains the following settings:

- Your name:** John Cottrell
- Email:** jcottrell@matrixscience.com
- Search title:** tripeacetyl_4
- Database:** NCBIInr
- Taxonomy:** Homo sapiens (human)
- Enzyme:** Trypsin/P
- Allow up to:** 2 missed cleavages
- Fixed modifications:** Carbamidomethyl (C), Carbamyl (K), Carbamyl (N-term), Carboxymethyl (C), Cation:Na (C-term)
- Variable modifications:** Methyl (DE), Methylthio (C), NIPCAM (C), Oxidation (HW), Oxidation (M)
- Quantitation:** SILAC R+0 R+10 (MD)
- Peptide tol.:** 0.200 Da
- MS/MS tol.:** 0.200 Da
- Peptide charge:** 2+ and 3+
- Monoisotopic:** checked
- Average:** unchecked
- Data file:** LOCALS-1\Temp\Dis166.tmp
- Data format:** Mascot generic
- Instrument:** MALDI-QUAD-TOF
- Decay:** unchecked
- Report top:** AUTO hits

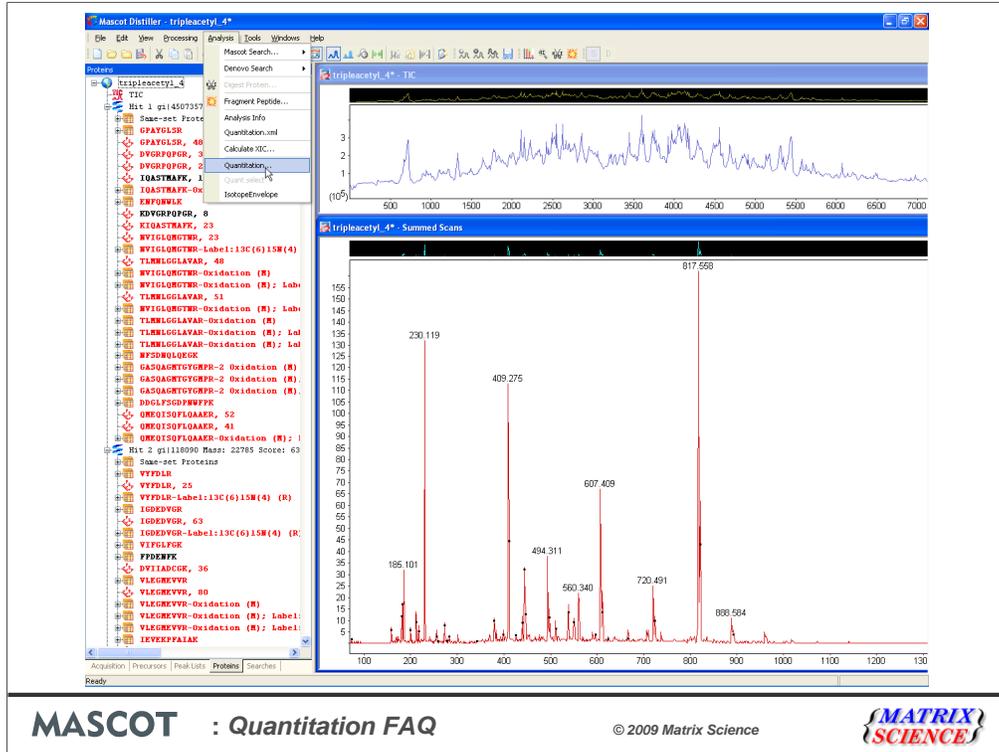
The background shows a peak list on the left and a mass spectrum plot at the bottom with peaks labeled at m/z 560.340, 720.491, and 888.584.

MASCOT : Quantitation FAQ © 2009 Matrix Science **MATRIX SCIENCE**

There's an opportunity to tweak the search conditions, but most of the important settings, including the modifications, are embedded in the selected quantitation method



We get progress reports while the search is running



MASCOT : Quantitation FAQ

© 2009 Matrix Science



When the search is complete, the results are imported into Distiller. All the required information is now available, so Distiller is in a position to generate a quantitation report

The screenshot displays the Mascot Distiller interface. On the left, a list of proteins is shown with their accession numbers, scores, and modification states. The main window shows a 'Quantitation Report' table with columns for Accession, Score, Mass, ML, S(geos), #, HL, S(D(geos)), #, and Description. Below the table, two mass spectrometry plots are visible: 'triplicate4_4 - NYVGLQMGTR(light)' and 'triplicate4_4 - Survey 5101 (I1-2489-4460)'. The first plot shows relative intensity versus m/z with a shaded region indicating a specific peak. The second plot shows a survey scan with multiple peaks across the m/z range from 608 to 619.

| Accession | Score | Mass | ML | S(geos) | # | HL | S(D(geos)) | # | Description |
|------------|-------|-------|--------|---------|---|--------|------------|---|--|
| g14507357 | 743 | 22945 | 1.0093 | 1.0626 | 7 | 0.9074 | 1.0569 | 7 | Transglutn 2 [Homo sapiens] |
| g1110390 | 638 | 22785 | 1.0184 | 1.0575 | 6 | 0.8991 | 1.0476 | 6 | Protein-tyrosine phosphatase B precursor |
| g14505591 | 515 | 22324 | 0.9887 | 1.0862 | 3 | 0.8906 | 1.1686 | 3 | peroxiredoxin 1 [Homo sapiens] |
| g14505403 | 444 | 23560 | 0.9779 | 1.0550 | 3 | 0.8573 | 1.0729 | 3 | Chain A, Crystal Structure Of Human Glutathione S-transferase [Homo sapiens] |
| g145054207 | 400 | 23665 | 0.9779 | 1.0550 | 3 | 0.8573 | 1.0729 | 3 | Chain A, Thioesterin Peroxidase B From Recombinant [Homo sapiens] |
| g145055005 | 296 | 23653 | 1.2162 | 1.0334 | 3 | 1.2760 | 1.0708 | 3 | Transglutn [Homo sapiens] |
| g131674085 | 263 | 28862 | 0.9503 | 1.2371 | 5 | 0.8562 | 1.1766 | 5 | hypothetical protein [Homo sapiens] |
| nc011169 | 252 | 24107 | 0.9476 | 1.1768 | 5 | 0.8198 | 1.1228 | 5 | neuronal-specific 13 [Human brain, Predicted] |

MASCOT : Quantitation FAQ

© 2009 Matrix Science



The quantitation results are displayed in a grid control, spreadsheet-style.

2. What software do I need for SILAC or 18O?

Mascot Distiller 2.2 or later

- Workstation + Search Toolbox + Quantitation Toolbox

Mascot Server 2.2 or later

So, if you go the Mascot route, this is what you need.

Please note that, if you use Distiller to peak pick data for reporter protocol quantitation, i.e. iTRAQ or TMT, and submit the Mascot search, you still have to go to the standard Mascot HTML report to see the quantitation results. The aim is for Distiller to import the quantitation results for reporter and multiplex, so that everything can be viewed and reported in Distiller, but that is scheduled for a future release.

