

Instrument Specific Calibration of PeptideProphet

Bryan Prazen¹, Erik Nilsson¹, Brian Pratt¹, Martin Sadilek², Daniel Martin³, John Klimek³, Andrew Gemmill³, Laura Hohmann³, Jennifer Jackson³

¹Insilicos, ² Department of Chemistry University of Washington, ³ Institute for Systems Biology

Review of PeptideProphet

PeptideProphet is a statistical approach to validation of peptide identifications from tandem MS search methods, such as SEQUEST. Employing database search scores, number of tryptic termini, number of missed cleavages, and other information, PeptideProphet computes the probability that the peptide was correctly identified.

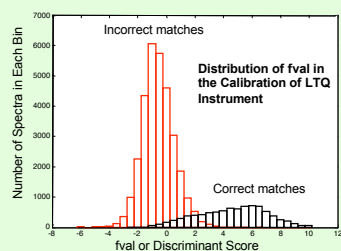
PeptideProphet replaces earlier techniques, where the researcher was forced to take uncalibrated quality scores from peptide search engines, then choose an essentially arbitrary threshold for acceptable "hits." PeptideProphet, in contrast, supplies statistically testable confidence values, and more fully exploits the MS data.

Step 1: Application of canonical discriminant analysis to the scores resulting from a MS/MS search for peptides. For a SEQUEST search these scores are Xcorr, deltaCn and SpRank.

$$fval = c_0 + c_1 \frac{\ln(Xcorr)}{\ln(length)} + c_2 \Delta Cn + c_3 \ln(SpRank)$$

Discriminant Function Coefficients – Coefficients were calibrated using a standard sample mixture in this experiment.

$c_0 = -1.5$
 $c_1 = 9.3$
 $c_2 = 7.3$
 $c_3 = -0.2$



Step 2: Application of Expectation Maximization (EM) to the fval, parent ion mass error between observed and theoretical (delta M) and number of tryptic termini (ntt).

• EM finds maximum likelihood estimates of parameters in a probabilistic models where the model is dependent on unobserved variables.

• EM is an iterative unsupervised learning method

• EM assigns probabilities to peptide matches.

• EM assumes the shape of the correct and incorrect match distributions.

• EM accounts for variation found in fval, nnt, and delta M distributions between experiments

Introduction

PeptideProphet was first calibrated using the ThermoFinnigan LCQ mass spectrometer. For this instrument it was shown that PeptideProphet computed probabilities have more correct matches for any given error rate than the conventional filtering criteria. Yet, little is known about how PeptideProphet performs on other instruments. We evaluate the need for instrument specific calibration of PeptideProphet when using two ThermoFinnigan LCQ instruments, a ThermoFinnigan LTQ instrument and an Applied Biosystems API QSTAR Pulsar Instrument.

Overview

If PeptideProphet is used to estimate the accuracy of peptide assignments and to serve as a common standard to compare peptide identification results across research groups, it is vital that the PeptideProphet model transfer between instruments.

PeptideProphet was applied across three tandem MS platforms and improvements resulting from instrument specific calibration were examined. To test PeptideProphet calibration transfer, results from the instrument generic PeptideProphet calibrations were quantitatively compared to results from the instrument specific PeptideProphet calibrations. Matches to common contaminants and the 18 proteins known to be in the standard were deemed correct, and all others incorrect.

