Absolute Protein Quantitation on an 'Omic Scale: Application to Cancer Biology

J. Will Thompson[1], Scott Geromanos[2], Wenle Xia[1], Neil Spector[1], M. Arthur Moseley[1]

[1]School of Medicine, Duke University, Durham, NC, [2]Waters Corp., Milford, MA

ABSTRACT

For decades, the field of differential expression proteomics has been capable of only providing relative quantitation information. A diverse array of technologies have been developed and applied to differential expression proteomic studies, from the initial 2D gels to modern LC/MS (isotope labeled and label free) methods. However, the biological impacts of these relative quantitation methods have been limited in that each experiment provides information relevant only to that specific experiment.

Relative and absolute quantitation information on all identified proteins greatly enhances the utility of any proteomics experiment. This unique ability provides a more complete understanding of the biology. Moreover the absolute quantitation information obtained on an 'omic scale allows the results from multiple experiments between different biological conditions, obtained at different times, and from different laboratories to be compared.

In 2005 Silva, et al. reported a method for absolute quantitation based on the empirically observed fact that the average LC/MS response for the three most intense tryptic peptides per mole of protein is constant (+/- 10 %) across a widely diverse array of proteins. Spiking a known amount of an internal standard protein into the sample permits the calculation of a universal signal response factor, if the LC/MS data acquisition method records accurate abundance information.

This work describes our initial efforts towards the creation of a proteomics database of absolute quantitation information for nuclear and cytoplasmic compartments of cancer cell lines. This database will addresses breast cancer and non-small cell lung cancer cell lines, specifically including those whose genotypic differences lead to refractory responses to drug therapies.

PRINCIPAL GOALS

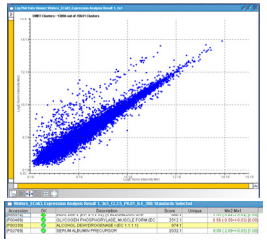
- **Compare data dependant and data independent modes of data acquisition**
  - Proteome coverage
  - Peptide coverage
- **Compare Relative and Absolute Modes of Protein Quantitation**
  - Accuracy
  - Precision

EXPERIMENTAL

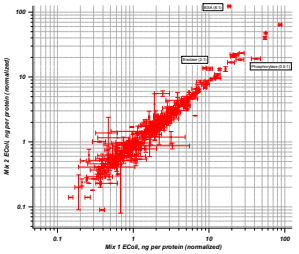
- **Protein Spiking Experiment**
  - E. coli cell lysate digest and standard protein digests were obtained commercially
    - Lysate divided into two aliquots, differentially spiked with standard protein digests
      - Serum Albumin (8:1)
      - Enolase (2:1)
      - Glycogen Phosphorylase (0.5:1)
      - Alcohol Dehydrogenase (1:1)
- **Cell Line Preparation**
  - BT-474 cells were cultured in RPMI media 1640 supplemented with or w/o drug. Whole cell extracts were prepared by scraping cells off petri dishes, washing the cell pellet twice in phosphate buffered saline (PBS), and then re-suspending the pellet in three-packed-cell volumes of 0.1% RapiGest buffer in 50 mM ammonium bicarbonate
- **UPLC system (NanoAcuity, Waters)**
  - 75 um x 250 mm column packed with 1.7 um BEH particles
  - 0.4 uL minute flow rate
  - 3% to 35% acetonitrile in water (0.1% formic acid in each)
  - Column temperature = 55C
- **Hybrid Quadrupole Time-of-Flight Tandem Mass Spectrometer (Q-Tof Premier, Waters)**
  - NanoLockSpray Source (5 ppm mass accuracy)
  - V-Mode (R=9,500)
  - Data Dependant Acquisition
    - 1 second survey scan (50-1990 m/z)
    - up to three targeted MS/MS acquisitions per survey scan
    - collision energy selected based on m/z and charge state
  - Data Independent Acquisition (MSE)
    - 1 second low energy scan (50 - 1,990 m/z)
    - 1 second high energy scan (50 – 1,990 m/z)
    - ramped collision energy
- **Data Processing**
  - Data Dependant Acquisition
    - -Mascot Distiller (Matrix Science)
    - Mascot Server (Matrix Science)
  - Data Independent Acquisition
    - ProteinLynx Global Server 2.3 (Waters)
    - Identity (Waters)
  - Data Correlation
    - Scaffold 2.0 beta (Proteome Software)

