Quantitation - Introduction

Protein quantitation is a large topic, and this module introduces the key concepts.

Quantitation - Overview

Protocol	Basis	Ratios	Examples
reporter	Specific reporter ion peaks within a single MS/MS spectrum	Inter-sample	iTRAQ, ExacTag, TMT, TMTpro
precursor	Extracted ion chromatograms for related precursors within a single dataset	Inter-sample	ICAT, SILAC, ^{18}O, ICPL, AQUA, Metabolic
multiplex (Neubert et. al.)	Pairs of sequence ion fragment peaks within a single MS/MS spectrum	Inter-sample	SILAC, ^{18}O
replicate	Extracted ion chromatograms for identical precursors across two or more datasets	Inter-sample	Label-free
empai (Ishihama et. al.)	Protein coverage from a database search result	Intra-sample	N/A
average (Silva et. al.)	Extracted ion chromatograms for selected peptides per protein within a single dataset	Intra-sample	N/A

We have classified the various approaches into a limited number of protocols. So far, we have identified 6 distinct protocols.

Reporter is quantitation based on the relative intensities of fragment peaks at fixed m/z values within an MS/MS spectrum. For example, iTRAQ or Tandem Mass Tags.

Precursor is quantitation based on the relative intensities of extracted ion chromatograms (XICs) for precursors within a single data set. This is by far the most widely used approach, which can be used with any chemistry that creates a precursor mass shift. For example, 18O, AQUA, ICAT, ICPL, Metabolic, SILAC, etc., etc.

Multiplex is quantitation based on the relative intensities of sequence ion fragment peaks within an MS/MS spectrum. This is a novel approach, which can be used with any chemistry that labels one peptide terminus, creating a small mass shift, such as 18O or SILAC under certain conditions.

Replicate is label free quantitation based on the relative intensities of extracted ion chromatograms (XICs) for precursors in multiple data sets aligned using mass and elution time.

All these four methods are used to measure the relative abundance of a protein from sample to sample. For example, whether a particular protein is up or down regulated when an organism is stressed or diseased. The next two methods are used to estimate

the relative abundances of different proteins within a single mixture.

emPAI is quantitation for the proteins in a mixture based on protein coverage by the peptide matches in a database search result.

Average is quantitation for the proteins in a mixture based on the application of a rule to the intensities of extracted ion chromatograms (XICs) for the peptide matches in a database search result. For example, the average intensity for the three strongest peptide matches per protein.

The rows with a blue background are the protocols that implemented in the search engine, and don’t require any additional software.

Quantitation - Overview

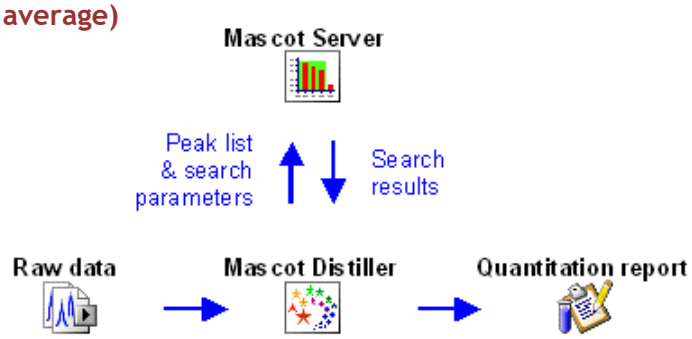
- **MS2: Quantitation methods that only require information available in the MS/MS peak list are supported in Mascot Server**
  - reporter
  - multiplex
  - emPAI
- **MS1: Methods that require additional information from the raw data file need Mascot Distiller + Quantitation Toolbox**
  - precursor
  - replicate
  - average

For the first three methods, the information required for quantitation is contained in the peak list. This is known as MS2 based quantitation.

The other three methods require additional information from the raw data file, either because it is necessary to integrate the elution profile of each peptide or because information is required for multiple peaks in the survey scan. These methods require that the raw data files are processed using Mascot Distiller. These are MS1 based methods.

Quantitation - Overview

Workflow for methods that require additional information from the raw data file, (precursor, replicate, average)

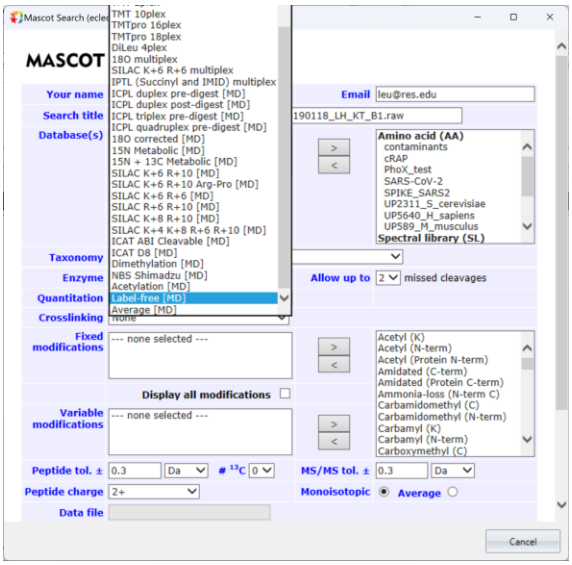


MASCOT : *Quantitation* MATRIX SCIENCE

For methods that require additional information from the raw data file, 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.

