

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

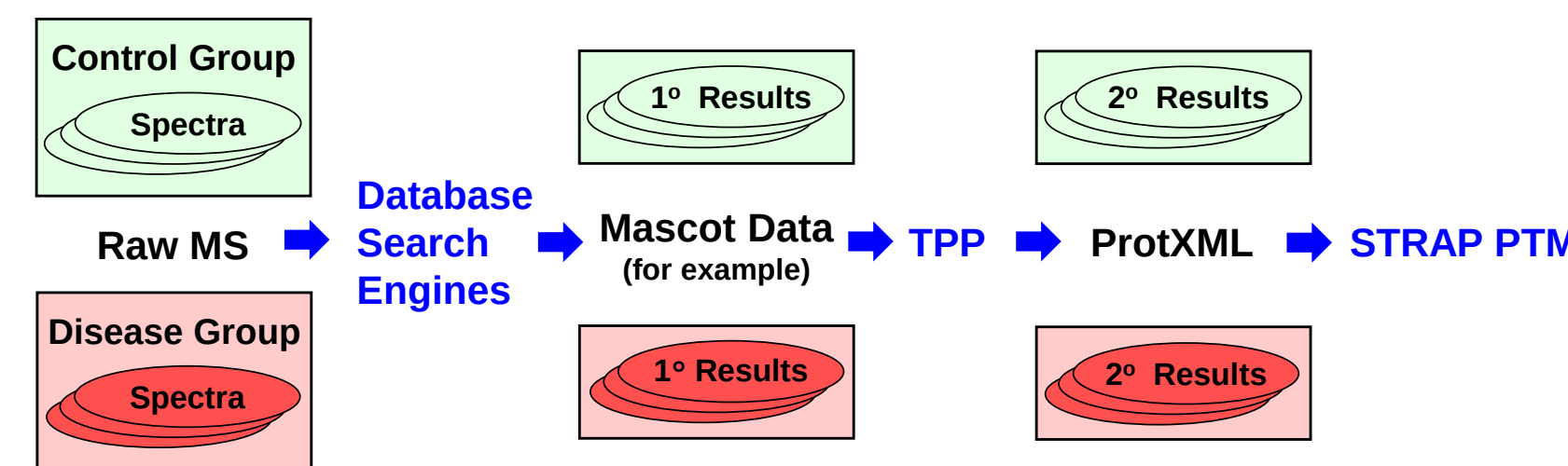
Introduction

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

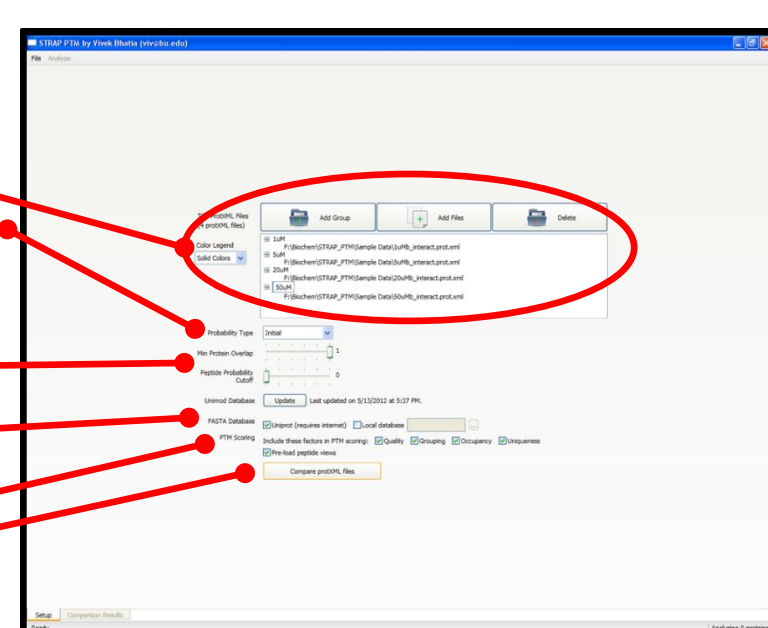
STRAP PTM: Counting and Scoring

Workflow

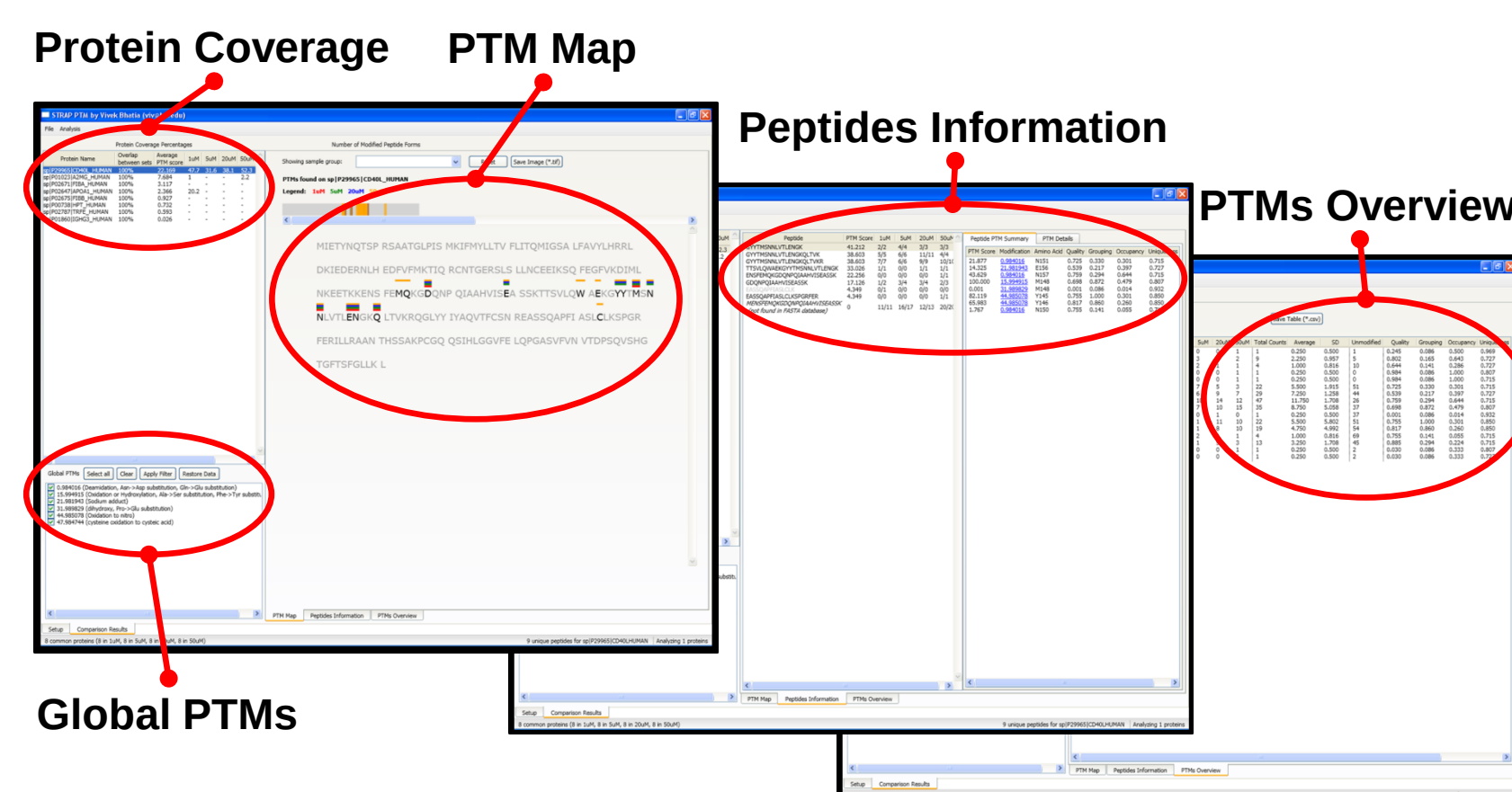


Setup Window

1. Enter protXML files per group.
2. Select ProteinProphet probability type.
3. Choose minimum protein overlap and peptide probability cutoff.