2. What software do I need for SILAC or 180?

"Toolkit" modules excluded. Packages in italics are primarily or exclusively for label-free

Commercial

BioWorks (Thermo)
DeCyder MS 2.0 (GE Healthcare)
Elucidator (Rosetta)
Excel (Microsoft)
GPS Explorer (Applied Biosystems)
Mascot Distiller (Matrix Science)
MsXelerator (MsMetrix)
PEAKS Studio (BSI)
PLGS (Waters)
Progenesis LCMS (NonLinear)
ProQuant (Applied Biosystems)
ProteinPilot (Applied Biosystems)
SIEVE (Thermo)
Spectrum Mill (Agilent)
WARP-LC (Bruker)

Shareware / Academic

ASAPRatio (TPP)
AYUMS (U. Tokyo)

Census (Scripps)
iTRACKER (Cranfield) ITRAQAnalyzer (Topp / OpenMS)
Libra (TPP)
MaxQuant (MPI Martinsried)
metAlign (Wageningen UR)
MFPaQ (IPBS Toulouse)
msInspect/AMT (CPL FHCRC)
MSQuant (SDU Odense)
MS-Spectre (Utrecht)
Multi-Q (Academia Sinica)
MZmine (VTT Finland)
Peakardt (Bochum)
ProRata (ORNL)
Quant (Wuerzburg)
QUIL (NIH)
RAAMS (Mayo)
RelEx (Scripps)
SILACAnalyzer (Topp / OpenMS)
VEMS (CIC bioGUNE)
VIPER (PNNL)
Vista (Harvard MS)
XCMS (Scripps)
Xpress (TPP)
ZoomQuant (Wisconsin)

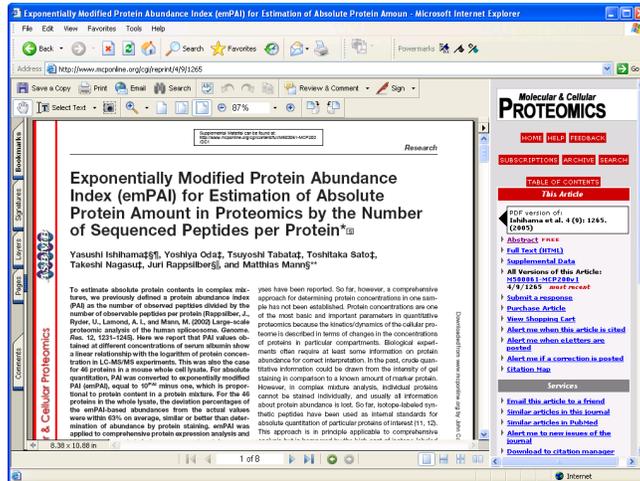
MASCOT : Quantitation FAQ

© 2009 Matrix Science



Of course, there are a host of other packages to generate a quantitation report based on Mascot search results. I know this list is incomplete; I just want to give you a sense of the choice that is available. Including Excel is only half joking. Seems like this is still the software of choice for many people when they want to do something novel, like use one manufacturer's instrument with another manufacturer's tag.

3. Does Mascot support spectral counting?



MASCOT : Quantitation FAQ

© 2009 Matrix Science



Yes. emPAI quantitation offers approximate, label-free, relative quantitation of the proteins in a mixture based on protein coverage by the peptide matches in a database search result. This approach was developed by Ishihama and colleagues

3. Does Mascot support spectral counting?

Very simple

$$emPAI = 10^{\frac{N_{observed}}{N_{observable}}} - 1$$

Very approximate

- Many assumptions in $N_{observed}$ and $N_{observable}$

'Always on'

1. [PPB1_HUMAN](#) Mass: 58259 Score: 452 Queries matched: 17 emPAI: 1.04
Alkaline phosphatase, placental type precursor (EC 3.1.3.1) (PLAP-1) (Regan isozyme) - Homo sapiens
 Check to include this hit in error tolerant search

| Query | Observed | Mr(expt) | Mr(calc) | Delta | Miss | Score | Expect | Rank | Peptide |
|--|----------|-----------|-----------|---------|------|-------|---------|------|-----------------|
| <input checked="" type="checkbox"/> 27 | 462.6807 | 923.3468 | 923.5116 | -0.1649 | 0 | 33 | 0.25 | 1 | R.FPYVALSK.T |
| <input checked="" type="checkbox"/> 41 | 517.1760 | 1032.3375 | 1032.5604 | -0.2229 | 0 | 71 | 6.4e-05 | 1 | R.GSSIFGLAPGK.A |
| <input checked="" type="checkbox"/> 62 | 564.6804 | 1127.3463 | 1127.5764 | -0.2301 | 0 | 9 | 1.2e+02 | 1 | R.GFFLVEGGK.I |

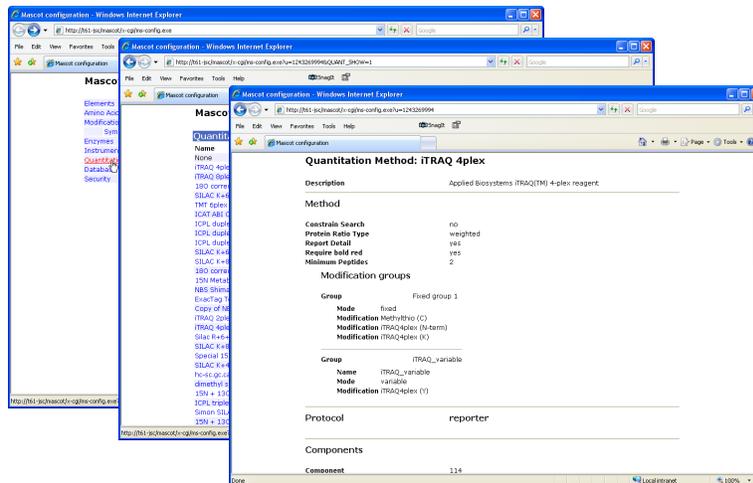
MASCOT : Quantitation FAQ

© 2009 Matrix Science



emPAI is very simple. It is also very approximate, because there are many arbitrary assumptions in the way that the number of observed and observable peptides are calculated. Nevertheless, Ishihama's paper shows that it can be a useful guide to relative amounts. emPAI doesn't require a label or special data processing, so it is always reported in a standard Mascot results report, as long as the number of MS/MS spectra is at least 100

4. Why do I keep getting errors about conflicting modifications?



MASCOT : Quantitation FAQ

© 2009 Matrix Science



To understand this, take a look at the quantitation method. You can get a pretty printed summary from the Configuration Editor. Choose Quantitation then Print for the method of interest. This is the original iTRAQ method, that shipped with Mascot 2.2. The iTRAQ modifications are defined in the quantitation method. In this particular case, this isn't strictly necessary, but for many quantitation methods the modifications have to be defined within the method because the modifications define the components. For example, a peptide with 16O at the C-terminus is light while a peptide with 18O at the C-terminus is heavy. So, the approach we took was that the isotopic labels should be defined within the method. We also felt that specifying as much as possible in the method was a good thing in terms of standard operating procedures. The original iTRAQ method was based on the Applied Biosystems procedure that used methythio (MMTS) to block the cysteines, so we also included this modification. The problem comes when someone uses a different alkylation agent, such as iodoacetamide, and chooses carbamidomethyl in the search form.

4. Why do I keep getting errors about conflicting modifications?

The screenshot shows a web browser window displaying a Mascot search results page. The page title is "Peptide Summary Report (Time files like an arrow) - Windows Internet Explorer". The browser address bar shows the URL: http://www.matrixscience.com/cgi/master_results.pl?raw=...

The main content is a table of search results. The table has columns for peptide number, mass, charge, pI, MW, and sequence. The sequences include modifications like iTRAQ4plex (K) and iTRAQ4plex (R-term).