Quantitation

Named quantitation methods keep the search form uncluttered



We want to keep the user interface simple. Quantitation adds a huge number of choices and parameters, but there is no point in exposing all of these in the search form.

The approach we have chosen is encapsulate these choices and parameters into named quantitation methods. This means that the search form has just a single control.

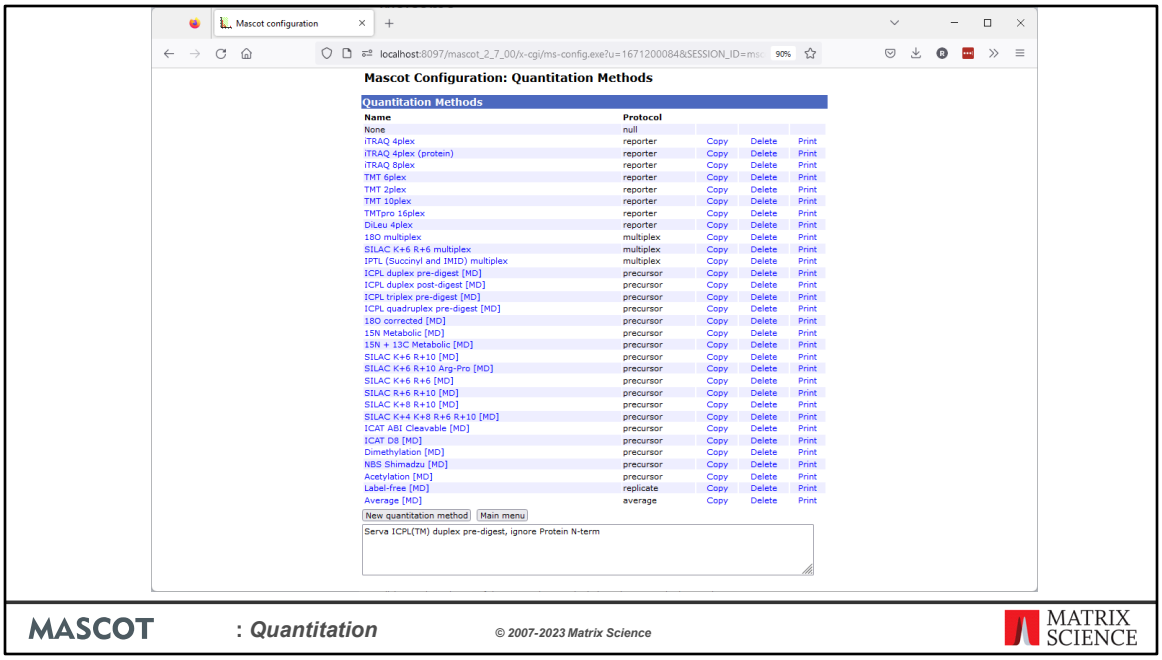
Methods that have [MD] at the end are the ones that require Mascot Distiller

Quantitation

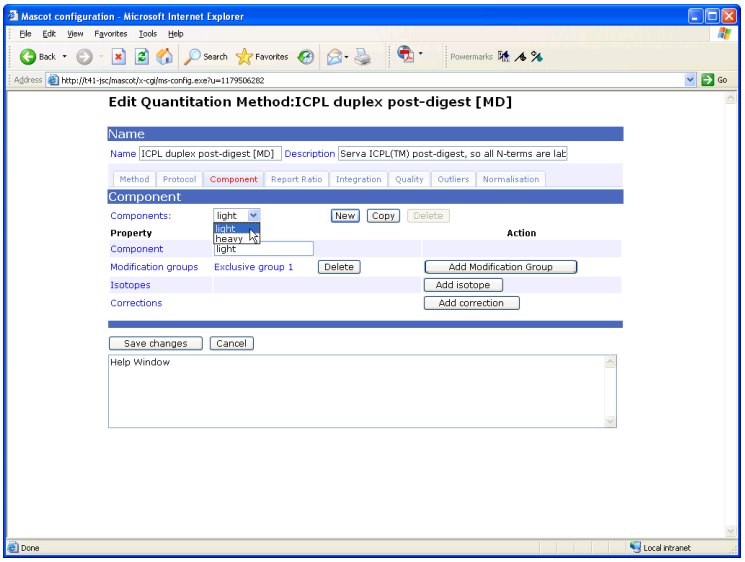
The quantitation methods are defined in a single XML configuration file

- quantitation.xml
- Browser based editor
- Add new methods as required
- Used by Mascot Server and Mascot Distiller

The configuration file that encapsulates the choices and parameters for each quantitation method is called quantitation.xml. This is an XML file, and there is a browser based editor for modifying methods and creating new ones. quantitation.xml lives on the Mascot server and is read by both the search engine and Mascot Distiller.



