Systematic Assessment of the Reproducibility of Relative Quantification Based on LC-MS with Replicates

Overview

Purpose: To develop a method to assess and display the reproducibility of protein quantification based on LC-MS with replicate analysis.

Methods: Technical replicates were used to assess the reproducibility of relative quantification at three levels:

Data level - Peptide features from each of the replicates were detected and aligned between the replicates. The data reproducibility was analyzed by paired T-test and correlation analysis on the intensities of paired features. The boxplot diagram was used to represent the distribution of the logarithmic ratios of feature pairs. Venn diagrams of peptide features between replicates were also used to display the reproducibility.

Identification level - Protein identification was conducted on each replicate. The reproducibility of the identified peptides and proteins between replicates was assessed by paired T-test and correlation analysis. The boxplot diagram was used to represent the distribution of the logarithmic ratios of identified feature pairs.

Quantification level - The relative peptide/protein abundance was computed based on the intensities of each peptide ion. The reproducibility of peptide and protein quantification was assessed by paired T-test and/or ANOVA.

Results: This method was implemented into the PEAKS® Studio software and tested by a standard protein mixture. The standard mixture was reduced and alkylated by iodoacetamide, then digested by trypsin overnight. MS data was obtained from a Waters Q-TOF Premier. Two replicate experiments were performed. MS raw data was analyzed with the PEAKS® de novo sequencing, database search and label-free quantification algorithms (Figure 2). S10 and S21 are two replicates of one sample and S20 and S21 are replicates of another sample. The results of reproducibility assessment at data level are shown in Figure 3. The number of peptide features in sample S10 and Sample 11 are 2794 and 2939, respectively. The common features is 2794 (Figure 3A). The scatterplot (Figure 3B) and boxplot of logarithm of the ratio of intensities (Figure 3C) of common peptide features showed a good data reproducibility. The Pearson correlation value is 85% for common features.

The results of reproducibility assessment at the identification level are shown in Figure 4. The total number of identified feature pairs (at least one feature in the feature pair is identified) in S10 and S11 is 563. The corresponding Pearson correlation value is 92%. Figures 4A and 4B are the score chart and Venn diagram for identified peptides and 4C and 4D for identified proteins S10 and S11.

Conclusions: This method proves to efficiently assess and display the reproducibility at three levels, which helps to evaluate both the LC-MS experiment and the quantification algorithm.

References


Acknowledgement

We thank Dr. Gilles A. Lajuste for the mass spectrometry experiment. PEAKS is a registered trademark of Bioinformatics Solutions Inc.