| Peptide | Mass | Charge | pI | MW | Sequence | | | | |
|---------|-----------|-----------|-----------|---------|----------|---|---------|---|--|
| 229 | 1096.9054 | 2287.6944 | 3286.8766 | 0.0170 | 2 | 0 | 2e+02 | 1 | HADGARDLGMALSTCLLPEPLGAARTK |
| 119 | 691.3824 | 1380.7602 | 1380.7647 | -0.0344 | 0 | 0 | 3.4e+02 | 1 | PFPAGDLSLAGAR |
| 233 | 1511.6445 | 3024.0912 | 3982.0250 | -0.0531 | 2 | 7 | 1.8e+02 | 1 | PTDGGSGFYLTPSNITVVGKDTFAYSIYR |
| 88 | 606.3040 | 1210.5934 | 1210.6293 | -0.0359 | 1 | 6 | 4.8e+02 | 1 | HSKAKSHTP |
| 51 | 562.2309 | 1122.4472 | 1122.6267 | -0.1795 | 0 | 6 | 2.9e+02 | 1 | ISELPHGAR |
| 108 | 877.7384 | 2630.1934 | 2630.3596 | -0.1662 | 1 | 5 | 1.9e+02 | 1 | ATDRLHGGKRWLICEVDVGR |
| 224 | 1090.1404 | 2187.4594 | 3306.3097 | 1.0096 | 1 | 4 | 1.2e+02 | 1 | DRITTTGHTMRLSTVDFGAMVNGAR |
| 202 | 894.1127 | 2679.3163 | 2679.3496 | -0.0334 | 0 | 4 | 5.3e+02 | 1 | QSASGYSERFALLDHPYTTGR + iTRAQ4plex (X) |
| 88 | 605.3054 | 1208.5962 | 1208.6887 | -0.0924 | 0 | 4 | 8.3e+02 | 1 | VVLETTLSR |
| 197 | 877.3937 | 2629.1593 | 2629.2587 | -0.0994 | 0 | 2 | 2.6e+02 | 1 | FSPGGAALVSHFPGQARDRCR + iTRAQ4plex (X) |
| 104 | 833.3796 | 2497.1170 | 2496.3104 | 0.0066 | 1 | 2 | 3.5e+02 | 1 | YTPSYVAFHREKLGDPAR |
| 220 | 1261.0090 | 2762.6452 | 3191.6942 | 0.9509 | 0 | 0 | 1e+02 | 1 | FSGGSGTFLIDHMRDSSAMTFCQTK |
| 223 | 1029.1577 | 3084.4513 | 3083.3974 | 1.0539 | 0 | 0 | 7.9e+02 | 1 | HSTYSALSSYYVDTIHGKADPVTGAR |
| 106 | 835.0141 | 2502.0205 | | | | | | | |
| 222 | 1028.0370 | 3083.4316 | | | | | | | |

Search Parameters

- Type of search : MS/MS Ion Search
- Enzyme : Trypsin/P
- Fixed modifications : Carbamidomethyl (C), iTRAQ4plex (K), iTRAQ4plex (R-term), iTRAQ4plex (R-term), iTRAQ4plex (K)
- Mass values : Monoisotopic
- Protein Mass : Unrestricted
- Peptide Mass Tolerance : ± 0.2 Da (# 13C = 1)
- Fragment Mass Tolerance : 0.2 Da
- Max Missed Cleavages : 2
- Instrument type : ESI-QUAD-TOF
- Number of queries : 234

Mascot: <http://www.matrixscience.com/>

MASCOT : Quantitation FAQ

© 2009 Matrix Science



Because it became clear that few people use MMTS, we took this out of the method when we updated it. If you haven't done so already, you should grab an updated quantitation.xml file from our web site, using the link on the Distiller support page. This fixes the MMTS problem, but it doesn't stop people using a method without looking at it. I suspect that many people have search results that look like this. The iTRAQ mods are there twice, because they have been chosen in the search form even though they are already in the method. This isn't an error, because there is no conflict, but it does indicate that we haven't got the message across about embedding the important search parameters in the method. Hopefully, this FAQ will help clarify that the modifications that are essential to the method should go in the method while those that are 'optional' and will vary from experiment to experiment, such as oxidation or acetylation, should be selected in the search form.

5. Can quantitation be automated?

Mascot Daemon & Distiller 2.2

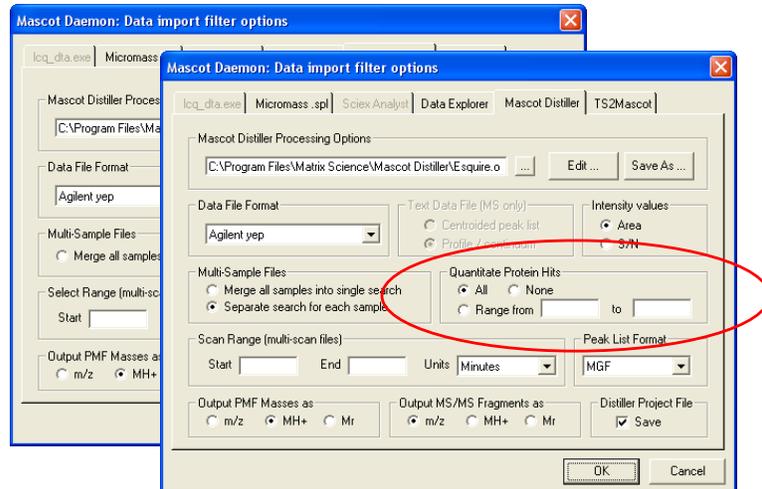
- Batch automation of peak picking, search submission, search result import, saving of Distiller Project file
- You still have to open the Distiller project file and click on the 'quantitate' tool button

Mascot Daemon & Distiller 2.3

- All steps are automated

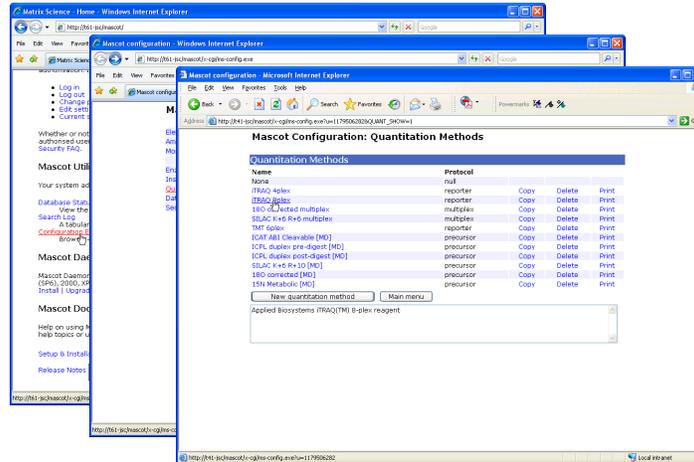
Yes, but not until Mascot 2.3 released.

5. Can quantitation be automated?



In Mascot Daemon 2.3, you have a choice whether to quantitate any, some or all of the protein hits

6. How do I create or edit a quantitation method?



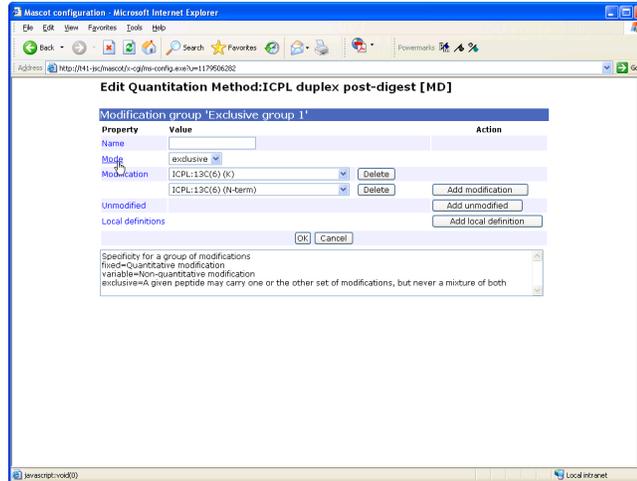
MASCOT : Quantitation FAQ

© 2009 Matrix Science



Quantitation methods are created or modified using the Configuration Editor. This was mentioned a few slides back, in the context of getting a pretty printed summary of a method. Choose Configuration Editor from your local Mascot home page, then Quantitation. For each method, besides Print, there is Delete, which is self-explanatory, and Copy, which creates a copy of an existing method. Starting from this is usually less work than choosing New and starting from scratch.

