

Characterization of Post-Translational Modifications Using Counting Approaches

Jean L. Spencer, Vivek N. Bhatia, Amanuel Kehasse, Stephen A. Whelan, Christian F. Heckendorf, Catherine E. Costello, and Mark E. McComb

Boston University School of Medicine, Boston, MA 02118

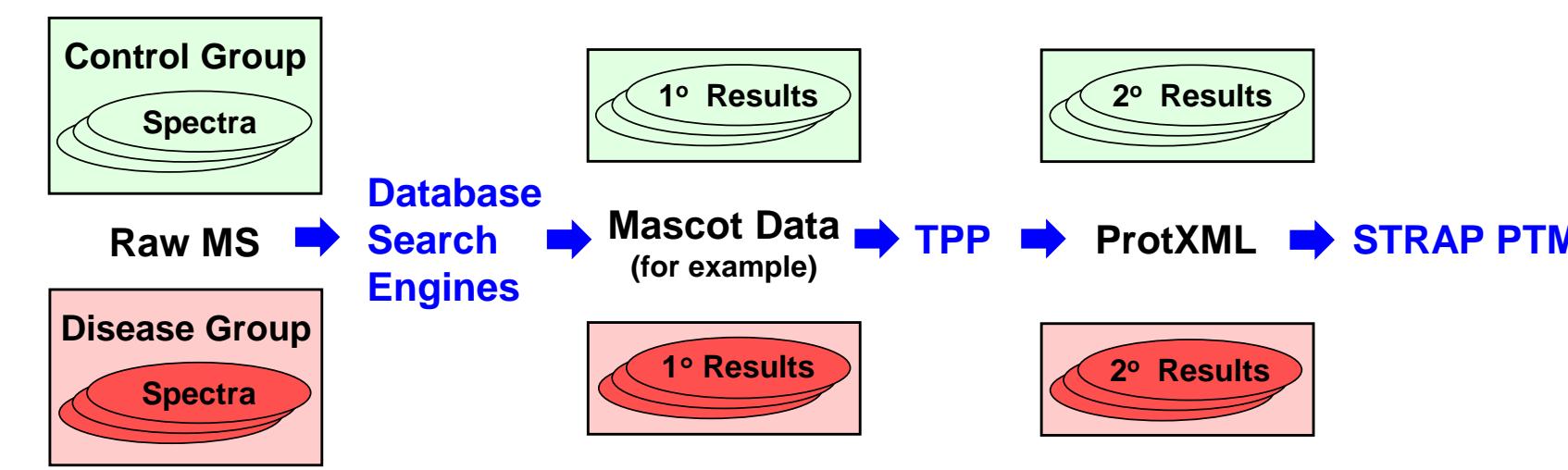
Introduction

There is an increasing interest in the identification and characterization of protein post-translational modifications (PTMs). The vast amount of information obtained within a typical differential proteomics study makes the measure of PTMs challenging. We have begun to explore counting methods for differential analysis of PTM changes in proteomes and have developed a software program to aid in this approach. The **Software Tool for Rapid Annotation of Proteins: Post-Translation Modification edition (STRAP PTM)** uses a new counting-based PTM scoring algorithm to facilitate multi-sample PTM comparison through collation and visualization. The application imports parsed prot.XML file data and ranks peptides based on differential PTMs. Here we show the utility of STRAP PTM across data sets with increasing complexity and compare the results with label-free analysis using Progenesis LC-MS (NonLinear Dynamics).