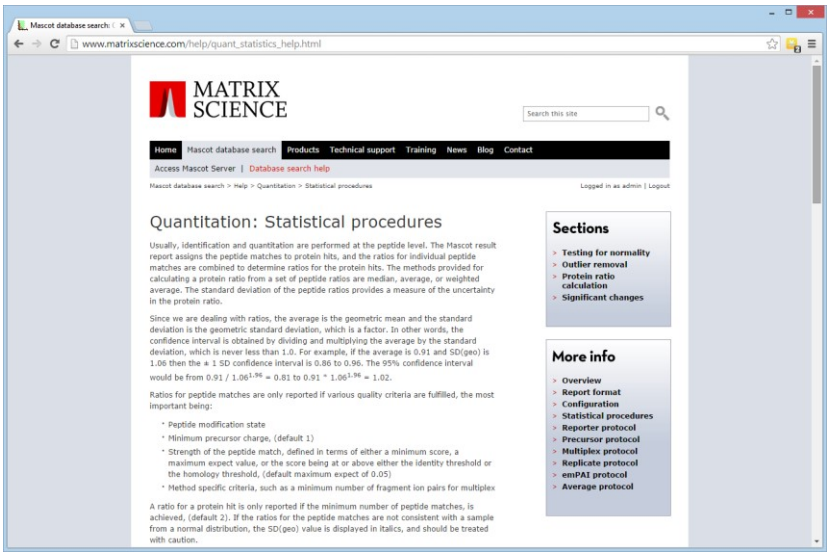
The browser-based Configuration Editor provides an interface to all the Mascot configuration files. In the case of quantitation, you can edit an existing method or make a copy of it as the basis for a new method.



MASCOT : *Quantitation* MATRIX SCIENCE

For each method, a tabbed dialog is used to navigate between property pages. In many cases, the property pages correspond to XML elements, but sometimes elements have been combined onto a single page or split across multiple pages so as to give a balanced layout.

Here, we can see a duplex ICPL method. The unlabelled and labelled components have been called heavy and light, but you are free to choose your own names so as to make the final report as clear as possible.



We have taken trouble to ensure that appropriate statistical procedures are correctly used. For example, we test that a set of peptide ratios is consistent with a normal distribution before rejecting outliers or reporting a standard deviation. Standard deviations are always geometric, because we are dealing with ratios that conform to a normal distribution in log space.

Selected Literature

- Ross, P. L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents*, Molecular & Cellular Proteomics 3 1154-1169 (2004) – iTRAQ
- Zhang, G. A. and Neubert, T. A., *Automated comparative proteomics based on multiplex tandem mass spectrometry and stable isotope labeling*, Molecular & Cellular Proteomics 5 401-411 (2006) - Multiplex
- Beynon, R. J. and Pratt, J. M., *Metabolic labeling of proteins for proteomics*, Molecular & Cellular Proteomics 4 857-872 (2005) - Metabolic
- Ong, S. E. and Mann, M., *Mass spectrometry-based proteomics turns quantitative*, Nature Chemical Biology 1 252-262 (2005) – General review
- Lill, J., *Proteomic tools for quantitation by mass spectrometry*, Mass Spectrometry Reviews 22 182-194 (2003) – General review
- Julka, S. and Regnier, F., *Quantification in proteomics through stable isotope coding: A review*, Journal of Proteome Research 3 350-363 (2004) – General review
- Bantscheff, M., et al., Quantitative mass spectrometry in proteomics: a critical review, Analytical and Bioanalytical Chemistry 389 1017-1031 (2007) – General review

These papers describe each approach in detail.

Quantitation talks

- K1. Quantitation - Introduction
- K2. Quantitation - MS2 based methods
- K3. Quantitation - MS1 based methods
- K4. Quantitation - Reporting

Please see the other quantitation presentations to learn about reporter ions, SILAC and label free quantitation as well as reporting formats for the results.

- K1. Quantitation – Introduction.
- K2. Quantitation - MS2 based methods. Quantitation methods that only require information available in the MS/MS peak list are supported in Mascot Server.
- K3. Quantitation - MS1 based methods. Methods that require additional information from the raw data file require Mascot Distiller + Quantitation Toolbox.
- K4. Quantitation – Reporting.