TMT-Based Quantitative Proteome Analysis with PEAKS

Jonathan R. Krieger and Baozhen Shan
Bioinformatics Solutions Inc., Waterloo, Canada

Abstract

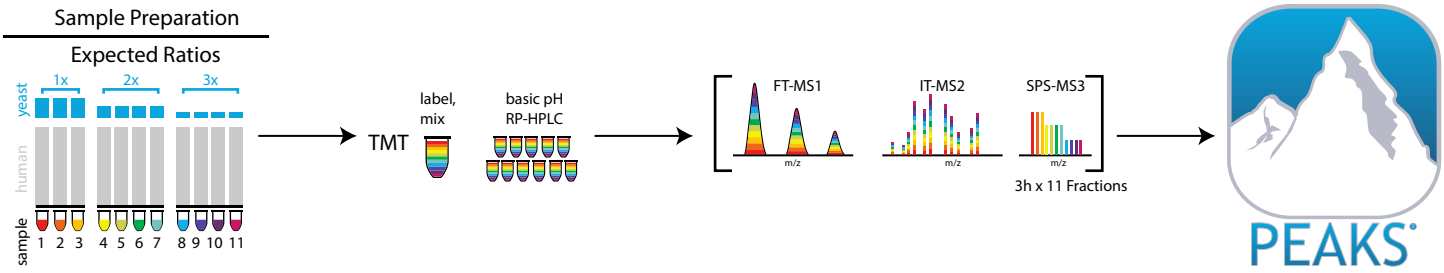
A published dataset [1] was used to illustrate the performance of PEAKS X database and quantification functions for Tandem Mass Tag (TMT)-labelled proteomic data. PEAKS provides both an overview as well as a samples-based view of proteins and peptides, as well as associated reporter ions. The data set described herein shows that PEAKS and the PEAKS Q module can effectively identify and quantify yeast proteins spiked into a complex human background, and quantify these proteins in the exact ratios they were spiked in.

Introduction

PEAKS is a complete data analysis tool for bottom-up, quantitative proteomics. Specifically, the add-on PEAKS Q module supports quantitative analysis of both label-free and label-based (e.g. SILAC, TMT, iTRAQ), data.

Once the database search is completed, the PEAKS Q module can be used to interrogate the data and determine the changes in relative protein abundance. Particularly in regard to TMT-labelled quantification, PEAKS Q can be customized by the end user to specify:

- Whether MS2 or MS3 is being used,
- TMT-lot specific correction factors,
- Sample organization (e.g. replicates, multiple TMT runs combined through common channel)



Modified from O' Connell et al., Journal of Proteome Research, 2018; 17(5):1934–1942

Figure 1: Schematic diagram of a TMT SPS-MS3 multiplexing experiment workflow, from sample preparation to data generation and PEAKS software analysis. In this experiment, a two-proteome sample was created by mixing human and yeast cell lysates in each of the 11 TMT-labeled samples according to the ratios indicated in the figure.

Tandem Mass Tags

The use of TMT has been increasing in quantitative proteomic experiments because of their robustness, precision, and reduction in mass spectrometry time compared to label free quantification [2]. TMT tags offers the ability to label up to 11 different samples which can then be multiplexed together to perform concurrent MS analysis. This type of labelling occurs at the peptide level, thus allowing cells to be derived from any biological source including cells, tissues,

and fluids. Every TMT reagent shares an identical structure consisting of a reactive region which binds primary amines, a mass balance region, and a reporter region. The reporter region of each tag (which breaks away from the rest of the molecule under HCD fragmentation) contains various combinations of 13C and 15N isotopes which differentiates each of the 11 TMT labels.

➢ TMT Quantification

Unlike precursor-based quantification methods (Label Free Quantification, SILAC) which use the MS1 for quantification, reporter-based quantification methods use the MS2 (or MS3) scans for quantification.

Proteins isolated from cells or tissues are reduced, alkylated and digested. Samples are then individually labelled and mixed before sample fractionation and cleanup. Once prepared, the samples are analyzed using high resolution LC-MS/MS to identify peptides and quantify the reporter ion abundance. Since all isobaric TMT tags have identical masses and chemical properties, all isotopologues will co-elute together.

Reporter ions can be quantified at the MS2 or MS3 level, depending on various factors including the type of instrument being used, and the complexity of the sample. Running complex samples in MS2 mode can lead to co-isolation of interfering species, potentially leading to a reporter ion ratio distortion, but will lead to a greater number of identified proteins. Conversely, the running samples in MultiNotch MS3 mode will result in more accurate quantification, but fewer identifications.