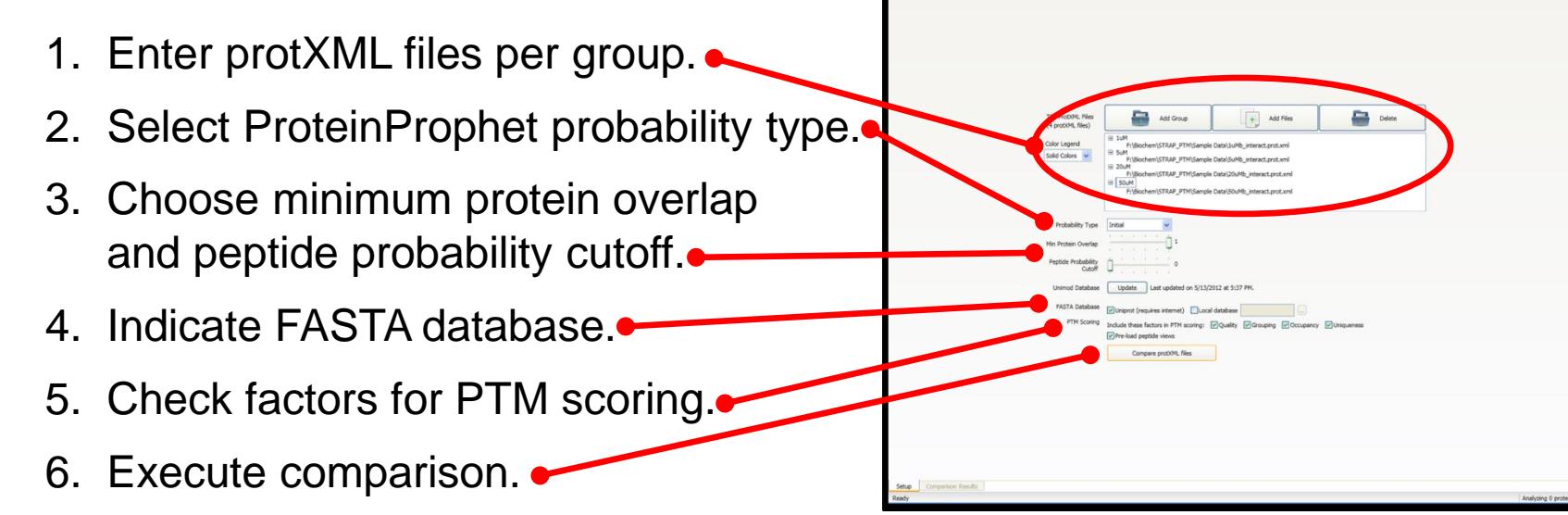
Methods

STRAP PTM: Counting and Scoring

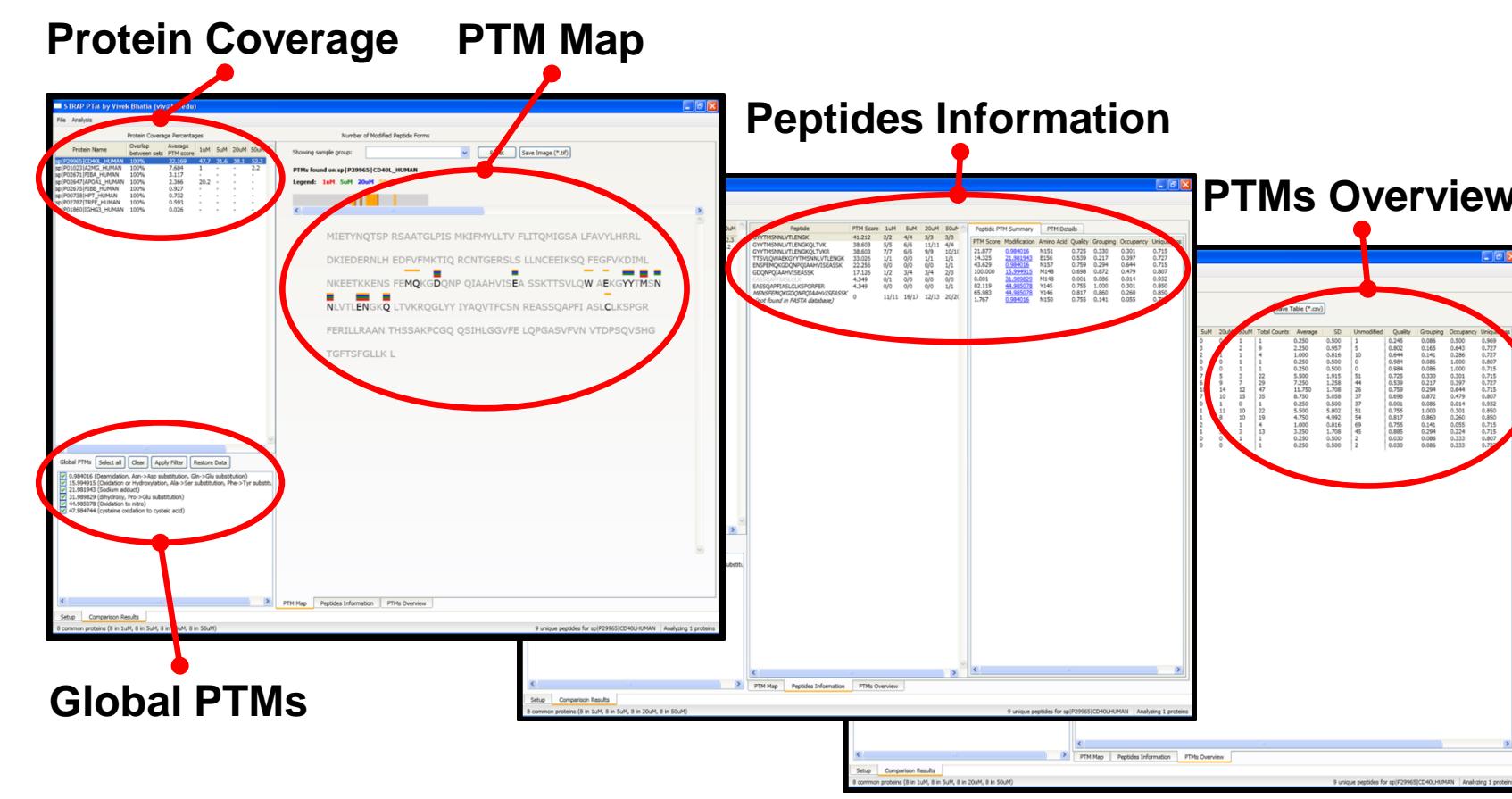
Workflow



Setup Window



Results Windows

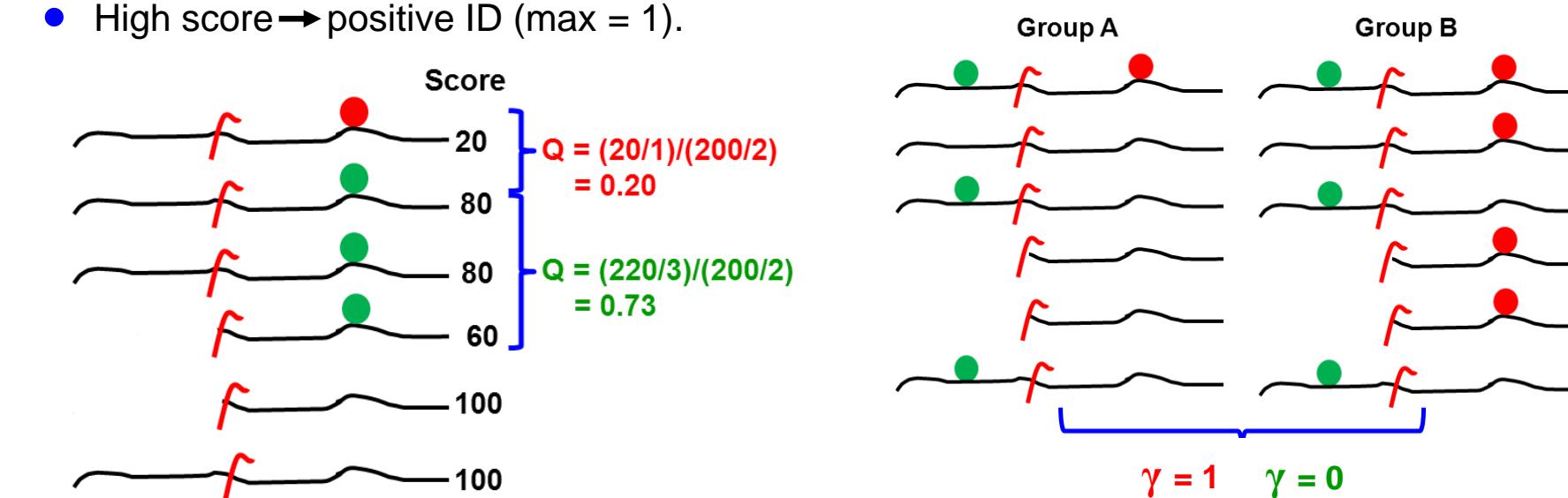


Quality

$$Q_{mip} = \frac{(B_{mip})}{(B^0_{mip})}$$

- Quality of database search based on MS/MS ID scores (B).

- Ratio: (avg score of peptides with specific PTM on specific site) to (avg score of unmodified peptides with specific site).
