A New Software for Automated, High-Throughput Quantitative Proteomics

Bernard Delanghe1, Rosa Viner2, Torsten Ueckert1
1Thermo Fisher Scientific (Bremen) GmbH, Bremen, Germany; 2Thermo Fisher Scientific, San Jose, CA, USA

Overview

Purpose: Quantification of precursor ion labeled peptides.

Methods: New algorithms.

Results: Fast, accurate and precise protein quantification for SILAC and isotopic labeling techniques.

Introduction

One of the major goals in proteomics research is the accurate quantification of protein expression in biological systems. Stable amino acid isotope labeling (SILAC) is widely used for relative protein quantification.1,2 However, analyzing those large and complex datasets is still challenging and requires sophisticated data reduction algorithms and quantification schemes.

Here we describe a new workflow in Thermo Scientific Proteome Discoverer software for automated, fast and accurate SILAC quantification. With the introduction of high resolution and high mass accuracy Orbitrap™ analyzers it now has become possible to analyze a wide range of SILAC experiments, including experiments using heavy and light lysine, arginine, secolucine, di-methyl lysine and other user defined labeling. The entire workflow from data acquisition to data analysis can be easily automated from within the acquisition sequence using Proteome Discoverer™ Daemon together with the Thermo Scientific Xcalibur sequence editor.

Methods

The complete workflow is shown in Figure 1 and consist of two parallel processes: identification and quantification.

Identification

The identification is done as usual and consists of 2 steps:
1. Selection of the spectra
2. Identification of the selected spectra using Mascot, Sequest or both of the search algorithms.

Quantification

Accurate detection and protein quantification consists of 3 steps:
1. Event detection: after noise removal all events (features) are detected and the individual peak areas are calculated.
2. Precursor quantification: The identified peptides are associated with the detected events based on mass accuracy and retention time. Using the information included in the quantification method, the theoretical mass(es) of the corresponding labeled pair (light, medium or heavy) are calculated. Those are then associated with the corresponding events the same way as the identified peptides. This information is also used to validate the results and errors are flagged. The peptide ratios are calculated using the same number of isotopes (Figure 3).
3. The protein ratio is calculated using the median peptide ratio (Figure 6).

All raw files selected in the Spectrum Files step will be automatically processed. This can be further automated by adding the Proteome Discoverer Daemon as post processing method in the Sequence Editor of the Xcalibur™ software. In this configuration the raw files acquired on the instrument computer are automatically send to the data analysis computer and processed with the specified workflow.

Results

A large number of raw files from different SILAC experiments have been processed to evaluate the performance of the quantification algorithm and the speed of the software. Samples with different types of labels as well as with different fixed ratios were analyzed. All samples were analyzed using Thermo Scientific LTQ Orbitrap XL or LTQ Orbitrap Velos hybrid mass spectrometers.

Robustness and quality

To evaluate the quality and robustness of the software we processed the benchmark data set consisting of 72 LC/MS runs provided with the MaxQuant software.3

The result are summarized in Figure 2A. In total around 210,000 peptides have been identified with a false discovery rate (FDR) better than 1%. About 90% of those peptides have been quantified.

FIGURE 2. Summary of the quantification result of 72 raw files (A) and the main parameters used for the protein identification (B).

A. Number of Peptides at 1% FDR
   Inconsistently labeled 437
   No quantification values 24130
   Not unique 12122
   Redundant 81804
   Unique 91484
   Total 209977
   Peptides quantified 89%

B. Search Conditions
   Search Engine: Mascot
   FASTA: IPF_Human
   Peptide Tolerance: 10 ppm
   Fragment Tolerance: 0.8 Da
   Dynamic Modifications: Oxidation (M); Lys8 (K); Arg10 (R);
   Fixed Modifications: Carbamidomethylation (C)

Speed

It takes about 8 minutes to quantitate an average raw file (~300MB, 12,500 MS2 spectra) of a SILAC experiment on a typical desktop computer. Computation time varies with file size, number of full scans, sample concentration and complexity.

FIGURE 3. Example of SILAC pair quantification. All detected isotopes are indicated by a blue circle (spectrum in lower panel). Isotopic peaks included in the quantification calculations are highlighted with filled blue circles, ensuring consistency and precision of peak area determination.

Conclusion

Proteome Discoverer software provides fast, accurate and precise protein quantification for SILAC and isotopic labeling techniques. Robustness and speed of the quantification algorithm enable quantification of the majority of peptides from large datasets in an economical fashion.

The entire workflow from data acquisition to data analysis can be easily automated from within the acquisition sequence using Xcalibur Sequence Editor together with the Proteome Discoverer Daemon.

References