6. How do I create or edit a quantitation method?



MASCOT : Quantitation FAQ

© 2009 Matrix Science



To ensure flexibility, there are very many quantitation parameters. These are documented in the help for both Mascot Server and Mascot Distiller. Also, when you are editing a method, help is displayed for individual fields whenever the mouse is held over a label, as shown

7. Why do I get no results for certain peptide matches?

| Query | Observed | Mr(calc) | Mr(calc) | Delta Miss Score | Expect | Rank | 115/114 | 116/114 | 117/114 | Peptide | | |
|-------|-----------|-----------|-----------|------------------|--------|------|---------|---------|---------|---------|--------|---------------------|
| 53 | 534.7767 | 1067.5288 | 1067.5491 | -0.0103 | 0 | 38 | 0.2 | 1 | 1.962 | 15.329 | 13.409 | K.FQGFQK.R |
| 26 | 625.0919 | 1249.7659 | 1249.7638 | 0.0033 | 0 | 36 | 0.35 | 1 | 2.148 | 2.180 | 3.430 | K.LTQDFQK.Y |
| 113 | 661.0789 | 1321.7432 | 1321.7597 | -0.0165 | 0 | 34 | 0.52 | 1 | 1.929 | 1.519 | 2.946 | K.YHQQLQK.W |
| 110 | 670.3065 | 1338.7584 | 1338.7387 | 0.0198 | 0 | (19) | 17 | 1 | --- | --- | --- | K.LDDVYVAK.A |
| 119 | 670.3067 | 1338.7589 | 1338.7307 | 0.0282 | 0 | 57 | 0.0023 | 1 | 1.912 | 1.956 | 3.475 | K.LDDVYVAK.A |
| 124 | 705.4144 | 1408.8142 | 1408.7805 | 0.0337 | 0 | 68 | 0.00018 | 1 | 1.027 | 5.873 | 10.012 | R.GDFYVDAINTSFAMGT |
| 141 | 723.4282 | 1444.8418 | 1444.8282 | 0.0137 | 0 | 35 | 0.4 | 1 | 1.174 | 0.851 | 1.783 | R.RFDFYVANSK.R.Q |
| 147 | 751.0827 | 1501.7508 | 1501.8092 | -0.0584 | 0 | 18 | 17 | 1 | --- | --- | --- | K.MAQFQQQLK.S |
| 148 | 757.9334 | 1513.8522 | 1513.8344 | 0.0179 | 0 | 42 | 0.077 | 1 | 3.981 | 4.020 | 6.726 | R.LEHQFLEQAK.V |
| 166 | 1119.1100 | 2236.2054 | 2236.2558 | -0.0504 | 0 | 61 | 0.00089 | 1 | 1.773 | 7.411 | 11.119 | K.AASDLIIDLQSLTWHF |
| 172 | 789.3096 | 2365.1470 | 2365.1495 | -0.0025 | 0 | (29) | 1.1 | 1 | 0.995 | 1.113 | 1.314 | R.GDAFVDAINTSFAMGT |
| 174 | 789.3903 | 2365.1491 | 2365.1495 | -0.0004 | 0 | (3) | 3.8e+02 | 1 | --- | --- | --- | R.GDAFVDAINTSFAMGT |
| 175 | 789.7163 | 2366.1271 | 2365.1495 | 0.9776 | 0 | (10) | 81 | 1 | --- | --- | --- | R.GDAFVDAINTSFAMGT |
| 222 | 1028.8178 | 3083.4316 | 3083.4572 | -0.0257 | 0 | (17) | 7.8 | 1 | --- | --- | --- | R.CDQDFYVDFPSTQVPGS |
| 223 | 1029.1377 | 3084.4513 | 3083.4572 | 0.9940 | 0 | 23 | 1.5 | 1 | --- | --- | --- | R.CDQDFYVDFPSTQVPGS |
| 224 | 1090.1604 | 3267.4594 | 3267.5411 | -0.0817 | 1 | 9 | 27 | 1 | --- | --- | --- | R.CDQDFYVDFPSTQVPGS |
| 229 | 1294.9831 | 3881.9275 | 3882.0043 | -0.0768 | 1 | (21) | 4 | 1 | 1.261 | 0.026 | 0.037 | K.WLATELVDFPSTQVPGS |
| 230 | 971.4973 | 3881.9681 | 3882.0043 | -0.0442 | 1 | 22 | 3.4 | 1 | --- | --- | --- | K.WLATELVDFPSTQVPGS |

MASCOT : Quantitation FAQ

© 2009 Matrix Science



A Mascot quantitation method includes certain rules and quality thresholds. One of the most important is the Mascot score. If you aren't confident of the peptide match, then you can't be confident of the protein assignment for that match and wouldn't want to include it in any calculations. A cut-off, which is often used in the default methods, is that the score should be above the homology threshold. So, first thing to check is that the ions score for the match is above threshold. Here, this is the reason for most of the missing ratios. It isn't the explanation for the match to CGDVDFYVDFPSTQVPGSENVIEAVAR with a score of 25. This has been excluded by a rule in the method, that any match with iTRAQ on tyrosine should be ignored. This is because the reaction with tyrosine is slower than with lysine and the amino terminus, which means that the ratio is likely to be skewed.

Fraction threshold

The fraction of the peak area in the precursor region accounted for by the components

| Rank | Sequence | Score | Delta Score | Probability | Intensity |
|------|------------------|---------|-------------|-------------|---------------------------|
| 1 | 24-TGDDGTEVVEAK | 1.1791 | 0.10482 | 0.8255 | 167552 |
| 2 | 24-DODFNNAQVK | 1.0747 | 0.20562 | 0.7992 | 511601 |
| 3 | 24-SEVYVWAKDEL | 0.7629 | 0.85811 | 0.6973 | 346249 |
| 4 | 24-AEAGSLVFAQ | 0.9138 | 0.15541 | 0.8468 | 459532 |
| 5 | 24-VALGAAGGQGG | 0.5887 | 0.18480 | 0.8440 | 249922 |
| 6 | 24-VALGAAGGQGG | 1.0291 | 0.02228 | 0.8099 | 205029 |
| 7 | 24-VALGAAGGQGG | 0.7426 | 0.18418 | 0.8326 | 100267 (16) Oxidation (M) |
| 8 | 24-LNPLVLSLVLVYI | -2.7924 | 1.95889 | 0.7081 | 122576 |

Quantitation Report

070207_nESLC2HCT.LB.DX014 - Survey 2432 (11-8790-4741)

Mass Spectrum (m/z vs Intensity)