- High score → positive ID (max = 1).



Grouping

$$Y_{mip} = \frac{\sigma_{mip}}{\max \sigma}$$

- Standard deviation (σ) of specific PTM on specific site across groups; normalized by highest σ across proteins.

- High score → group specificity for specific PTM and site (max = 1).

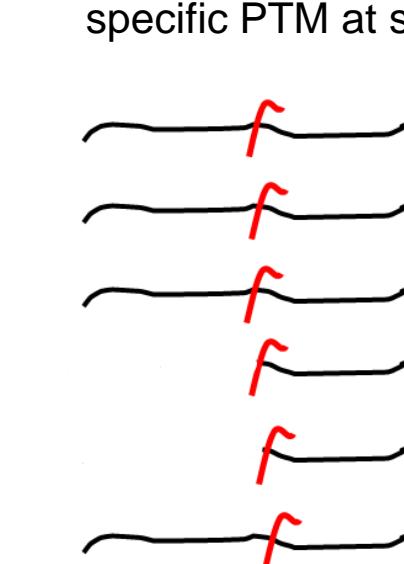
Occupancy

$$W_{mip} = \frac{N_{mip}}{(\sum_{i=1}^M N_{mip}) + (N^0_{ip})}$$

- Degree of modification of specific site with specific PTM on specific protein.

- Ratio of (specific PTM on specific site) to (all PTMs and vacancies on site).

- High score → more susceptible to specific PTM at site (max = 1).



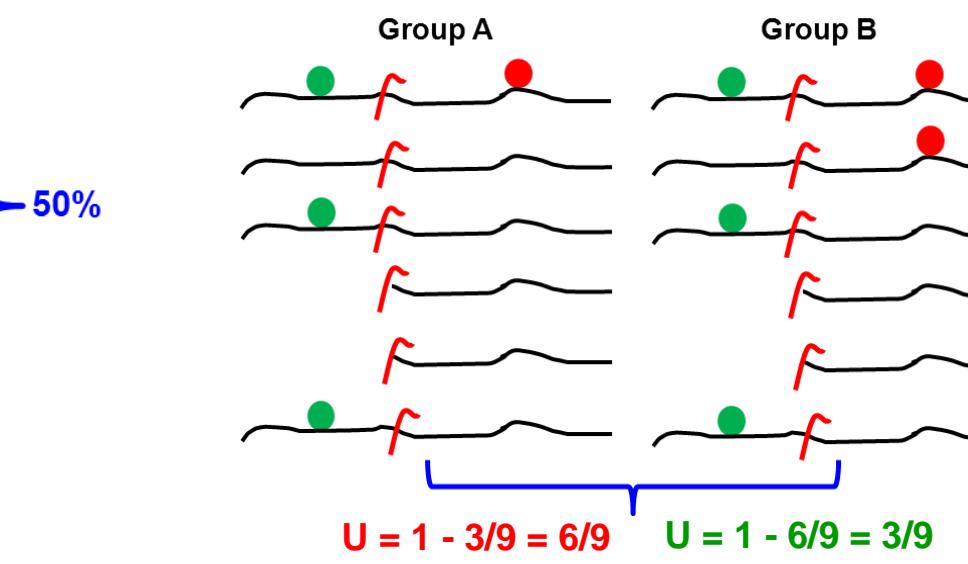
Uniqueness

$$U_{mp} = 1 - \frac{\sum_{i=1}^I N_{mip}}{\sum_{i=1}^M \sum_{j=1}^I N_{mip}}$$

- Rarity of specific PTM on specific protein.