4. Indicate FASTA database.
5. Check factors for PTM scoring.
6. Execute comparison.



Results Windows



PTM Score

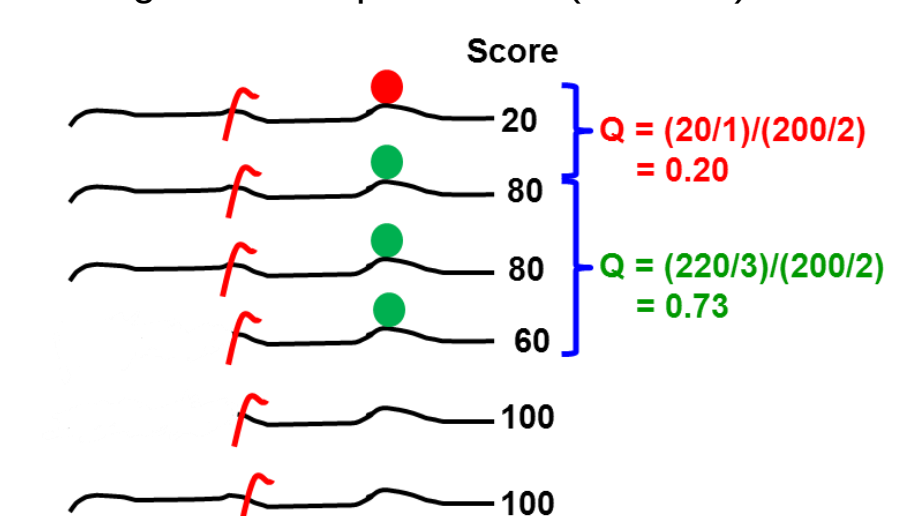
$$S_{mip} = 100 \frac{Q_{mip} Y_{mip} W_{mip} U_{mip}}{\max S}$$

m = specific PTM ($m = 1, \dots, M$).
 i = specific site ($i = 1, \dots, I$).
 p = specific protein ($p = 1, \dots, P$).
 $\max S$ = highest S across proteins.