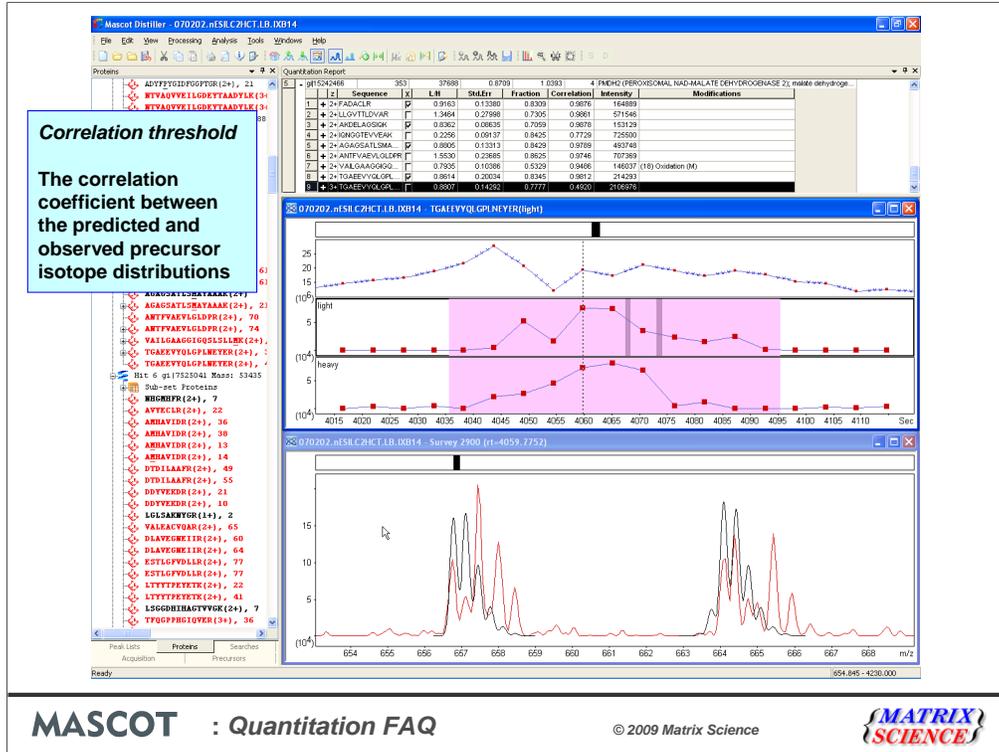
MASCOT : Quantitation FAQ

© 2009 Matrix Science

MATRIX SCIENCE

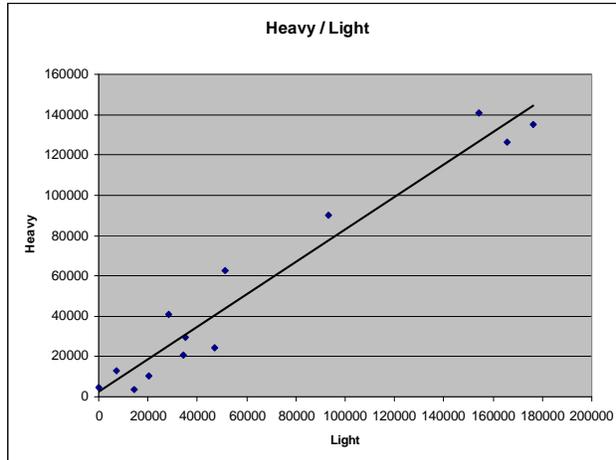
For the methods that use Distiller, there are three quality thresholds that are explained in detail in the help. These are:

Fraction is the fraction of the peak area in the precursor region accounted for by the components. Here is a spectrum where there are a lot of interfering peaks. When we add up the areas, the expected precursors only account for 44% of the area, so the ratio is rejected. This threshold, like the others, is set at the quantitation method. You can change it in Distiller or you can over-ride it at the individual peptide match level by checking or clearing the checkboxes



What about a case where the interfering peaks come right on top of the precursor peaks? A test on the matched fraction won't help. This is why we have a second test on the shape of the distribution. This uses the correlation coefficient between the predicted and observed precursor isotope distributions. Here is a case where a ratio fails the test, because the correlation coefficient is only 0.49

Quantitation - Precursor



Each point represents the heavy and light intensities in a scan from the XIC peak. The straight line is a least squares fit

- The gradient is the best estimate of the ratio.
- Any background is corrected
- The standard error for the fit is a good measure of the reliability of the ratio

Another column reports the estimated standard error for the calculated ratio. Each ratio comes from making a least squares fit to the component intensities from the scans in the XIC peak. Here, for example, each point represents the heavy and light intensities in one scan. The gradient of the fitted line is the best estimate of the ratio. The standard error for the fit is a good measure of the reliability of the ratio, and can simply be tested against another threshold.

7. Why do I get no results for certain peptide matches?

Quantitation Method: SILAC R-6 R-10 [MD]

| | |
|---|-------------------------------------|
| <input type="checkbox"/> Method | |
| Protein search | <input checked="" type="checkbox"/> |
| Protein Ratio Type | median |
| Protein Score | msdpt |
| Report Detail | <input checked="" type="checkbox"/> |
| Show subsets | 0.20 |
| Require Bolded | <input checked="" type="checkbox"/> |
| Minimum Peptides | 2 |
| Significance Threshold | 0.05 |
| <input type="checkbox"/> Protocol | |
| Protocol | precursor |
| <input checked="" type="checkbox"/> Allow mass time match | <input checked="" type="checkbox"/> |
| <input type="checkbox"/> Components | |
| <input type="checkbox"/> Report Ratios | |
| <input type="checkbox"/> Integration | |
| Integration Method | trapezium |
| Single Ratio | <input type="checkbox"/> |
| Integration Source | survey |
| Allow elution time shift | <input type="checkbox"/> |
| Elution Time Delta | 10 seconds |
| Std. Err. Threshold | 0.2 |
| Correlation Threshold | 0.9 |
| XIC Threshold | 0.1 |
| All charge states | <input type="checkbox"/> |
| All charge state threshold | 0.20 |
| Max XIC width | 250 |
| XIC Smooth | 3 |
| <input type="checkbox"/> Quality | |
| <input type="checkbox"/> Outlier Removal | |
| <input type="checkbox"/> Normalisation | |

Allow mass time match
If false, we only report a ratio if we have peptide matches to both components. Also, if false, then all_charge_states treated as false, even if true in method. Almost always set true in practice.

OK Cancel

MASCOT : Quantitation FAQ

© 2009 Matrix Science



If you have less peptide matches than expected, check that ‘Allow mass time match’ is true. If this is false, then you must have matches for all components before a ratio can be reported. If true, you only need a match to one component. Its partner or partners will be identified from their molecular mass and elution time. This is almost always what is wanted.

8. Can I use centroided data for quantitation?

Yes, but ...

- Quantitation will only be as good as the original peak picking

This is mainly an issue for trap data, where raw data is often saved as centroids (i.e. a peak list) to minimise the size of the raw file. This is not a serious problem as long as the peak picking performed by the firmware or instrument data system is accurate. How can you tell? Look critically at the spectra in the instrument data system or in Distiller. Are the isotope envelopes for weak peaks complete, or does it look like weak isotope peaks are missing? Are the isotope peak spacings accurate? Are the peak positions stable from scan to scan? Ideally, make duplicate runs of a dilution standard and save one as profile and the other as centroid. Process them both and compare the results. If the accuracy and precision of one is significantly better than the other, that answers this question for you.

8. Can I use centroided data for quantitation?

Reporter (iTRAQ, TMT)

- Need at least unit resolution in the reporter ion region
- Ensure centroiding does not attempt to de-isotope
- Reporter ion intensity may be the most critical factor

Precursor (SILAC, ^{18}O , metabolic)

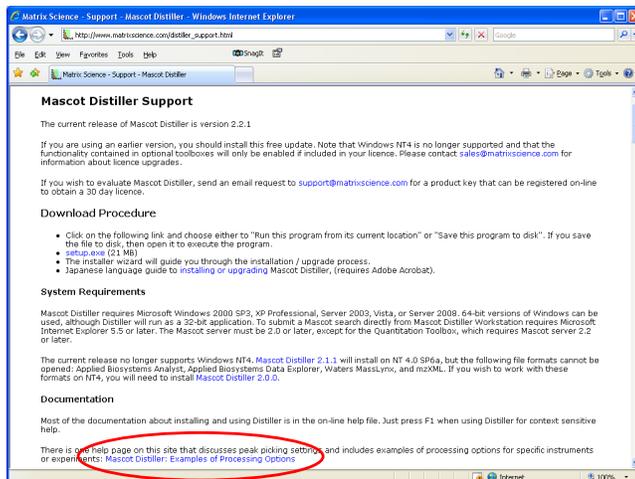
- For standard traps, use zoom scans for best results (requires Distiller 2.3)
- For hybrids, (Orbi, FT), advisable to save survey as profile and MS/MS as centroid

A common choice, particularly for hybrid traps is to save survey scans as profile and MS/MS scans as profile. If you are using iTRAQ or TMT, the quantitation uses the MS/MS data, so this is the one to worry about. As long as you have unit resolution or better in the reporter ion region and as long as the centroiding does not attempt to de-isotope, it probably makes little difference whether the MS/MS data are saved as centroid or profile. The most critical factor is the intensity of the reporter ions peaks. They need to be of comparable intensity to the base peak to give good counting statistics

If quantitation uses MS data, such as SILAC or ^{18}O or metabolic, then it is advisable to save the MS as profile. Distiller will be able to do a much better job of fitting the isotope distributions if it has true profile data. No disadvantage to save the MS/MS as centroids, if you want to keep the file size down.

