

Quantitation - Introduction

MASCOT

: *Quantitation*

© 2007-2023 Matrix Science



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

MASCOT

: *Quantitation*

© 2007-2023 Matrix Science

MATRIX
SCIENCE

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

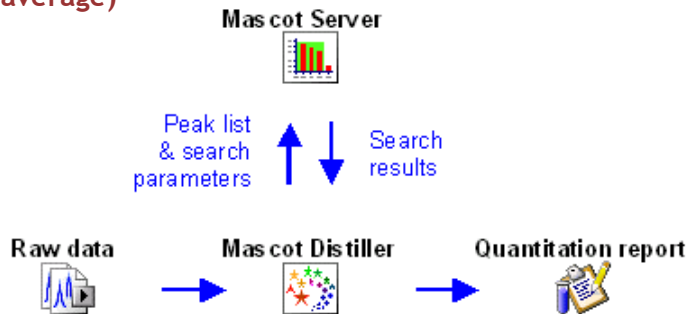
- **MS2: Quantitation methods that only require information available in the MS/MS peak list are supported in Mascot Server**
 - reporter
 - multiplex
 - emPAI
- **MS1: Methods that require additional information from the raw data file need Mascot Distiller + Quantitation Toolbox**
 - precursor
 - replicate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

© 2007-2023 Matrix Science

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

The screenshot shows the Mascot Search interface with the 'Quantitation' section highlighted. The 'Quantitation' dropdown menu is open, showing options: 'Label-free [MD]', 'Average [MD]', and 'none'. The 'Average [MD]' option is selected. The interface also shows other sections like 'Database(s)', 'Enzyme', 'Taxonomy', 'Crosslinking', 'Fixed modifications', and 'Variable modifications'. The 'Peptide tol.' is set to 0.3 Da, 'Peptide charge' is 2+, and 'MS/MS tol.' is 0.3 Da. The 'Monoisotopic' radio button is selected, and the 'Average' radio button is also selected. The 'Data file' field is empty.

MASCOT

: *Quantitation*

© 2007-2023 Matrix Science

**MATRIX
SCIENCE**

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.

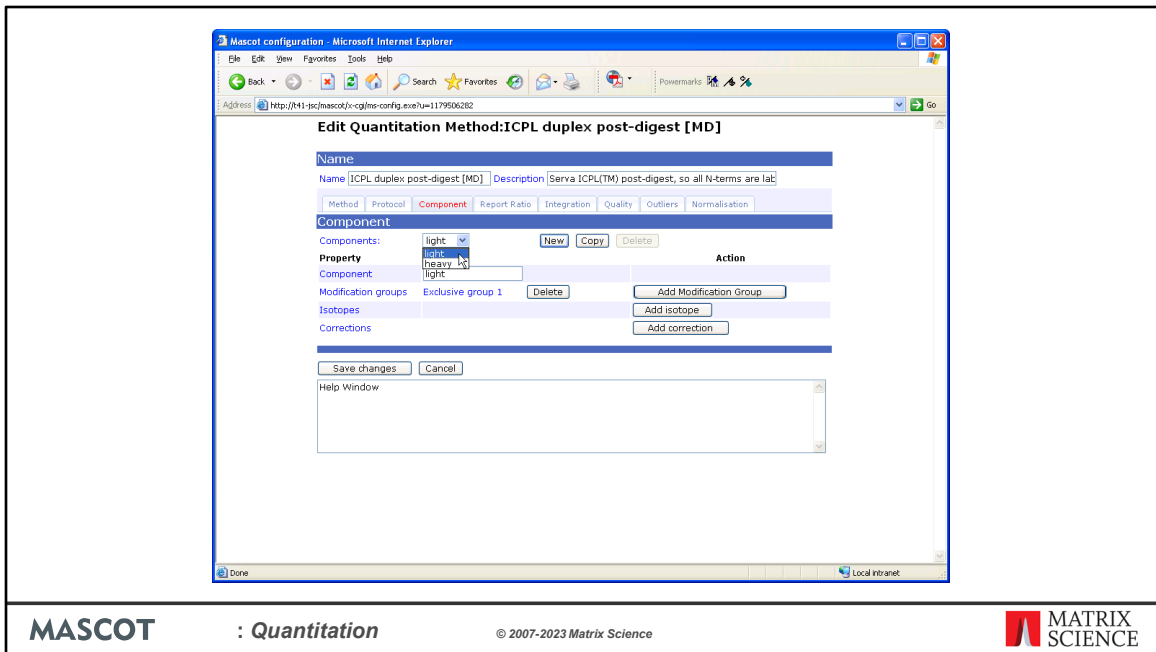
Mascot Configuration: Quantitation Methods

Name	Protocol	Copy	Delete	Print
None	null			
ITRAQ 4plex	reporter	Copy	Delete	Print
ITRAQ 4plex (protein)	reporter	Copy	Delete	Print
ITRAQ 8plex	reporter	Copy	Delete	Print
TMT 6plex	reporter	Copy	Delete	Print
TMT 2plex	reporter	Copy	Delete	Print
TMT 10plex	reporter	Copy	Delete	Print
TMTpro 16plex	reporter	Copy	Delete	Print
DiLeu 4plex	reporter	Copy	Delete	Print
18O multiplex	multiplex	Copy	Delete	Print
SILAC K+6 R+6 multiplex	multiplex	Copy	Delete	Print
IFTL (Suconyl and IMD) multiplex	multiplex	Copy	Delete	Print
ICPL duplex pre-digest [MD]	precursor	Copy	Delete	Print
ICPL duplex post-digest [MD]	precursor	Copy	Delete	Print
ICPL triplex pre-digest [MD]	precursor	Copy	Delete	Print
ICPL quadruplex pre-digest [MD]	precursor	Copy	Delete	Print
18O corrected [MD]	precursor	Copy	Delete	Print
15N Metabolic [MD]	precursor	Copy	Delete	Print
15N + 13C Metabolic [MD]	precursor	Copy	Delete	Print
SILAC K+6 R+10 [MD]	precursor	Copy	Delete	Print
SILAC K+6 R+10 Arg-Phe [MD]	precursor	Copy	Delete	Print
SILAC K+6 R+6 [MD]	precursor	Copy	Delete	Print
SILAC R+6 R+10 [MD]	precursor	Copy	Delete	Print
SILAC K+8 R+10 [MD]	precursor	Copy	Delete	Print
SILAC K+4 K+8 R+6 R+10 [MD]	precursor	Copy	Delete	Print
ICAT ABI Cleavable [MD]	precursor	Copy	Delete	Print
ICAT D8 [MD]	precursor	Copy	Delete	Print
Dimethylation [MD]	precursor	Copy	Delete	Print
NBS Shimadzu [MD]	precursor	Copy	Delete	Print
Acetylation [MD]	precursor	Copy	Delete	Print
Label-free [MD]	replicate	Copy	Delete	Print
Average [MD]	average	Copy	Delete	Print

New quantitation method

Serve ICPL(TM) duplex pre-digest, ignore Protein N-term

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.



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.

The screenshot shows a web browser window displaying the Mascot database search help page. The page title is "Quantitation: Statistical procedures". The content explains that identification and quantitation are performed at the peptide level. It details how peptide matches are combined to determine ratios for protein hits and how the standard deviation of peptide ratios provides a measure of uncertainty. It also discusses the geometric mean and standard deviation for ratios, and provides a formula for the 95% confidence interval. The page includes a navigation menu, a search bar, and a sidebar with sections like "Testing for normality" and "Outlier removal".

MASCOT : *Quantitation* © 2007-2023 Matrix Science

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.