

# Analysis Platform for Differential Quantitative Proteomics Analysis

Fang R<sup>1</sup>, Fitzgibbon M<sup>1</sup>, May D<sup>1</sup>, Mannova P<sup>1</sup>, Wang H<sup>1</sup>, Eng J<sup>1</sup>, Goodlett D<sup>2</sup>, Hanash S<sup>1</sup>, Beretta L<sup>1</sup>, and McIntosh M<sup>1</sup>

<sup>1</sup>Fred Hutchinson Cancer Research Center, <sup>2</sup>University of Washington, Seattle, WA

## OVERVIEW

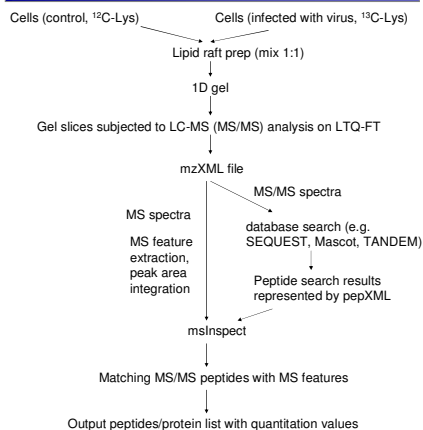
Differential quantitation is an important problem in the field of comparative proteomics and biomarker discovery. For LC-MS based approaches, accurate quantitation requires high performance software tools for the analysis of mass spectra data. We describe the use of our open-source software application called msInspect (1) for differential quantitation analysis using stable isotope labeling e.g. SILAC, and a label-free approach. MS features are extracted using a novel 2D algorithm (1), and MS/MS search results were loaded from pepXML files making this approach compatible with any search engine with an available pepXML converter.

Only a small subset of MS peaks is selected for CID and peptide sequencing in a typical LC-MS/MS analysis. We implemented an accurate mass and time tag approach (AMT) in msInspect, which has been successfully used to improve proteomics coverage (2) by identifying those MS features not selected for CID but matched with peptides identified in relevant LC-MS/MS runs.

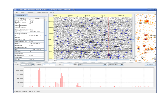
By operating on blocks of MS scans and carefully caching those most recently accessed, msInspect is able to quickly extract features from large high-resolution LC-MS files. For a typical 90 minute LC gradient LTQ-FT data of complex biological sample, msInspect only took 11 minutes to find all the MS features (~ 15,000 feature extraction) on a single processor Dell computer with CPU of 2.8 GHz and 512 MB RAM.

msInspect is freely available and may be downloaded as an executable Java JAR file, or launched via Java Webstart by a single click on the Web at <http://proteomics.thrcr.org/CPAS>.

## Quantitation by SILAC

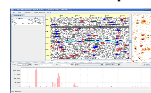


### MS Feature Extraction



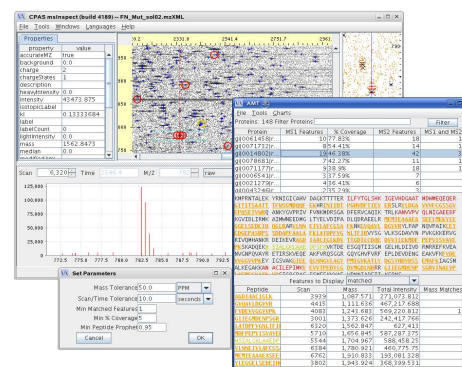
Dots in black are peaks from the raw spectrum file  
Crosses in blue are MS features extracted by msInspect  
Crosses in red are peptides identified by database searching of MS/MS spectra matched with MS features identified by msInspect

### Matching MS Features with MS/MS Peptides



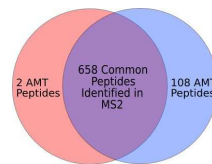
126 proteins with two or more peptide hits of high confidence (PeptideProphet >= 0.95) were found to have been differentially expressed, many of which have been confirmed by Western blot (3).

## AMT Matching



Highlighted in red, peptides found by MS/MS only  
Highlighted in yellow, peptide features found by MS only  
Highlighted in orange, peptide/features found by both MS and MS/MS

### Distribution of Peptide ID by MS/MS and AMT



Circles in red and blue are peptide IDs from replicate 1 and 2 respectively. Overlap area is the common peptides identified by MS/MS in both replicates

In two replicates of LTQ-FT analysis of *Francisella* cell lysates, 110 additional peptides were identified through matching of no-CID MS features found in replicate 1, 2 (R1, R2) with peptides identified with high confidence by MS/MS in replicate 2, 1 (R2, R1) respectively, representing a 17% increase.

## CONCLUSIONS

- msInspect was successfully used for differential quantitation of lipidraft proteomics following infection by HCV virus. Results were confirmed by Western blot.
- AMT approach implemented in msInspect increased peptide identification by 17% in two replicates of LC-MS analyses of *Francisella novicida* cell lysates.

## METHODS

Differential quantitation by SILAC: Normal and HCV-infected human liver cells (Huh7, SFL-3) were cultured with <sup>13</sup>C<sub>6</sub> L-lysine supplemented for SFL-3. Lipid raft proteins were isolated from both cell cultures and mixed with 1:1 and analyzed on SDS gel. Gel slices were digested by trypsin and subjected to LTQ-FT analysis and the spectra data were converted to mzXML. MS/MS spectra were submitted to TANDEM search with plug-in K-score algorithm (5) against human IPI database appended with HCV protein sequences. msInspect was used to analyze the MS spectra and obtain protein/peptide quantitation results.

AMT mapping: *Francisella novicida* U112 cell lysates were digested by trypsin and analyzed by LTQ-FT. Data were analyzed as described above except that the MS/MS spectra were searched using TANDEM (K-score) against *Francisella novicida* ORF database. To match the no-CID MS features in replicate 1 (R1) with potentially corresponding peptides identified by MS/MS in replicate 2 (R2), The elution times of the peptides in R2 were transformed into LC-gradient independent values based on the predicted hydrophobicity index of peptides (4) then converted to times in the R1 gradients using R1 MS/MS IDs as reference. The no-CID MS features in R1 and R2 were then matched with peptides identified by MS/MS with PeptideProphet score >= 0.95 in R2 and R1 respectively using a mass and elution time tolerance window of ± 10 ppm and ± 3.5%, respectively.

## REFERENCES

- Bellew M *et al.*, A suite of algorithms for the comprehensive analysis of complex protein mixtures using high-resolution LC-MS, under revision, *Bioinformatics*, 2006
- Smith RD, *et al.*, An accurate mass tag strategy for quantitative and high-throughput proteome measurements. *Proteomics* 2, 519-529, 2002.
- Mannova P *et al.*, Modification of host lipid raft proteome upon hepatitis C virus replication, under revision, *Molecular & Cellular Proteomics*, 2006.
- Krokhin OV *et al.*, An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC, *Mol Cell Proteomics*, 2004, 3(9), 908-919.
- Maclean B, Eng J, and McIntosh M, An open-source framework for evaluation MS/MS score functions: "plugable scoring" in X! Tandem. *ASMS2006 poster TP655*

This work was funded by NCI subcontract 23XS144A.