For a standard trap, you will get much better results by using the zoom scans for quantitation, because you need isotopic resolution for 2+ and 3+ precursors to get reliable results. This requires Distiller 2.3

8. Can I use centroided data for quantitation?



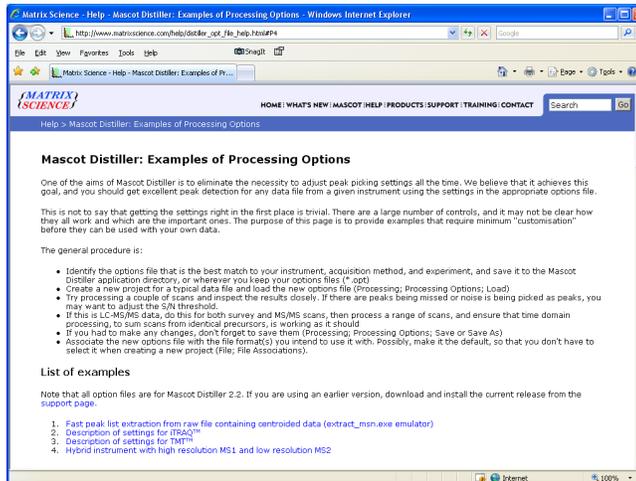
MASCOT : Quantitation FAQ

© 2009 Matrix Science



Setting the peak picking options in Distiller for a hybrid, where the MS and MS/MS spectra have very different characteristics, can be a little complicated. Some settings have a big impact on processing time, which can be an important consideration. There is a help page on our web site which goes into some detail about this. Go to Support; Mascot Distiller. The link is under Documentation.

8. Can I use centroided data for quantitation?



MASCOT : Quantitation FAQ

© 2009 Matrix Science



There are links to download example option files. Although this isn't relevant to quantitation, one of these is designed for raw files where both MS and MS/MS are centroids. It is a replacement for `extract_msn.exe`, which Thermo no longer distributes, and you'll find it is even faster than `extract_msn.exe` because it doesn't have to create a large number of individual *.DTA files

9. How do I export my quantitation results?

Copy table to clipboard

- Paste into Excel

Display HTML report in browser

- Print or save

Save XML

- For import into a relational database or another software application

There are three ways to get quantitation data out of Distiller.

The screenshot displays the Mascot Distiller interface. On the left is a 'Protein List' with various protein entries and their scores. The main window shows a 'Quantitation Table' with columns for Accession, Score, Mass, ML, SD(ipo), #, HL, SD(ipo), #, and Description. Below the table are two mass spectra plots: 'Triple Encoding SILAC sample data - Triple Encoding SILAC sample data.wiff - HWKLDQGTNR (light)' and 'Triple Encoding SILAC sample data - Survey 5105 (11-2494.2740)'. The bottom of the window features the Mascot logo, 'Quantitation FAQ', copyright information for Matrix Science, and the Matrix Science logo.

| Accession | Score | Mass | ML | SD(ipo) | # | HL | SD(ipo) | # | Description |
|-----------|-------|--------|--------|---------|--------|---------|---------|--------|-----------------------------|
| gH507357 | 743 | 22548 | 1.0273 | 1.0874 | 7 | 0.8999 | 1.1015 | 7 | transgelin 2 (Homo sapiens) |
| 1 | 22548 | 3 | 1.1271 | 0.02078 | 1.0754 | 0.00688 | 0.0583 | 0.9886 | 3067 |
| 2 | 22548 | 1196 | 1.0273 | 0.03133 | 0.9619 | 0.04612 | 0.5044 | 0.9091 | 483 |
| 3 | 22548 | heavy1 | 1.1417 | 0.03780 | 1.0009 | 0.03923 | 0.5794 | 0.9823 | 488 |
| 4 | 22548 | heavy2 | 0.5292 | 0.01931 | 0.7050 | 0.02470 | 0.4477 | 0.8222 | 6259 (2) Oxidatio... |
| 5 | 22548 | heavy3 | 0.8201 | 0.01768 | 0.6242 | 0.01699 | 0.9319 | 0.9832 | 3261 (3) Oxidatio... |
| 6 | 22548 | heavy4 | 1.0383 | 0.01219 | 0.8969 | 0.01708 | 0.9130 | 0.9886 | 3272 (7) Oxidatio... |
| 7 | 22548 | heavy5 | 0.9410 | 0.01878 | 0.8404 | 0.02097 | 0.8774 | 0.9917 | 658 (2) Oxidatio... |
| 8 | 22548 | heavy6 | 0.7761 | 0.0643 | 1.0741 | - | - | - | 1.938 |

Copy table to clipboard for Excel is often convenient for a quick calculation or to display something as a chart

| Accession | Score | Mass | ML | SD(geo) | # | HL | SD(geo) | # | Description | Correlation | Intensity | Modifications | |
|------------|-------|---------|----|---------|---------|----------|-------------|-----------|---------------|-------------|-----------|---------------|---|
| Sequence X | ML | Std Err | X | H/L | Std Err | Fraction | Correlation | Intensity | Modifications | | | | |
| 1 | 1 | + | 2+ | GPAYGLE | 1 | 1.1271 | 0.020788 | 1 | 1.0734 | 0.03089 | 0.8563 | 0.9698 | 3067.302 |
| 2 | 1 | + | 2+ | Query | Rank | M(calc) | Observed | Delta | Score | Component | | | |
| 3 | 38 | | | | 1 | 819.4239 | 410.7086 | -0.0189 | 51 | light | | | |
| 4 | 39 | | | | 1 | 819.4239 | 410.7135 | -0.01138 | 51 | light | | | |
| 5 | 66 | | | | 1 | 829.4322 | 415.7096 | -0.03555 | 48 | heavy | | | |
| 6 | 2 | + | 2+ | NVISLQM | 1 | 1.0273 | 0.031333 | 1 | 0.9619 | 0.046123 | 0.5044 | 0.9091 | 482.7506 |
| 7 | 1998 | | | | 1 | 1211.632 | 606.8346 | 0.022644 | 46 | heavy | | | |
| 8 | 3 | + | 2+ | TLMNLGG | 1 | 1.1417 | 0.03978 | 1 | 1.0009 | 0.036233 | 0.5794 | 0.9623 | 489.2471 |
| 9 | 1610 | | | | 1 | 1214.681 | 608.3565 | 0.017917 | 48 | light | | | |
| 10 | 1661 | | | | 1 | 1224.689 | 613.3569 | 0.010486 | 51 | heavy | | | |
| 11 | 4 | + | 2+ | NVISLQM | 1 | 0.9692 | 0.019312 | 1 | 0.876 | 0.02476 | 0.9107 | 0.9952 | 6929.108 (7) Oxidation (M) |
| 12 | 1617 | | | | 1 | 1217.619 | 609.8145 | -0.00412 | 59 | light | | | |
| 13 | 1618 | | | | 1 | 1217.619 | 609.8232 | 0.013196 | 40 | light | | | |
| 14 | 1669 | | | | 1 | 1227.627 | 614.8091 | -0.02338 | 54 | heavy | | | |
| 15 | 1970 | | | | 1 | 1227.627 | 614.8289 | 0.019102 | 34 | heavy | | | |
| 16 | 22 | | | | 1 | 1223.639 | 612.8172 | -0.01891 | 48 | medium | | | |
| 17 | 5 | + | 2+ | TLMNLGG | 1 | 0.9231 | 0.017678 | 1 | 0.8242 | 0.018058 | 0.9319 | 0.9926 | 9260.677 (3) Oxidation (M) |
| 18 | 1678 | | | | 1 | 1230.675 | 616.3393 | -0.01134 | 84 | light | | | |
| 19 | 1679 | | | | 1 | 1230.675 | 616.3438 | -0.00246 | 102 | light | | | |
| 20 | 1727 | | | | 1 | 1240.684 | 621.3354 | -0.02755 | 90 | heavy | | | |
| 21 | 1735 | | | | 1 | 1240.684 | 621.3628 | 0.023307 | 61 | heavy | | | |
| 22 | 1736 | | | | 1 | 1240.684 | 621.362 | 0.025727 | 93 | heavy | | | |
| 23 | 1737 | | | | 1 | 1240.684 | 621.3666 | 0.034907 | 64 | heavy | | | |
| 24 | 1708 | | | | 1 | 1236.696 | 619.3428 | -0.02262 | 60 | medium | | | |
| 25 | 1710 | | | | 1 | 1236.696 | 619.3619 | 0.013676 | 55 | medium | | | |
| 26 | 1711 | | | | 1 | 1236.696 | 619.3685 | 0.026936 | 60 | medium | | | |
| 27 | 6 | + | 2+ | GASQAGI | 1 | 1.0393 | 0.012189 | 1 | 0.8999 | 0.01706 | 0.913 | 0.9866 | 3272.118 (7) Oxidation (M);(12) Oxidation (M) |
| 28 | 2262 | | | | 1 | 1414.597 | 708.3155 | 0.019489 | 54 | light | | | |
| 29 | 2268 | | | | 1 | 1424.605 | 713.3186 | 0.017519 | 34 | heavy | | | |
| 30 | 2390 | | | | 1 | 1424.605 | 713.3228 | 0.027899 | 29 | heavy | | | |
| 31 | 2376 | | | | 1 | 1420.617 | 711.3319 | 0.024167 | 31 | medium | | | |
| 32 | 7 | + | 2+ | QMEDISQ | 1 | 0.941 | 0.018762 | 1 | 0.8404 | 0.029973 | 0.8774 | 0.9917 | 656.3021 (2) Oxidation (M) |
| 33 | 2631 | | | | 2 | 1693.829 | 847.9184 | 0.012927 | 52 | light | | | |
| 34 | 2636 | | | | 2 | 1703.818 | 852.9165 | 0.000796 | 47 | heavy | | | |
| 35 | 2634 | | | | 2 | 1699.829 | 850.9286 | 0.013245 | 41 | medium | | | |