Quality

$$Q_{mip} = \frac{\langle B_{mip} \rangle}{\langle B^0_{ip} \rangle}$$

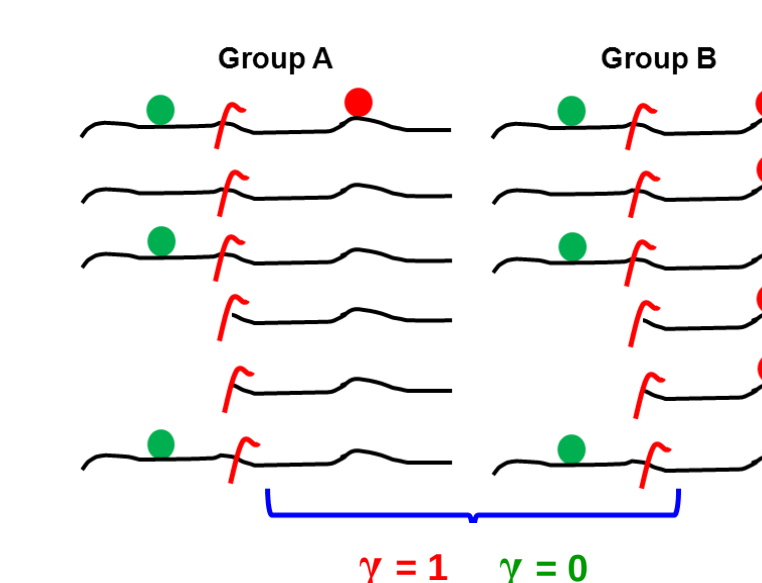
- Quality of database search based on MS/MS ID scores (B).
- Ratio: (avg score of peptides with specific PTM on specific site) to (avg score of unmodified peptides with specific site).
- High score $\rightarrow$ positive ID ($\max = 1$).



Grouping

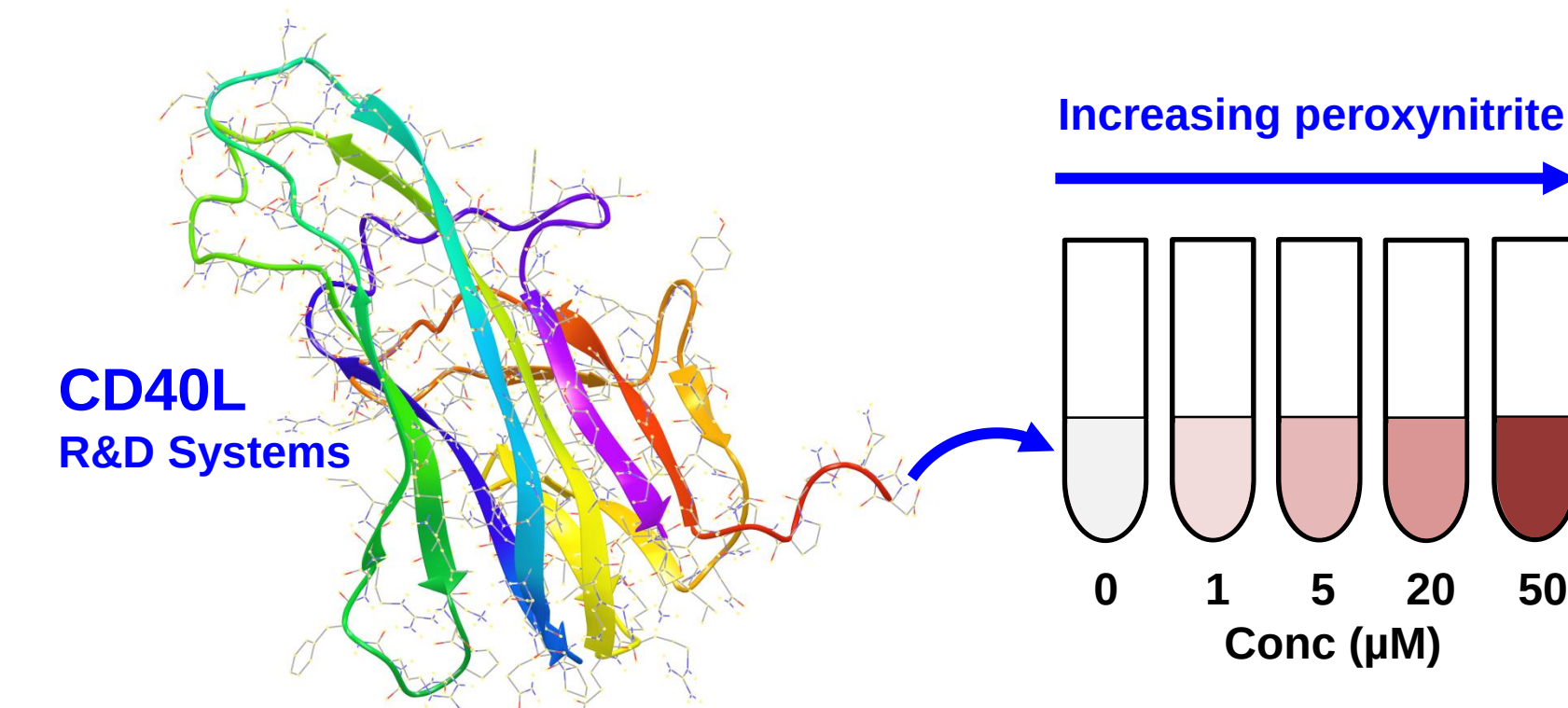
$$Y_{mip} = \frac{\sigma_{mip}}{\max \sigma}$$

- Standard deviation (σ) of specific PTM on specific site across groups; normalized by highest σ across proteins.
- High score $\rightarrow$ group specificity for specific PTM and site ($\max = 1$).