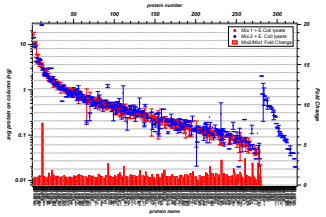
PROTEIN SPIKING STUDY

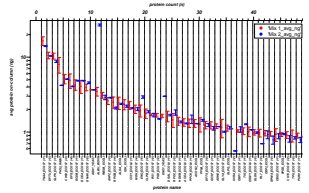
Relative Quantitation – Peptide Level



Absolute Quantitation – Protein Level



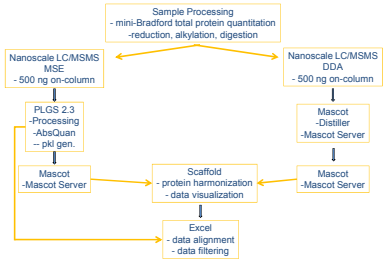




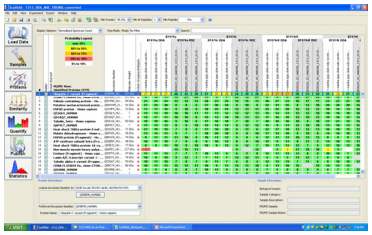
BT474 Cancer Cell Experimental Design

- Four treatment conditions
  - Native (alone)
  - Drug treated
  - Drug resistant
  - Drug resistant removed
- Samples analyzed by Data Dependant and Data Independent Analyses
  - Triplicate analyses by each method
  - Qualitative identification comparison
  - Absolute quantitation using Data Independent Analysis

Data Workflow



Scaffold for Protein Harmonization and Data Visualization

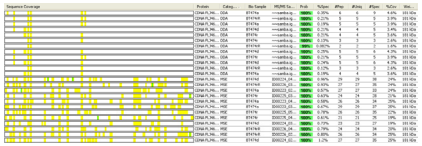


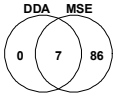
Data Dependant (DDA) and Data Independent (MSE) Acquisition Metrics

-Comparison of coverage at the proteome and peptide level

Number of Unique Proteins

Number of Unique Peptides



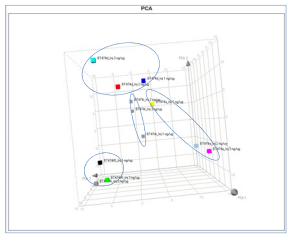


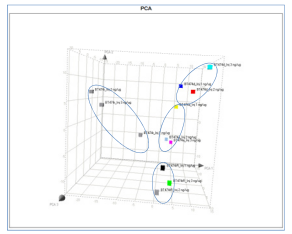
CANCER CELL LINE STUDY

Replication Metrics for Absolute Quantitation

Replication Rate	Protein Count	CV (%)		
3o3	209	13.0		
2o3	68	21.6		
1o3	148			
3 of 3 Replication		**Count (%)**	**Cumulative Count (%)**	**% ng on-col**
Count <10%	128	61	61	66
Count >10%<20%	39	19	80	14
Count >20%<30%	10	5	85	1
Count >30%<40%	22	11	95	4
Count >40%+	10	5	100	2
2 of 3 Replication				
Count <10%	22	32	32	1
Count >10%<20%	23	34	66	1
Count >20%<30%	4	6	72	1
Count >30%<40%	6	9	81	0
Count >40%+	13	19	100	1

Separation of Samples by Treatment Condition Using Absolute Quantitation

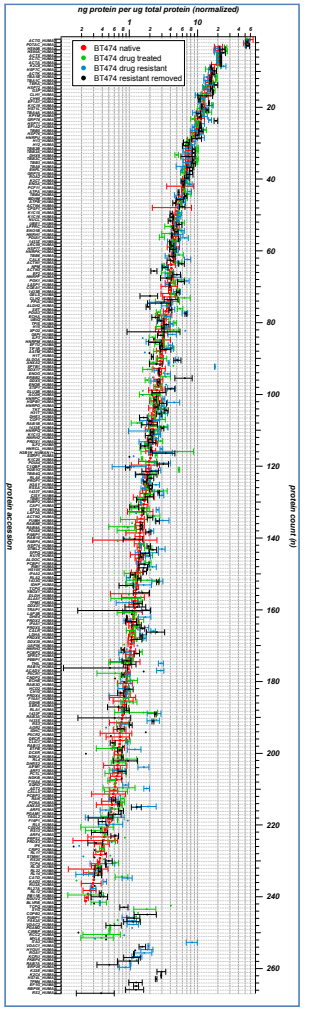




Conclusions

- Qualitative analyses comparison DDA and MSE
  - MSE provides significant increases in
    - Proteome coverage - 469 vs 278 proteins
    - Peptide coverage – 5,658 vs 1,017 peptides
- Both relative quantitation and absolute quantitation
  - Provide good precision and accuracy
    - Absolute Quantitation
      - 209 proteins with < 13% CV (3 of 3 analyses)
      - 68 proteins with < 22% CV (2 of 3 analyses)
- Absolute quantitation provides class separation between different treatment conditions
- Absolute quantitation will facilitate creation of databases of results which can mined across experiments and between labs

Proteome Coverage by Absolute Quantitation -2.5 orders of magnitude dynamic range



Proteomics Core Facility for the Duke University School of Medicine Institute for Genome Sciences & Policy

www.genome.duke.edu/proteomics/

720002503EN