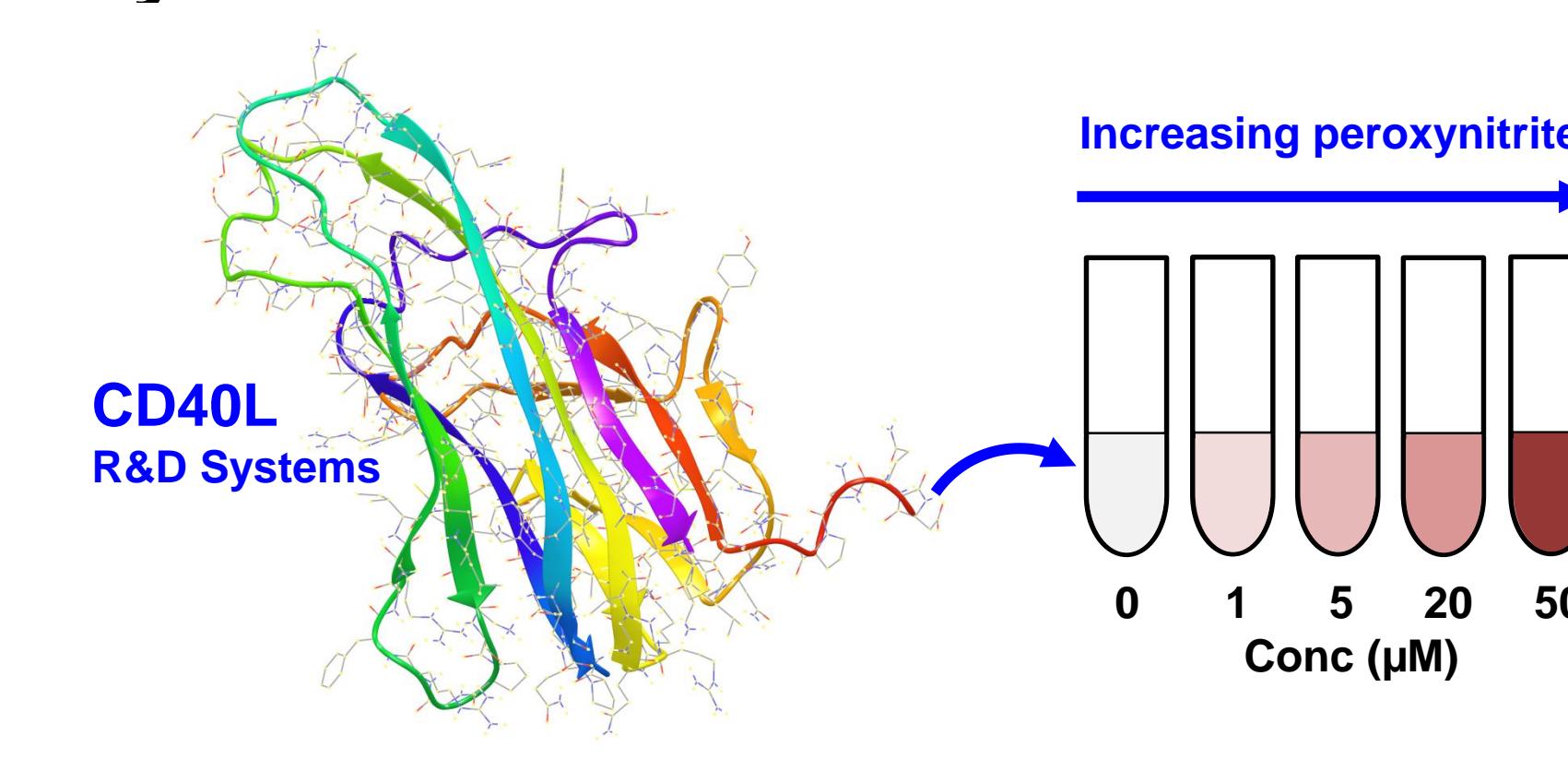
- Ratio of (specific PTM on all sites) to (all PTMs on all sites) subtracted from 1.