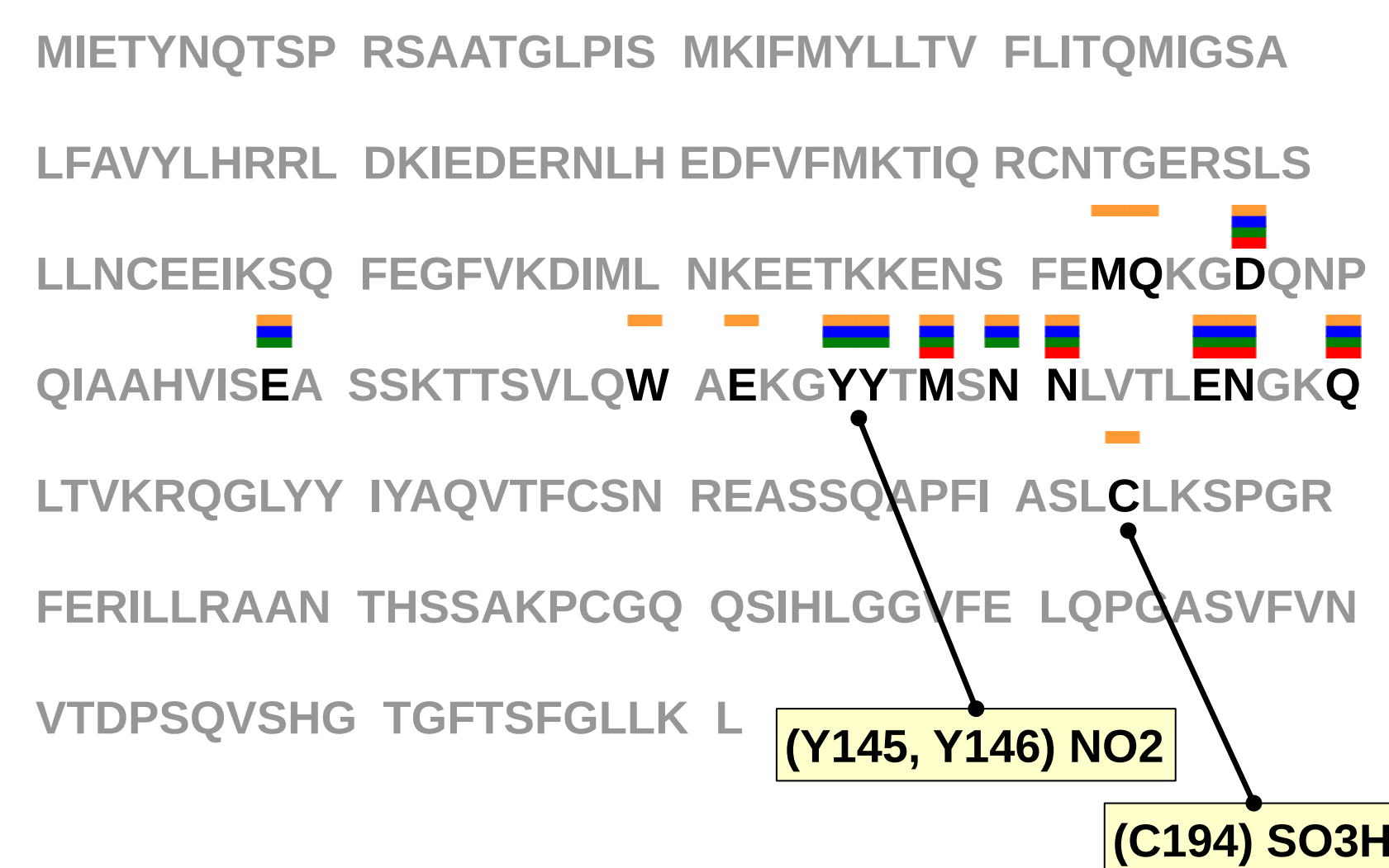
Results

System 1: Oxidation of CD40L



PTM Map

Legend: 1 μM 5 μM 20 μM 50 μM

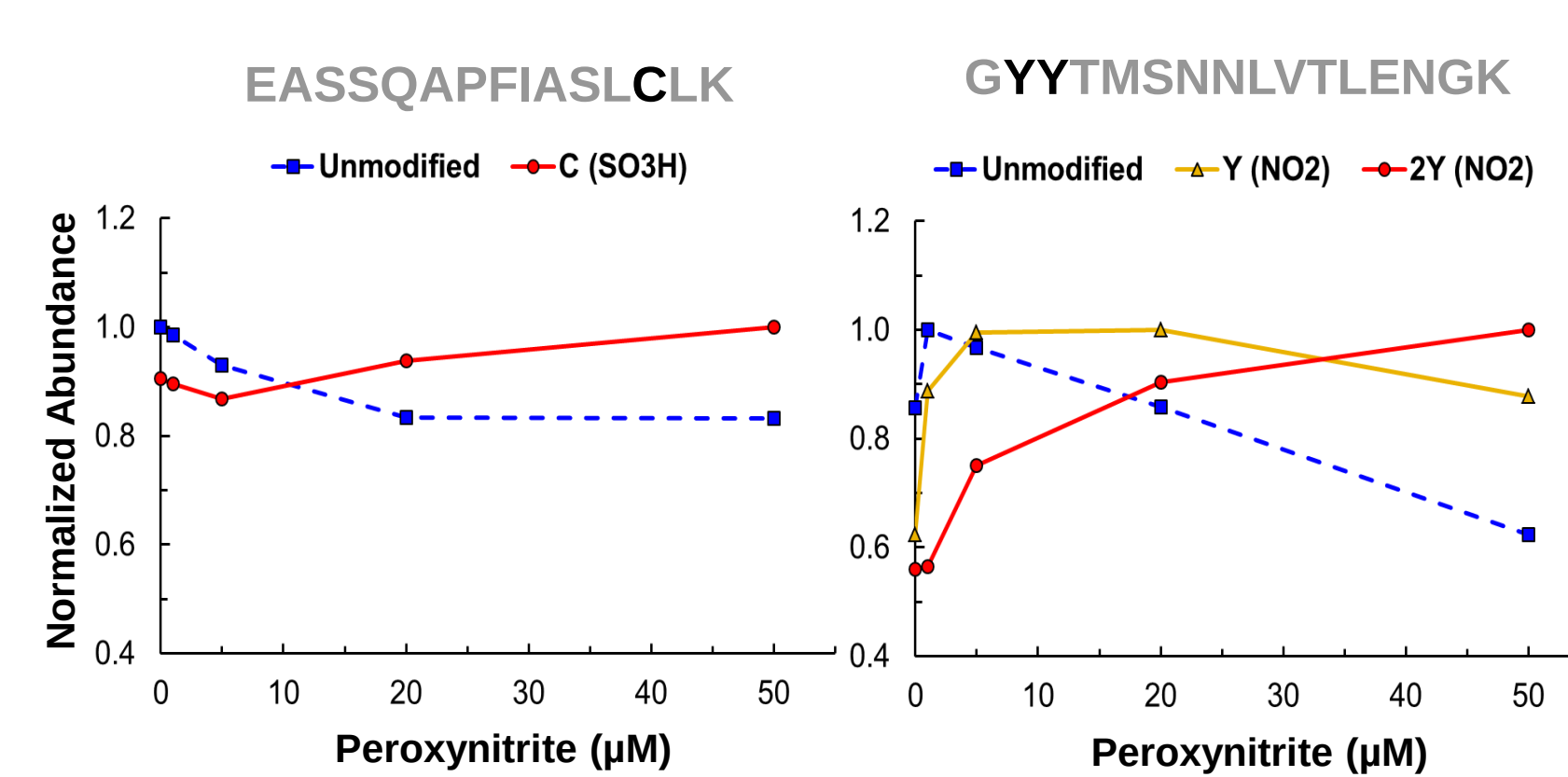


PTM Scoring

Position	AA	Mod Mass	Conc (μM)			Counts			Unmodified Forms	Q	Y	W	U	PTM Score	
			1	5	20	Total	Avg	SD							
148	M	15.994915	3	7	10	15	35	8.75	5.06	37	0.698	0.872	0.479	0.807	100.0
145	Y	44.985078	0	1	11	10	22	5.50	5.80	51	0.755	1.000	0.301	0.850	82.1
146	Y	44.985078	0	1	8	10	19	4.75	4.99	54	0.817	0.860	0.260	0.850	66.0
157	N	0.984016	11	10	14	12	47	11.75	1.71	26	0.759	0.294	0.644	0.715	43.6
113	M	15.994915	0	0	0	1	1	0.25	0.50	0	0.984	0.086	1.000	0.807	29.0
117	D	21.981943	1	3	3	2	9	2.25	0.96	5	0.802	0.165	0.643	0.727	26.3
114	Q	0.984016	0	0	0	1	1	0.25	0.50	0	0.984	0.086	1.000	0.715	25.7
151	N	0.984016	7	7	5	3	22	5.50	1.92	51	0.725	0.330	0.301	0.715	21.9
160	Q	0.984016	4	1	5	3	13	3.25	1.71	45	0.885	0.294	0.224	0.715	17.7
156	E	21.981943	7	6	9	7	29	7.25	1.26	44	0.539	0.217	0.397	0.727	14.3
129	E	21.981943	0	2	1	1	4	1.00	0.82	10	0.644	0.141	0.286	0.727	8.0
194	C	47.984744	0	0	0	1	1	0.25	0.50	1	0.245	0.086	0.500	0.969	4.3
150	N	0.984016	0	2	1	1	4	1.00	0.82	69	0.755	0.141	0.055	0.715	1.8
140	W	15.994915	0	0	0	1	1	0.25	0.50	2	0.030	0.086	0.333	0.807	0.3
142	E	21.981943	0	0	0	1	1	0.25	0.50	2	0.030	0.086	0.333	0.727	0.3
148	M	31.989829	0	0	1	0	1	0.25	0.50	37	0.001	0.086	0.014	0.932	0.0