The screenshot displays the Mascot Distiller software interface. The main window shows a list of proteins on the left and a table of search results in the center. The table includes columns for Sequence, Score, Mass, ML, SD(geo), #, HL, SD(geo), #, and Description. Below the table, there are two plots: a mass spectrum plot showing relative intensity versus m/z, and a chromatogram plot showing relative intensity versus time (min).

| Sequence | Score | Mass | ML | SD(geo) | # | HL | SD(geo) | # | Description | |
|---------------|--------|---------|--------|---------|---------|---------|---------|--------|---|------------------------|
| 2-IPAVQLQSR | 1.1271 | 102019 | 1.0273 | 1.0874 | 7 | 0.8989 | 1.1015 | 7 | transgelin 2 (Homo sapiens) | |
| 2-NVXQLGQTR | 1.0273 | 0.03133 | 1.1417 | 0.03780 | 1 | 0.9619 | 0.04812 | 0.5044 | 0.5091 | 483 |
| 2-TLNLGLAVAR | 0.9229 | 0.0231 | 0.9231 | 0.01703 | 0.01410 | 0.9261 | 0.9523 | 0.9261 | 0.9523 | 8258 (2) Oxidation (M) |
| 2-TLNLGLAVAR | 0.9231 | 0.01788 | 0.9231 | 0.01788 | 0.9242 | 0.01938 | 0.9519 | 0.9526 | 8201 (3) Oxidation (M) | |
| 2-GLSGAGMFLVY | 1.0393 | 0.01219 | 0.9899 | 0.01706 | 0.9130 | 0.8666 | 0.9130 | 0.8666 | 3272 (7) Oxidation (M) | |
| 1-peptide | 0.8410 | 0.01870 | 0.8464 | 0.02687 | 0.8714 | 0.9917 | 0.8714 | 0.9917 | 155 (2) Oxidation (M) | |
| 2-peptides | 2.7795 | 0.9463 | 1.0744 | 1.0744 | 4 | 0.9263 | 1.1104 | 4 | Dendrokinin (Dendroica striata) (Dendroica striata) (Dendroica striata) | |

Below the table, there are two plots. The top plot is a mass spectrum showing relative intensity versus m/z, with peaks labeled at 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, and 620. The bottom plot is a chromatogram showing relative intensity versus time (min), with peaks labeled at 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, and 620.

MASCOT : Quantitation FAQ

© 2009 Matrix Science



The quantitation reports are HTML and can be displayed in a browser. These reports are created by transforming the XML result file into HTML by means of an XSLT style sheet. This is a powerful technology, because anyone can modify these style sheets or create new ones to make custom reports.

Mascot Distiller Quantitation Report - Opera

File Edit View Bookmarks Widgets Tools Help

Mascot Distiller Quantitation Report

file://localhost/C:/Program%20Files/Matrix%20Science/Mascot%20Distiller%20Prototype/quant_report.html

H/L

| H/L | Accession | Score | Mass | M/L | SD(geo) | # | H/L | SD(geo) | # |
|-----|-----------|-------|-------|-------|---------|---|--------|---------|---|
| 1 | gi4507357 | 743 | 22548 | 1.027 | 1.087 | 7 | 0.8999 | 1.101 | 7 |

transgelin 2 [Homo sapiens]

| # | Sequence | Incl. | M/L | Std.Err. | Incl. | H/L | Std.Err. | Fraction | Correlation | Intensity | Modifications |
|---|------------------|-------|--------|----------|-------|--------|----------|----------|-------------|-----------|--------------------------------------|
| 1 | 2 GPAYGLSR | X | 1.127 | 0.02079 | X | 1.073 | 0.03089 | 0.8563 | 0.9598 | 3067 | |
| 2 | 2 NVIGLQMGITNR | X | 1.027 | 0.03133 | X | 0.9619 | 0.04812 | 0.5044 | 0.8091 | 482.8 | |
| 4 | 2 TLMNLGGLAVAR | X | 1.142 | 0.03780 | X | 1.001 | 0.03823 | 0.5794 | 0.9823 | 488.2 | |
| 4 | 2 NVIGLQMGITNR | X | 0.9692 | 0.01931 | X | 0.8760 | 0.02476 | 0.9107 | 0.9952 | 6929 | (7) Oxidation (M) |
| 5 | 2 TLMNLGGLAVAR | X | 0.9231 | 0.01768 | X | 0.8242 | 0.01806 | 0.9319 | 0.9926 | 9261 | (3) Oxidation (M) |
| 6 | 2 CASQAGMTGYGMPR | X | 1.039 | 0.01219 | X | 0.8999 | 0.01706 | 0.9130 | 0.9866 | 3272 | (7) Oxidation (M);(12) Oxidation (M) |
| 7 | 2 QMEQISQFLQAEER | X | 0.9410 | 0.01878 | X | 0.8404 | 0.02997 | 0.8774 | 0.9917 | 856.3 | (2) Oxidation (M) |

| H/L | Accession | Score | Mass | M/L | SD(geo) | # | H/L | SD(geo) | # |
|-----|-----------|-------|-------|--------|---------|---|--------|---------|---|
| 2 | gi118090 | 638 | 22785 | 0.9943 | 1.074 | 5 | 0.9263 | 1.132 | 5 |

Peptidyl-prolyl cis-trans isomerase B precursor (PPIase) (Rotamase) (Cyclophilin B) (S-cyclophilin) (SCVLP) (CYP-S1)

| # | Sequence | Incl. | M/L | Std.Err. | Incl. | H/L | Std.Err. | Fraction | Correlation | Intensity | Modifications |
|---|-------------|-------|--------|----------|-------|--------|----------|----------|-------------|-----------|---------------|
| 1 | 2 VVFDLR | X | 1.137 | 0.01448 | X | 1.169 | 0.03031 | 0.8299 | 0.9893 | 1994 | |
| 2 | 2 IGDDEDVGR | X | 0.9433 | 0.01618 | X | 0.8474 | 0.02149 | 0.8691 | 0.9936 | 5198 | |

MASCOT : Quantitation FAQ

© 2009 Matrix Science

MATRIX SCIENCE

The screenshot displays the Mascot Distiller software interface. The main window shows a list of proteins and their quantitation results. A table of results is visible, including columns for Sequence, Score, Mass, ML, SD(geo), #, HL, SD(geo), #, Intensity, and Modifications. The table lists several protein sequences and their corresponding scores and intensities. A 'Quantitation Reports' menu is open, showing options like 'Delete Quantitation Results', '1 peptides', and '2 proteins'. A 'Save As XML...' option is highlighted. Below the table, there are two mass spectrum plots. The top plot shows relative intensity versus m/z for a specific protein, and the bottom plot shows a survey mass spectrum for the sample data.