- High score → less enrichment of specific PTM (max = 1).



Results

System 1: Oxidation of CD40L



PTM Map

Legend: 1 μM 5 μM 20 μM 50 μM

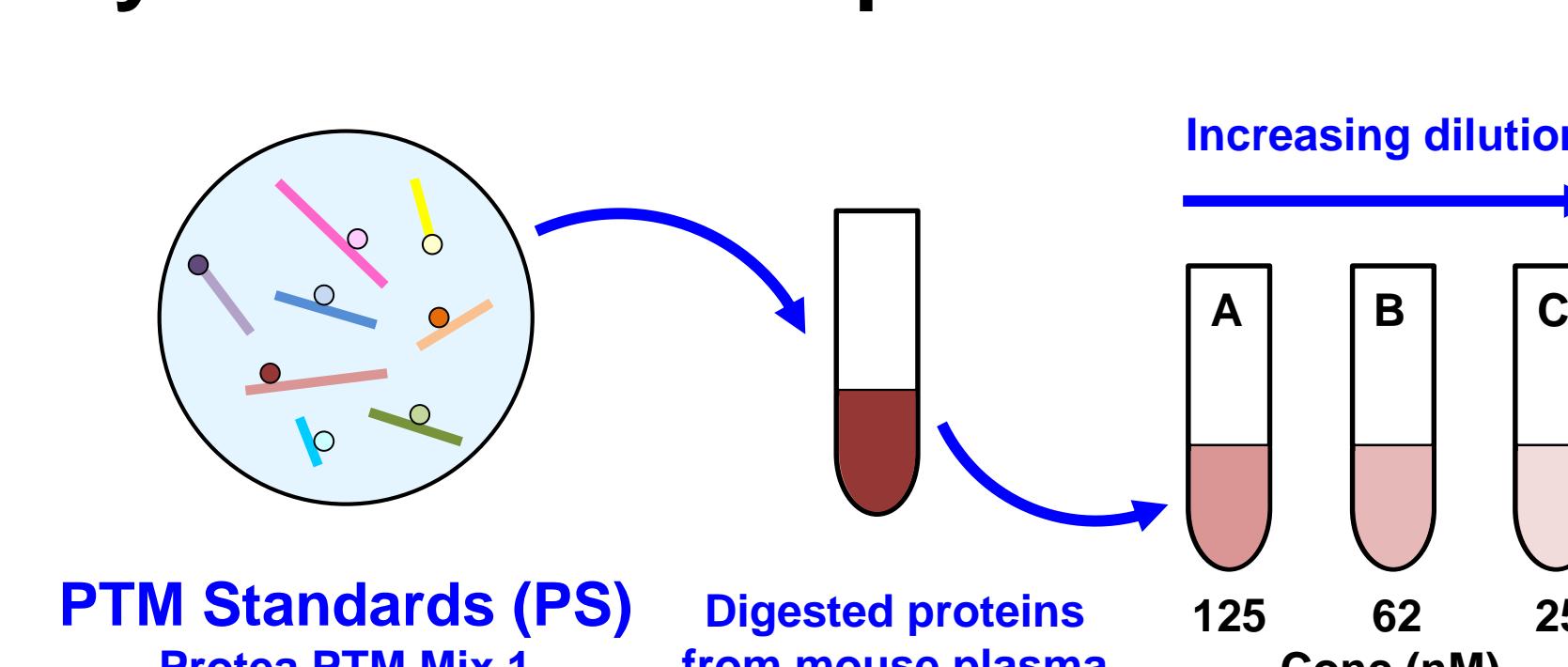
MIETYNTSP RSAATGLPIS MKIFMYLLTV FLITQMIGSA
LFAVYLHRLR DKIEDERNLH EDFVFMKTIQ RCNTGERSLS
LLNCEEIKSQ FEGFVKDIMAL NKEETKKENS FE**M**MQKG**D**QNP
QIAAHVISEA SSKTTSVLQ**W** AEKG**Y**YTMSN **N**LVTLENG**K**
LTVKRQGLYY IYAQVTFCSN REASSQAPFI ASL**C**LKSPGR
FERILLRAAN THSSAKPCGQ QSIHLGG**F**E LQP**G**ASVFVN
VTDPSQVSHG TGFTSFLKL**K** **(Y145, Y146) NO2** **(C194) SO3H**