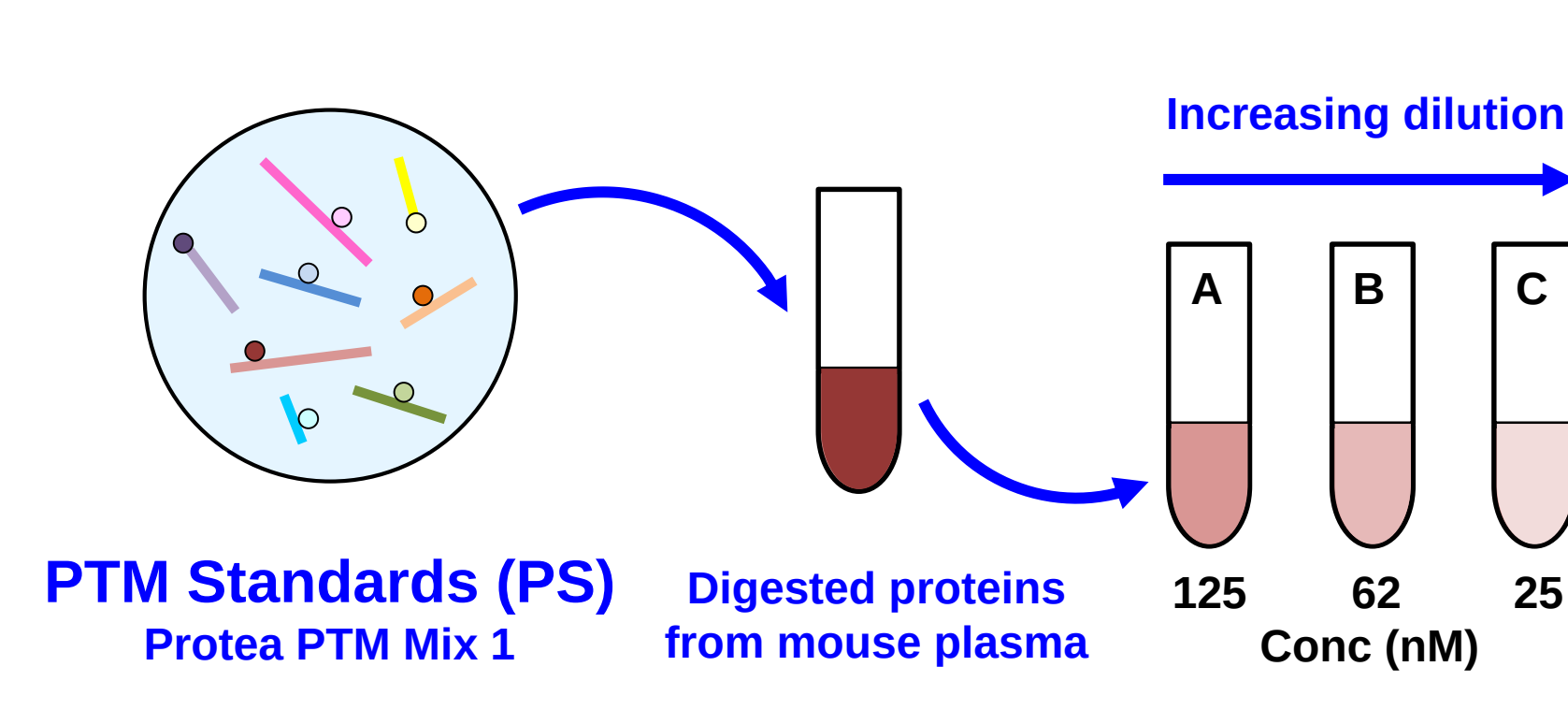
Label-Free Comparison

Progenesis LC-MS



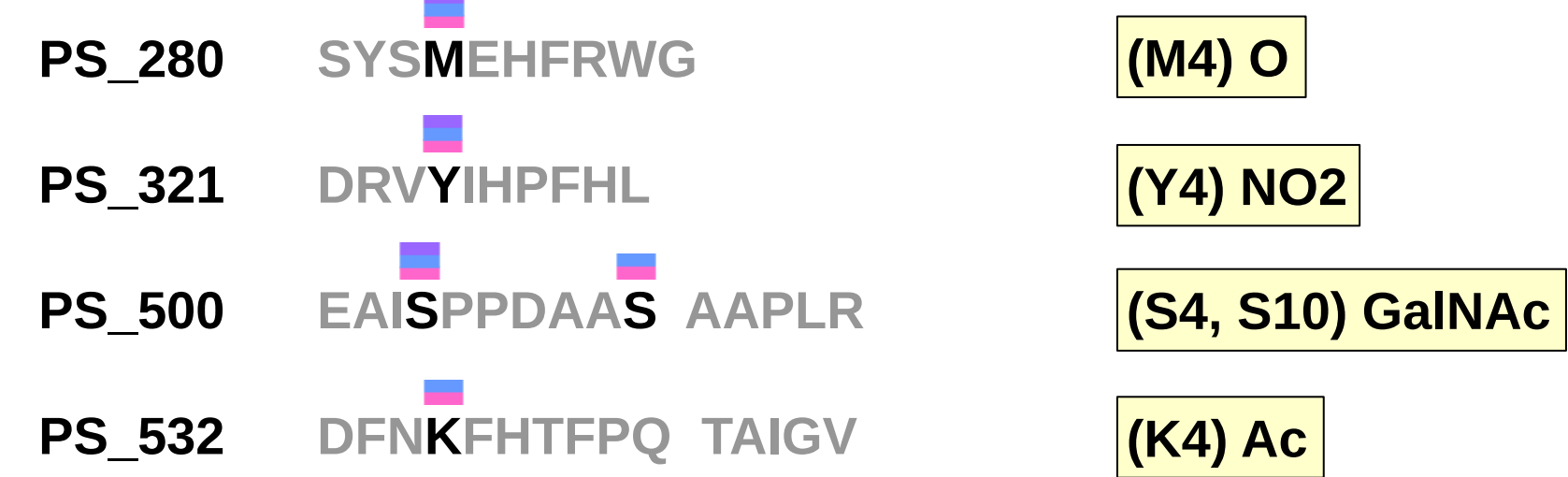
Results

System 2: PTM Peptides in Plasma



PTM Map

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



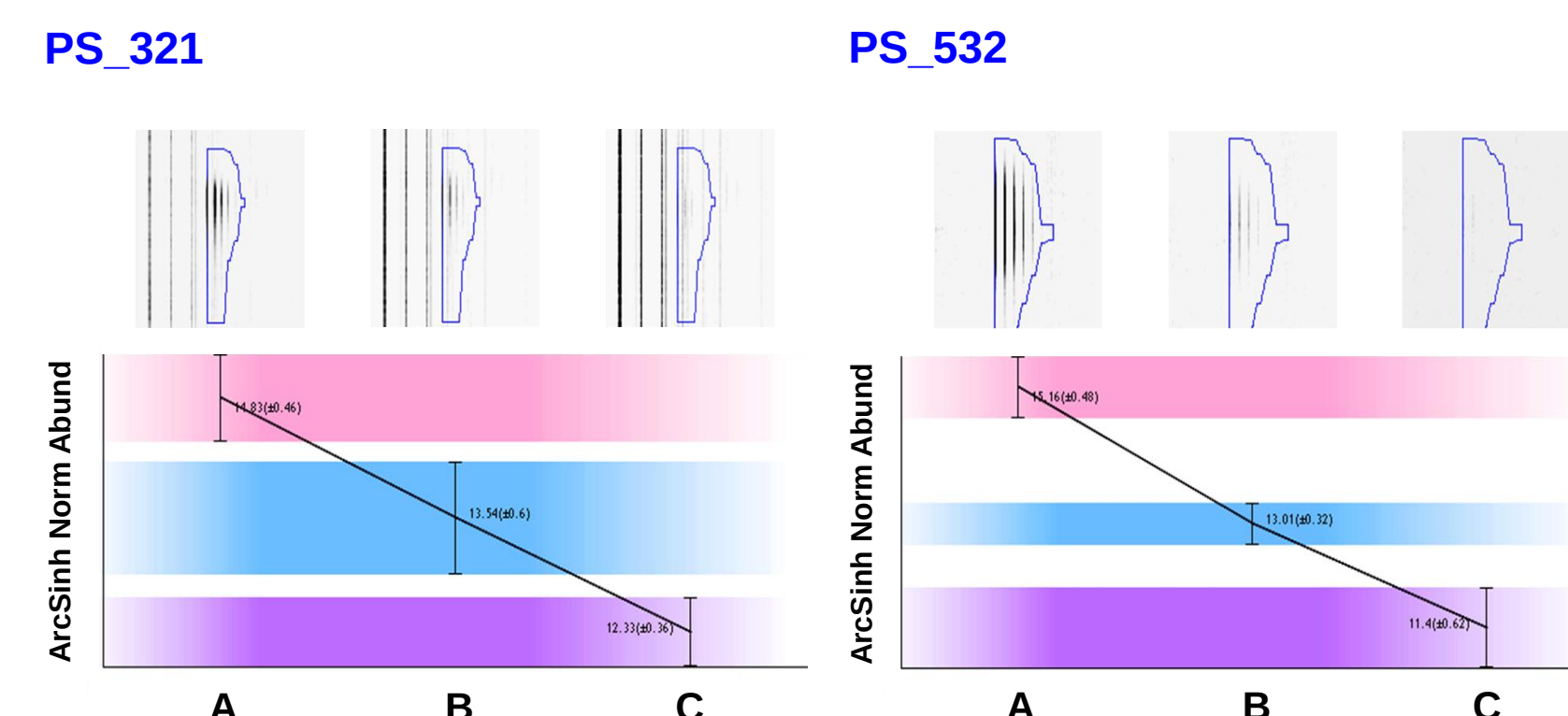
PTM Scoring

Peptide	Position	AA	Mod Mass	A	B	C	Counts Total	Avg	SD	Unmodified Forms	Q	Y	W	U	PTM Score
PS_321	4	Y	44.985078	8	5	3	16	5.33	2.52	0	0.899	1.000	1.000	0.885	100.0
PS_532	4	K	42.010565	4	1	9	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	8	8	7	23	7.67	0.58	0	0.899	0.229	1.000	0.839	21.8

* Peptide PS_500 was also observed as a native mouse plasma peptide and was scored with unmodified forms.

Label-Free Comparison

Progenesis LC-MS



Summary of Peptide Results

Peptide	Confidence Score	Anova (p) $\times 10^7$	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