| Sequence | Score | Mass | ML | SD(geo) | # | HL | SD(geo) | # | Intensity | Modifications |
|--------------------|--------|---------|--------|---------|--------|--------|--|---|-----------|---------------|
| 2-GRVYGLSR | 1.1271 | 0.02079 | 1.0734 | 0.03889 | 0.8553 | 0.8553 | 30871 | | | |
| 2-NVYKLGQTR | 1.0273 | 0.03133 | 0.9619 | 0.04812 | 0.5044 | 0.8091 | 463 | | | |
| 2-TLNLGGLAVAR | 1.1417 | 0.03780 | 1.0008 | 0.03823 | 0.5794 | 0.8233 | 488 | | | |
| 2-NVYKLGQTR | 0.9228 | 0.01931 | 0.7170 | 0.04170 | 0.9562 | 0.8523 | 3229 (3) Oxidation (M) | | | |
| 2-TLNLGGLAVAR | 0.9231 | 0.01788 | 0.8242 | 0.01938 | 0.9319 | 0.9026 | 9261 (3) Oxidation (M) | | | |
| 3-GLASGAGMTDGLV... | 1.0393 | 0.01219 | 0.8999 | 0.01706 | 0.9130 | 0.8966 | 3272 (7) Oxidation (M) | | | |
| 1-peptides | 0.8410 | 0.01879 | 0.8484 | 0.02297 | 0.8774 | 0.9977 | 656 (2) Oxidation (M) | | | |
| 2-proteins | 22795 | 0.8483 | 1.0744 | 4 | 6.9063 | 1.1159 | 6 (Swedish Lymphoblastoid Interferon B precursor (S99en) (Ratemeas) (Carbonyl) | | | |

MASCOT : Quantitation FAQ © 2009 Matrix Science **MATRIX SCIENCE**

The quantitation results, together with all the intermediate data required to calculate or display these results, can be saved as an XML format file by choosing Save As XML... from the Quantitation Reports... sub-menu. This XML document provides a clean and powerful way to integrate Distiller quantitation with other applications, or to export the results to a relational database. The file format is described by a schema called `distiller_quantitation_1.xsd`, which can be found in the Distiller application directory.

The screenshot shows the Altova XML Spy interface displaying an XML file named 'Triple Encoding SILAC sample data.xml'. The main window shows a tree view on the left and a data grid on the right. The data grid is titled 'peptide (345)' and contains the following columns: m/z Calc, obs, delta, score, component, peptideStr, varModStr, and queryBank. The data is sorted by m/z Calc in ascending order.

| m/z Calc | obs | delta | score | component | peptideStr | varModStr | queryBank |
|------------|---------|-----------|-------|-----------|------------|-----------|-----------|
| 1 811 423 | 406 71 | -0.017868 | 41 | light | VVTELR | 00000000 | queryBank |
| 7 910 413 | 410 202 | -0.024853 | 45 | light | SVTEELR | 00000000 | queryBank |
| 3 919 424 | 410 71 | -0.018901 | 51 | light | GPAYQLSR | 00000000 | queryBank |
| 4 819 424 | 410 714 | -0.011381 | 51 | light | GPAYQLSR | 00000000 | queryBank |
| 5 821 435 | 411 714 | -0.017279 | 35 | heavy | VVTELR | 00000000 | queryBank |
| 6 828 422 | 415 197 | -0.041553 | 45 | heavy | SVTEELR | 00000000 | queryBank |
| 7 829 432 | 415 706 | -0.035551 | 48 | heavy | GPAYQLSR | 00000000 | queryBank |
| 8 853 453 | 427 723 | -0.02465 | 42 | heavy | LLDVNRR | 00000000 | queryBank |
| 9 959 404 | 430 894 | -0.020516 | 71 | light | IKDEYVRR | 00000000 | queryBank |
| 10 965 424 | 433 703 | -0.032698 | 83 | medium | IKDEYVRR | 00000000 | queryBank |
| 11 869 412 | 435 684 | -0.037887 | 74 | heavy | IKDEYVRR | 00000000 | queryBank |
| 12 973 492 | 437 737 | -0.031866 | 48 | light | DQLLTSR | 00000000 | queryBank |
| 13 979 512 | 440 743 | -0.040687 | 38 | medium | DQLLTSR | 00000000 | queryBank |
| 14 983 5 | 442 727 | -0.060876 | 56 | heavy | DQLLTSR | 00000000 | queryBank |
| 15 993 424 | 447 883 | -0.073458 | 83 | light | ADEGSFR | 00000000 | queryBank |
| 16 993 424 | 447 701 | -0.037758 | 56 | light | ADEGSFR | 00000000 | queryBank |
| 17 998 444 | 450 886 | -0.092539 | 87 | medium | ADEGSFR | 00000000 | queryBank |
| 18 998 444 | 450 709 | -0.040219 | 53 | medium | ADEGSFR | 00000000 | queryBank |
| 19 999 444 | 450 716 | -0.027919 | 53 | medium | ADEGSFR | 00000000 | queryBank |
| 20 903 433 | 452 876 | -0.091568 | 56 | heavy | ADEGSFR | 00000000 | queryBank |
| 21 903 433 | 453 879 | -0.088398 | 83 | heavy | ADEGSFR | 00000000 | queryBank |
| 22 914 457 | 458 223 | -0.025742 | 44 | light | EQGKILAR | 00000000 | queryBank |
| 23 923 435 | 462 707 | -0.034577 | 46 | light | TDEGAVR | 00000000 | queryBank |
| 24 929 455 | 465 72 | -0.020239 | 53 | medium | TDEGAVR | 00000000 | queryBank |
| 25 929 455 | 465 725 | -0.020519 | 53 | medium | TDEGAVR | 00000000 | queryBank |
| 26 933 443 | 467 716 | -0.024667 | 57 | heavy | TDEGAVR | 00000000 | queryBank |

Here, we see the XML file displayed in XMLSpy as a data grid

MASCOT : Quantitation FAQ

© 2009 Matrix Science



10. Should I choose normalisation? If so, which type is best?

When to use normalisation

- Analysing equal total weights of protein from a complex mixture, e.g. cell lysate - YES
- Starting with equal numbers of cells - NO
- Isolating a sub-set of proteins by affinity methods - NO
- Looking at a synthetic dilution series - NO

The correct time to use normalisation is when you have two or more samples, each of which is supposed to contain the same total amount of a reasonably complex mixture of proteins. Some proteins may be up-regulated, others down-regulated, but the majority are unchanged. In such a case, the mean ratio is expected to be close to unity. If it is not unity, this probably reflects systematic errors, such as failure to fully label or digest a sample, or maybe the sample aliquots were not identical. By forcing the mean ratio to unity, we can eliminate these systematic errors.

That is, normalisation forces the average or median peptide ratio to be 1. If the average or median ratio is supposed to be 1, this is the smart thing to do. In other cases, it is absolutely the wrong thing to do. For example:

- If you are analysing a standard, where the ratio is supposed to be 3:1, you wouldn't want to force it to be 1:1, would you?
- If you start with equal numbers of cells and want to see the effect of starvation on one line then normalisation would be incorrect because you expect less protein from starved cells
- If you are isolating a sub-set of proteins by affinity methods then normalisation is not appropriate

Whether you get better results using average or median depends on various factors. The safest way to choose is to run some standards (e.g. similar samples spiked with varying amounts of a known protein) and see which gives the best accuracy and precision.

Quantitation FAQ

1. What software do I need for iTRAQ or TMT?
2. What software do I need for SILAC or 18O?
3. Does Mascot support spectral counting?
4. Why do I keep getting errors about conflicting modifications?
5. Can quantitation be automated?
6. How do I create or edit a quantitation method?
7. Why do I get no results for certain peptide matches?
8. Can I use centroided data for quantitation?
9. How do I export my quantitation results?
10. Should I choose normalisation? If so, which type is best?