A New Software for Automated, High-Throughput Quantitative Proteomics

Bernard Delanghe, Rosa Viner, Torsten Ueckert

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. However, analyzing large and complex datasets is still challenging and requires sophisticated data reduction algorithms and quantitation 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 has now become possible to analyze a wide range of SILAC experiments, including experiments using heavy and light lysine, arginine, isoleucine, 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 consists of three 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.

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

The results shows that the median peptide ratio is 7.6 with a standard deviation (SD) of 1.5 for this example. As the heavy labels are not completely incorporated, the labeling efficiency is ~96% for this example and quantification is resulting in an overall lower peptide ratio.

This example demonstrates that accurate peptide quantification can be obtained over 4 orders of magnitude of precursors ion abundances at 10:1 ratios (Figure 4B).

Speed

It takes about 6 minutes to quantify 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.

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


This is a trademark of Thermo Fisher Scientific. Thermo Scientific is a registered trademark of the University of Washington. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. No permission is granted to use any trademark without the written consent of the trademark owner.