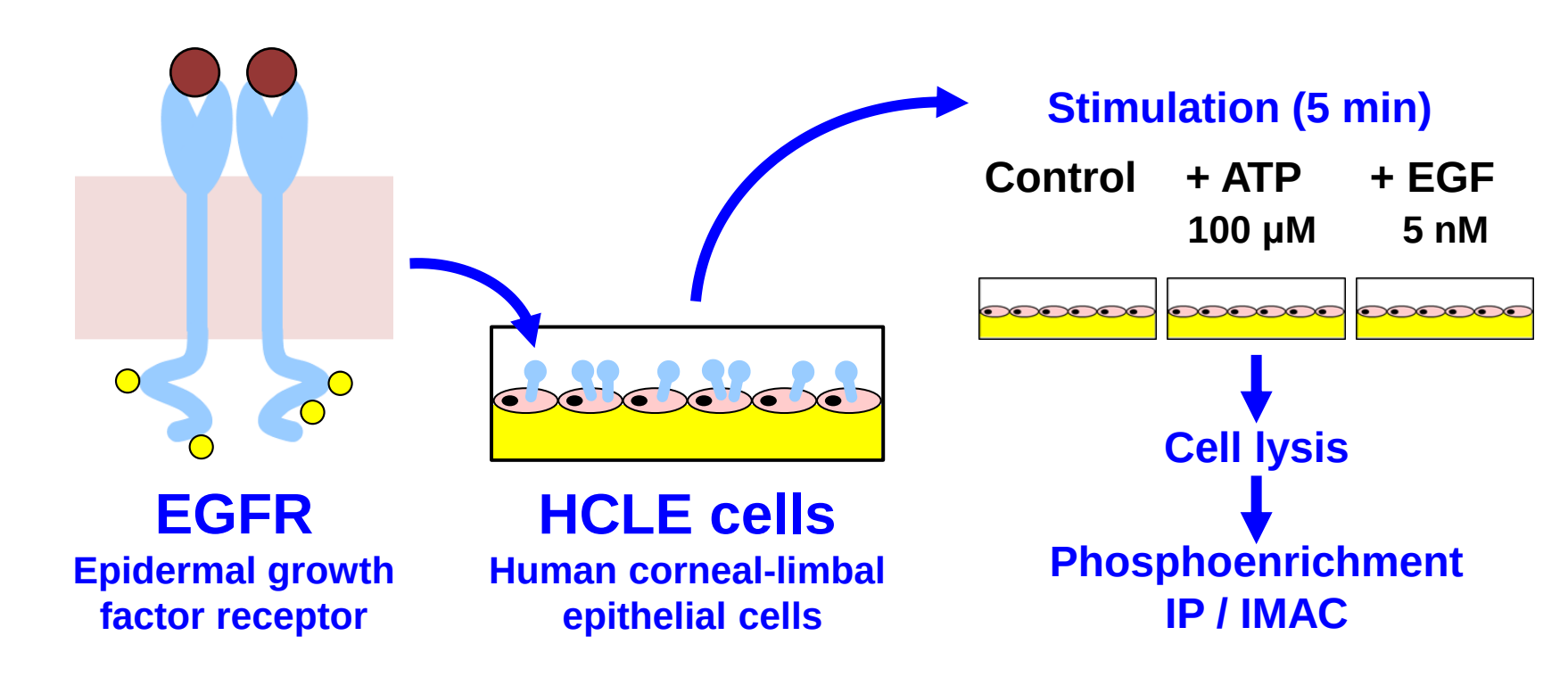
Conclusions

- 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.

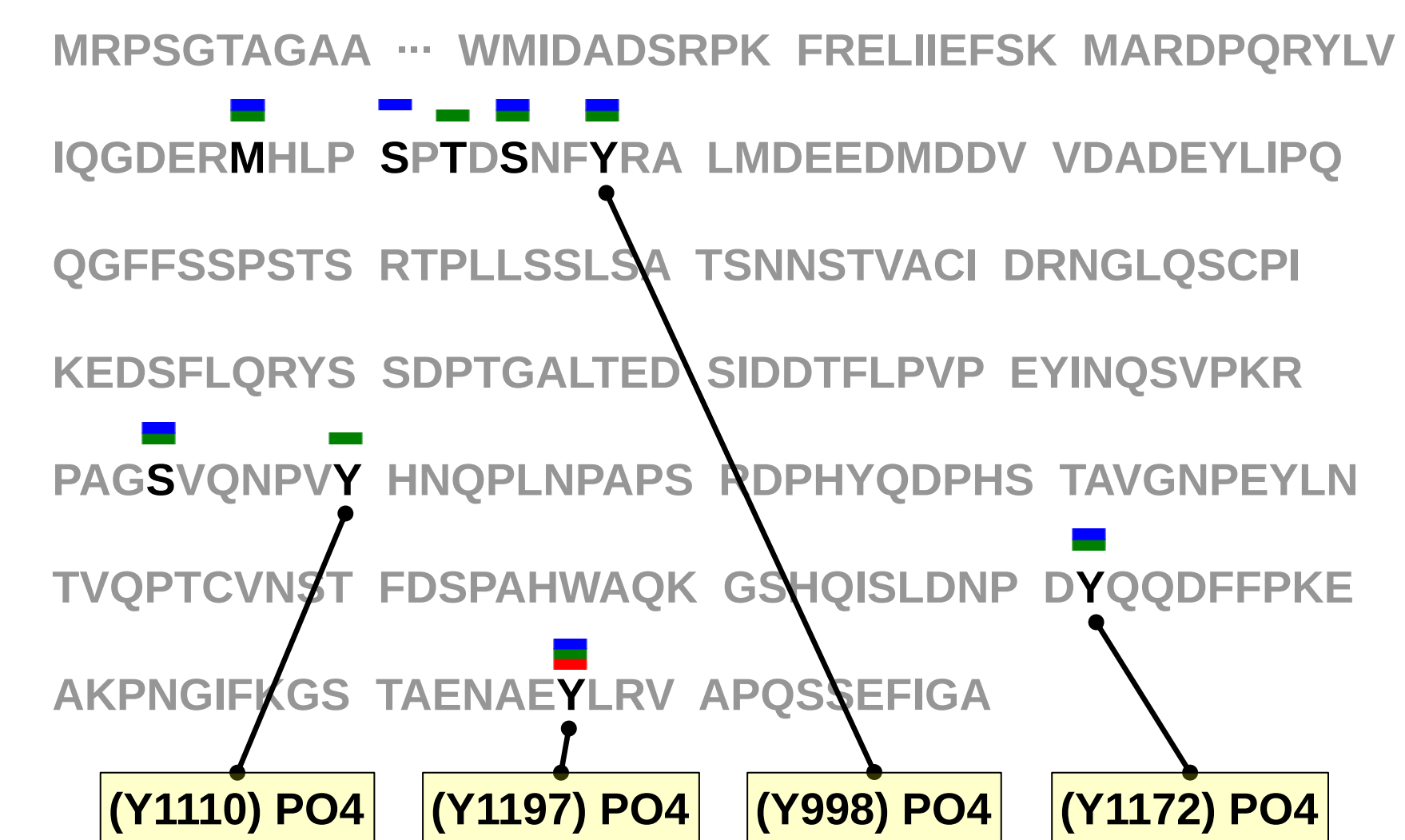
Results

System 3: EGFR Phosphorylation



PTM Map

Legend: Control ATP EGF



PTM Scoring

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	0.265	19.0
1110	Y	79.966331	0	1	0	1	0.33	0.58	2	0.983	0.125	0.333	0.265	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