PTM Scoring

Peptide	Position	AA	Mod Mass	Conc (μM)	Counts	Total	Avg	SD	Unmodified	Forms	Q	Y	W	U	PTM Score		
PS_321	4	Y	44.985078	0 1 11 10	8	5	3		16	5.33	2.52	0	0.899	1.000	1.000	0.885	100.0
PS_532	4	K	42.010865	0 1 8 10	5	1.67	2.08		0	0.899	0.827	1.000	0.617	57.7			
PS_500*	10	S	203.079373	4 3 0	7	2.33	2.08		12	0.899	0.827	0.368	0.821	28.3			
PS_280	4	M	15.994915	0 1 8 10	23	7.67	0.58		0	0.899	0.229	1.000	0.839	21.8			

Results

System 2: PTM Peptides in Plasma



PTM Map

Legend: A (125 nM) B (62 nM) C (25 nM)

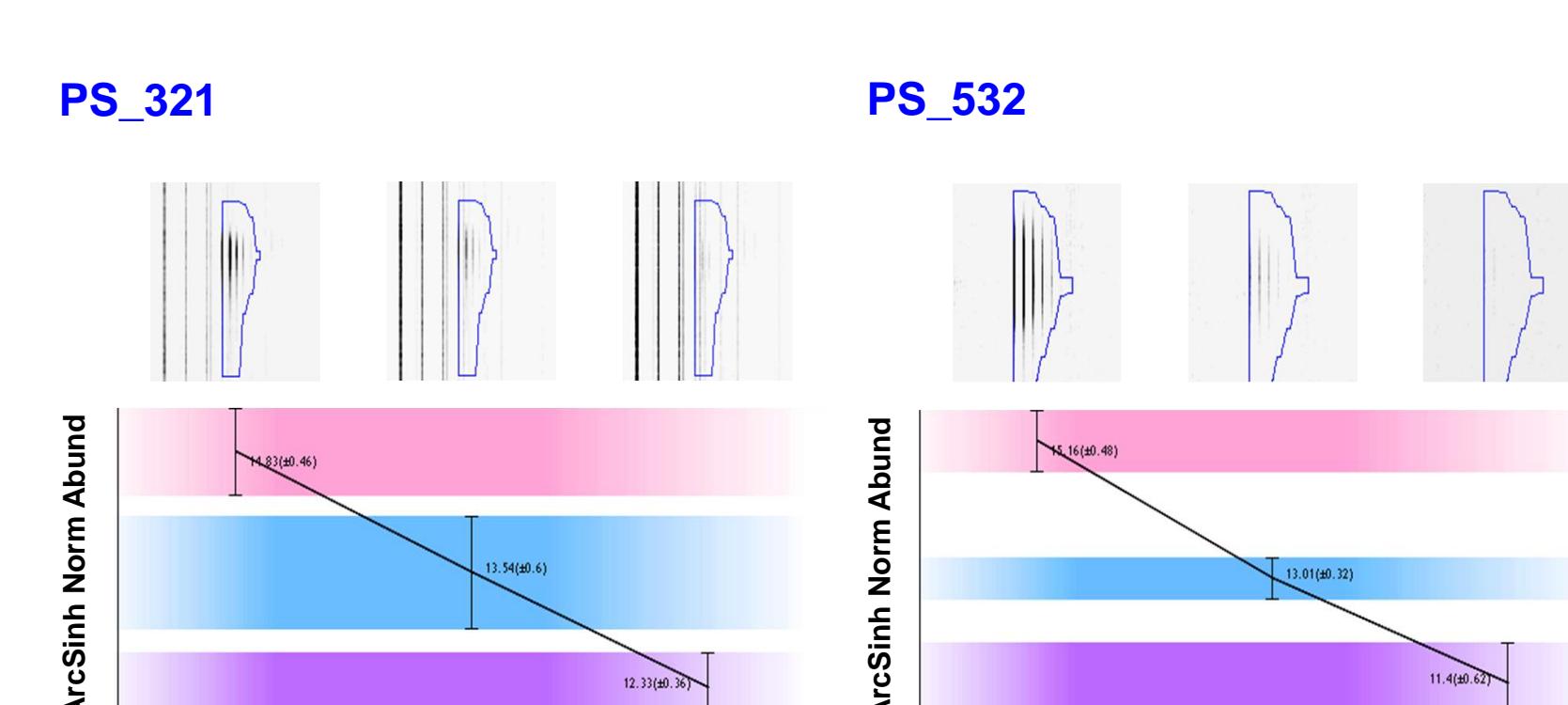
PS_280 SYSMEHFRWG (M4) O
PS_321 DRVYIHPFL (Y4) NO2
PS_500 EAISPPDAAS AAPLR (S4, S10) GalNAc
PS_532 DFNKFHTFPQ TAIVG (K4) Ac