Methods

A published dataset [1] was downloaded and used to illustrate TMT quantification in PEAKS Studio X. Details of the dataset are detailed as follows.

➢ Study Aims

This study was designed in a way to benchmark TMT-labelled quantification to Label-Free Quantification (LFQ) through the comparison of the two techniques to determine which method is better to measure various fold changes in low abundant proteins.

➢ Experimental Design

Yeast lysate was spiked into human lysate at 3 different percentages relative to total protein content: Group 1 – 10%; Group 2 – 5%; and Group 3 – 3.3%. This was done to achieve a 2x-fold change between the first and second group, a 3x-fold change between group 1 and group 3, and a 1.5x-fold change between the final two groups. The human proteins formed a background and were at consistent levels across all groups. Three replicates from group 1 (126,127N, 127C) and four replicates from group 2 (128N,128C,129N,129C) and group 3 (130N, 130C, 131, 131C) were labelled, mixed and then subjected to offline high pH reverse phase separation. Fractions were collected, concatenated into 11 samples and run using 3-hour gradients on an Orbitrap Fusion Lumos instrument in MS3 mode.

➢ Data analysis

The TMT fractions were searched as one sample against a combined database of SwissProt human and *Saccharomyces cerevisiae* sequences (May 2019) in PEAKS Studio X. Both TMT-11 plex (229.16 @K, Peptide N-term), Carbamidomethylation (+57.02 @ C) were set as fixed modifications. Oxidation (15.99 @ M) and deamidation (0.98 @ NQ) were added as variable modifications. A precursor mass tolerance of 10 ppm and a fragment mass tolerance of 0.6 Da were applied. A maximum of 2 missed cleavages were allowed and 1% FDR was applied at the peptide-spectrum match (PSM) level. Normalized TMT reporter ion intensities were used for data analysis.

Results and Discussion

➢ TMT Quantification in PEAKS X

- In total, the database search identified 8840 protein groups with 1% FDR set to the PSM. 8556 of these proteins were able to be quantified (96.8%) with the same 1% FDR cutoff.
- The PEAKS X Quantification module normalizes the TMT data in each channel based on TIC or a user specified protein. Box plots of the data confirm the normalization across all samples (Figure 2).
- PEAKS X accurately quantifies the expected fold changes differences in the yeast proteins, and this is displayed in the generated heat map (Figure 3).

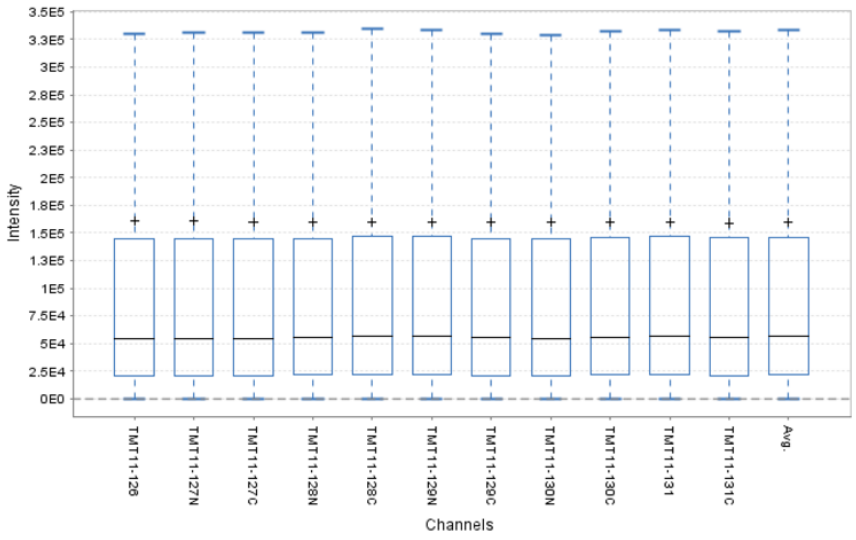


Figure 2: Box Plots show average intensity of all proteins across each TMT channel.

