Rapid and Accurate Global PTM Discovery (GPTM-D) Using Post-Acquisition Spectral Calibration and Defined Mass Windows

# Introduction

The calibration of spectra files and algorithmic improvements in the peptide dictionary search improves the accuracy and efficiency of new posttranslational-modification (PTM) identification.

# Motivation

Physical measurements introduce noise, and instruments that take multiple measurements often have correlated noise between different samples. Knowledge of what some of the measurements should be, paired with a correlated noise assumption, enables us to make an intelligent guess at what the noise is for the rest of the measurem­­­ents. This is indeed the scenario with mass spectrometry data, where the knowledge of the true mz values for some peaks comes from identified peptide sequences.

# Calibration

The goal of the calibration process is to shift each peak in the MS and MS/MS spectra by an appropriate amount

In order to decide the

## Data point acquisition

Every peptide sequence identification corresponds to multiple peaks in the spectra. First, the MS/MS scans should ideally include peaks corresponding to all of the b and y ions of the peptide produced by the collision-induced dissociation technique employed in the mass spectrometer. The neighboring MS scans should have evidence of the un-fragmented peptide. Each of the matches corresponds to peaks at different charge states, and different isotopic peaks. All of those have a true mz value, and most of them should have corresponding peaks in the acquired spectra.

For a concrete example, assume that an identification tells us that an MS/MS spectrum corresponds to peptide sequence HVVQSISTQQEKETIAK, identified with a precursor charge 3. Since the sequence contains 17 amino acids, the total number of b and y ions that should be present in the MS/MS spectrum is 32. Each of those ions can have either 1, 2, or 3 charges, so the number of monoisotopic peaks to look for is 96. Every ion-charge state match still corresponds to multiple peaks, since

Describe the differences and similarities with the original calibration paper. Explain how we extended the idea. Now we calibrate the MS2 spectra as well. Show the difference with only calibrating MS1.

# Notch Search

Using mass differences between the observed peptide mass and the theoretical mass is an efficient method for identifying post-translational modifications. A popular approach is to consider the MS/MS spectrum in isolation, and perform a database search that identifies

## Comb Search

A search of the unimod database reveals that known modifications with mass difference within 200 daltons have values that are within [-0.1, 0.2] of every integer. PTM combinations also have this property. This allows us to ignore

## Notch Search

# Results

a. Only spectra files calibration, no additional search done. Show MSE values for different calibration functions, pick one, and say why it’s better than others

b. A new normal Morpheus run comparison: 10ppm, vs 3ppm (4 runs)

i. Show that precursor mass errors are centered at zero and have smaller variance

ii. Show that there are more PTMs identified we are confident it

iii. Show that decoy PTMs are pushed down the list

Compare with 11 cell line paper, and with Jurkat/Mouse paper

c. Notches

i. Show that now are able to discern between PTMs with similar mass errors (ones that are only different because of the mass defect)

d. Show the improvement in search time due to the new notches option