PTM Scoring

Peptide	Position	AA	Mod Mass	Treatment	Ctrl	ATP	EGF	Counts	Total	Avg	SD	Unmodified	Forms	Q	Y	W	U	PTM Score
987	M	15.994915	0	4	6	10	3.33	3.06	7	0.401	0.661	0.588	0.920	56.0				
1172	Y	79.966331	0	1	2	3	1.00	1.00	0	0.878	0.217	1.000	0.265	19.6				
998	Y	79.966331	0	8	8	16	5.33	4.62	16	0.369	1.000	0.500	0.285	18.0				
1110	Y	79.966331	0	1	0	1	0.33	0.58	2	0.983	0.125	0.333	0.285	4.2				
1104	S	79.966331	0	1	1	2	0.67	0.58	1	0.109	0.125	0.667	0.265	0.9				
993	T	79.966331	0	2	0	2	0.67	1.16	22	0.267	0.250	0.083	0.265	0.6				
995	S	79.966331	0	2	2	4	1.33	1.16	21	0.083	0.250	0.160	0.265	0.4				
991	S	79.966331	0	0	2	2	0.67	1.16	23	0.166	0.250	0.080	0.265	0.4				
1197	Y	79.966331	1	1	1	3	1.00	0.00	0	0.862	0.000	1.000	0.265	0.0				

Label-Free Comparison

Progenesis LC-MS



Summary of Peptide Results

Peptide	Confidence Score	Anova (p) x 10 ⁻³	Max Fold Change
PS_321	22.9	58.7	12.9
PS_532	17.7	1.98	41.7
PS_500	37.5	5.84	11.2
PS_280	80.3	161	11.0

Conclusions

- PTM counting is a powerful technique for rapid semi-quantitative integration of large MS-based proteomics data sets.
- STRAP PTM is a novel counting approach with a new scoring algorithm to rank PTMs in differential proteomics experiments.
- STRAP PTM software is easy to implement on a PC and provides fast turnaround for large data sets.
- Select STRAP PTM results correlate well with label-free results from both simple and complex data sets.

Acknowledgments: Funding was provided by NIH-NCRR grants P41 RR010888/GM104603, S10 RR015942, S10 RR020946, S10 RR025082 and NIH-NHLBI grant R21 HL107993 and contract HHSN268201000031C.