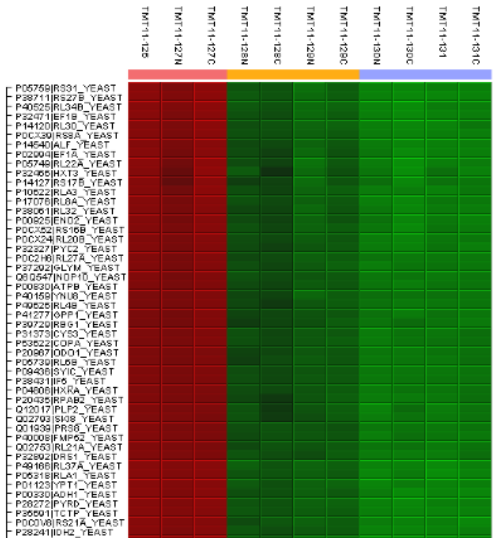


Figure 3: Partial heat map shows the top proteins with an abundance fold change ≥ 2.

- Individual proteins and peptides can be further accessed in detail in an organized, easy-to-navigate, view. In the protein level view (Figure 4), PEAKS Studio displays an interactive coverage map of the protein and a clickable volcano plot to quickly locate the information from the protein table for the protein of interest. In the peptide level view (Figure 5), PEAKS Studio displays an

Figure 4: Screenshot shows individual protein view of TMT labelled data in PEAKS Studio X.

interactive, annotated fragmentation spectrum which corresponds to the peptide selected from the peptide table. The matching fragment ions are then summarized in a table below. In addition, the reporter ions associated with the peptide selected from the peptide table as well as the peptide-feature envelope from the LC-MS, is displayed in the same view.

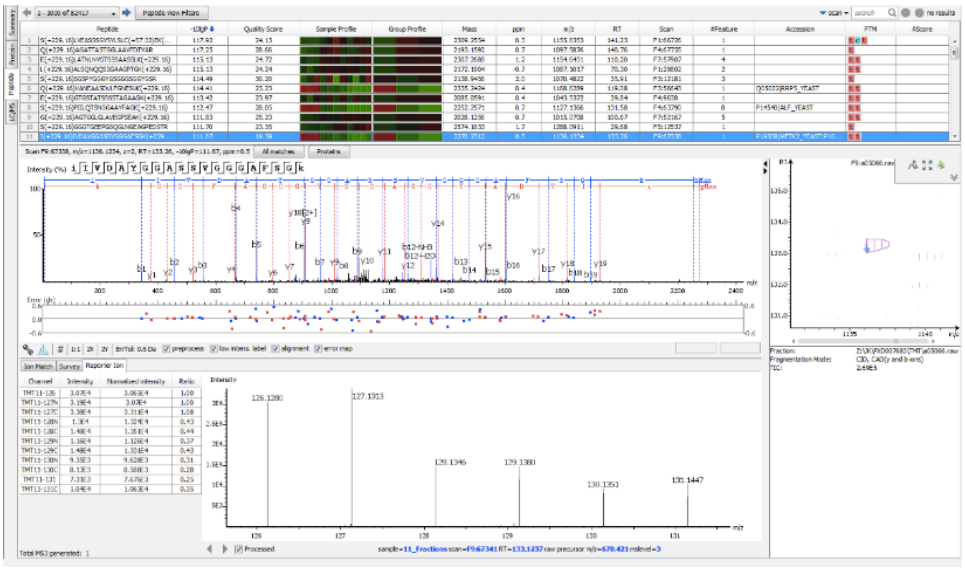


Figure 5: Screenshot shows individual peptide view of TMT-labelled data in PEAKS Studio X.

Conclusions

PEAKS X provides accurate identification and quantification of TMT-labelled samples. Over 96% of identified proteins were accurately quantified. Using the PEAKS Studio GUI, the protein and peptide results can then be analyzed in detail and validated from the project level, to a comparison between groups, and finally further down as a comparison between each individual sample. The interactive interface and various export functions, specifically for TMT analyzed data, facilitates users performing proteome quantification. As a result, PEAKS Studio X provides a reliable and accurate way to identify and quantify LC-MS/MS data obtained from complex fractionated TMT-based proteomic studies.

1. O'Connell, J.D., Paulo, J.A., O'Brien, J.J., Gygi, S.P.: Proteome-Wide Evaluation of Two Common Protein Quantification Methods. - PubMed - NCBI. J. Proteome Res. 17, 1934–1942 (2018).
2. Krieger, J.R., Wybenga-Groot, L.E., Tong, J., Bache, N., Tsao, M.S., Moran, M.F.: Evosep One Enables Robust Deep Proteome Coverage Using Tandem Mass Tags while Significantly Reducing Instrument Time. J. Proteome Res. 18, 2346–2353 (2019).