Methods and Materials

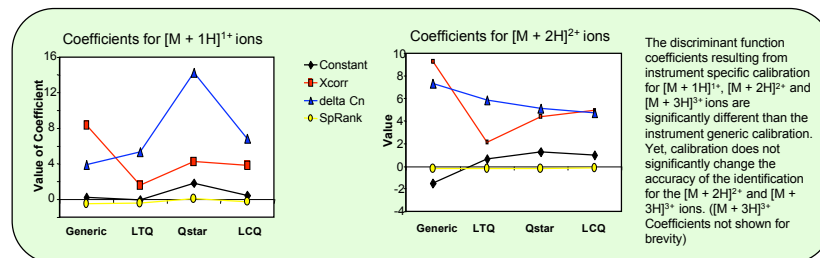
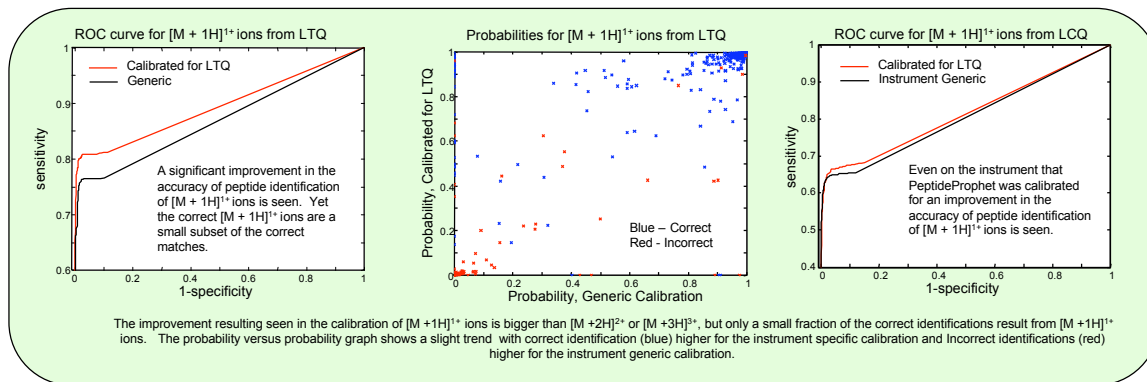
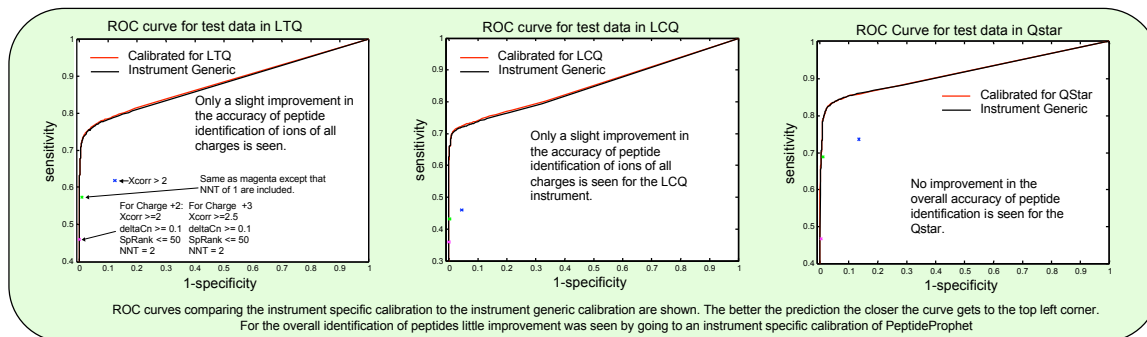
- Standard mixture of 18 proteins.
- Measured 10 times on each instrument (5 for calibration, 5 for test).
- Instruments: ThermoFinnigan LCQ, ThermoFinnigan LTQ and Applied Biosystems API QSTAR Pulsar.
- Enzyme un-constrained SEQUEST search of data base made from standard proteins and ubiquitous proteins appended to haemophilus influenza peptide database.

Conclusions

• Instrument specific PeptideProphet calibrations for the ThermoFinnigan LTQ instrument and the Applied Biosystems API QSTAR Pulsar Instrument do not significantly perform the instrument generic calibration that was done with the ThermoFinnigan LCQ.

• The Instrument specific PeptideProphet calibration for [M + 1H]⁺ ions does not perform the instrument generic calibration, for the ThermoFinnigan LTQ and LCQ. The instrument generic calibration was done with a LCQ. This implies that calibration for [M + 1H]⁺ ions might be more susceptible to experimental variation than previously thought.

Comparison of Results form Instrument Specific Calibration and Instrument Generic Calibration



References

PeptideProphet: Keller A, Nesvizhskii AI, Kolker E, Aebersold R "Empirical Statistical Model to Estimate the Accuracy of Peptide Identifications Made by MS/MS and Database Search" Analytical Chemistry 2003, 75, 4646-4658.

Acknowledgments

We gratefully acknowledge Alexey Nesvizhskii of University of Michigan Department of Pathology for his insight and recommendations and recommendation regarding this project. This work was funded by NIH grant # HL079807