**ABSTRACT**

The most popular acquisition strategies for multiplexed quantitation using tandem mass tags (TMT) are broadly grouped into MS2 and MS3 based approaches. Here we demonstrate the benefit of adding machine learning (ML) algorithms to TMT data analysis tools. Addition of INFERYS and CHIMERYS increase identified and quantified proteins. Furthermore, we recommend data analysis workflows that best suit MS2 and MS3 based methods.

**INTRODUCTION**

Multiplexed proteomics quantitation using Thermo Scientific™ Tandem Mass Tags™ (TMT™) is a powerful tool to measure differences in animal states. Much effort has been applied to instrument development and development of smart acquisition strategies such as real time search (RTS) and synchronous precursor selection MS2 (SPS-MS2) to improve the quantitative performance of TMT workflows. At the same time, there has been significant progress leveraging data mining and AI to enhance TMT quantification. The integration of machine learning (ML) has become a valuable approach in the application of deep learning algorithms (INFERYS and CHIMERYS) to increase identified and quantified TMT labeled peptides in MS2 and MS3 based acquisition methods.

**MATERIALS AND METHODS**

Sample preparation and data acquisition were carried out as described in Flintewanger et al. MCP, 2022 (PXD020007); Kisting et al., Proteomics, 2022 (PXD012777); and Paulo et al. JAMB, 2021 (PXD028515). From these studies were used here. CHIMERYS and INFERYS algorithms were implemented as noise in the Proteome Discoverer 3.1 software, as described from Paul et al., in Thermo Scientific™ Proteome Discoverer™ 3.1 using the updated INFERYS/CHIMERYS model that now supports phosphorylation modifications. CHIMERYS was deployed in the cloud and received prepared data from a local instance of Proteome Discoverer 3.0. Once processed by the CHIMERYS service, results were transmitted to the local instance for final processing and results viewing.

**CHIMERYS AND INFERYYS WORKFLOWS**

Figure 1. Representative TMT Processing Workflows in Proteome Discoverer

**IMPROVEMENTS FOR MS2 AND MS3 TMT DATA ANALYSIS**

While CHIMERYS and INFERYS show improvements over SEQUEST in most cases (Figure 3 & 4) demonstrate that CHIMERYS offers the greatest improvement for MS2 based TMT data with INFERYS and CHIMERYS providing comparable performance for MS3 based approaches. These results show the impact of data complexity on the results. For more complex data such as the MS2 data that CHIMERYS can accurately decompose and predict, complex chimeric spectra (Figure 5) while INFERYS is limited to the PSMs assigned by SEQUEST (via Precursor Detector node), which assigns fewer chimeric spectra. However, for MS3 based approaches (Figure 4) INFERYS and CHIMERYS perform similarly.

**REFERENCES**


**TRADEMARKS/LICENSEING**

© 2023 Thermo Fisher Scientific Inc. All rights reserved. SEQUEST is a trademark of the University of Washington. CHIMERYS and INFERYS are trademarks of MASCOT. TMT™ is a trademark of Pro陕西省 Sci. Technol. Limited. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others. P24203-58EN