VISIT OUR **WEBSITE**

strand

Annotation and Differential Analysis of Protein Post-Translational Modifications for Target Discovery

Badri Padhukasahasram, Durairaj Renu, Ramesh Hariharan
Strand Life Sciences, Bangalore, India

Contact
Badri Padhukasahasram
VP, Data Science
badri.p@strandls.com
VISIT **LINKEDIN**

Background

- Post translational modifications (PTMs) are covalent modifications of proteins and can range from small chemical changes to addition of entire proteins
- PTMs regulate folding, localization, interactions, degradation and activity of proteins and can also play a role in signal transmission
- Aberrant PTMs can drive disease and disorders by altering protein folding and dysregulation of cell signalling.
- We seek to understand how differences in post translational modifcation patterns can enable new target discovery using data from tandem MS/MS

Highlights

- A validated PTM quantification and differential analysis pipeline with multiple metrics for quantifying PTM levels
- Pipeline currently takes ~ 2.75 hrs per sample (AMD EPYC 7763 64-Core Processor 64 threads) and can be set up for parallel execution on larger datasets
- Statistically significant differences in post-translational modifications identified muscle contraction and cardiomyopathy pathways in heart as well Wnt/β-Catenin Signaling Pathway targets in a colorectal cancer dataset.

Methods

2 types of normalization metrics to quantify differential post-translational modifications across groups

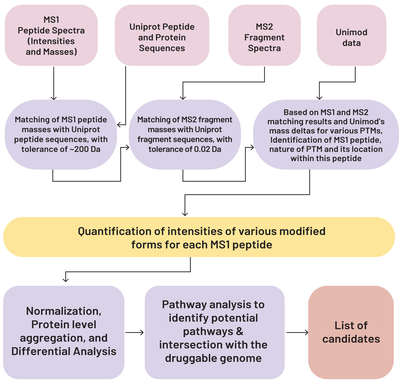
- **Differences in abundance of PTM peptide: log2(PTM_peptide_intensity/median_sample_intensity)**

  Intensity is normalized against the sample median intensity. This metric reflects any increase/decrease in the amount of post-translational modified peptide including situations where the increase might be driven by increase in the overall abundance of parent peptide for a given group/condition.

- **Differences in PTM usage: log2(PTM_peptide_intensity/Total_peptide_intensity)**

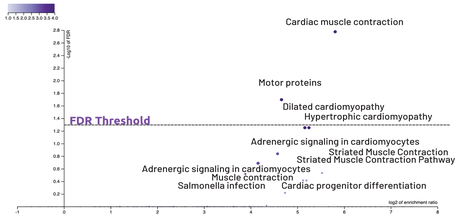
  This metric corrects for the overall peptide abundance for a given condition. It reflects an increase/decrease in the relative fraction of post-translational modified peptide versus total peptide for a given group/condition.

Strand post translational modification analysis pipeline for target discovery

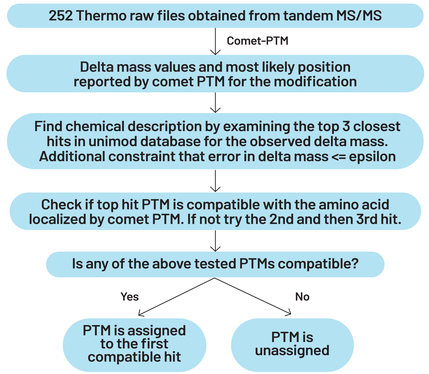


Results for heart dataset

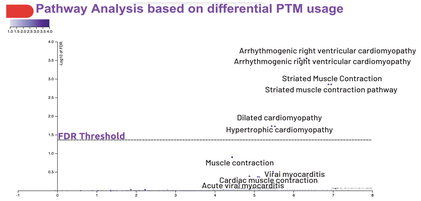
Pathway Analysis based on differential PTM abundance



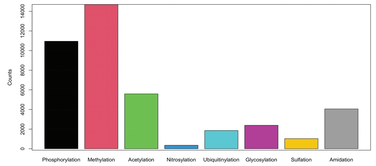
Post translational modification annotation method



Pathway Analysis based on differential PTM usage



Annotation of post-translational modifications in human heart data



Total signal intensity of PTMs by chambers maxQuant

Modification	Left Atrium	Right Atrium	Left Ventricle	As a fraction of total signal
Phosphorylation	34886.56	37348.97	40926.26	0.992%
Methylation	52744.62	47219.51	57857.18	1.378%
Acetylation	27598.66	25946.34	26465.63	0.703%

Aggregated signal intensity from all 3 PTMs is ~3% of total signal intensity

Counts of PTM peptides quantified by Comet-PTM

Modification	Left Atrium	Right Atrium	Left Ventricle
Phosphorylation	4758	4637	4755
Methylation	6291	6219	6564
Acetylation	2438	2381	2451
Nitrosylation	163	143	157
Ubiquitinylation	791	832	775
Glycosylation	1059	1003	1068
Sulfation	464	404	458
Amidation	1729	1719	1767

Results

Top gene targets from differential PTM analysis

Dataset	Metric	Genes
Cardiac Biopsy	PTM abundance	MYH6, MYL2 ,MYH7 ,AHNAK ,DNAJC5, YWHAB ,DNAH3, RTN4
Cardiac Biopsy	PTM usage	PKP2,GJA1,MYH6,DMD,MAP2K2,HRC,YAE1, CAMSAP2,CPS1,MAP1B,DES,H4C1
Colorectal Cancer	PTM abundance	DCDC2
Colorectal Cancer	PTM usage	CTNNB1, LIMA1, EHD4, TGFB1L1

Conclusions

- We have developed analysis pipelines and scripts for PTM identification , annotation as well as differential analysis across conditions from tandem MS/MS data.
- Methods were tested on i) Cardiac biopsy data from 3 heart chambers (Bagwan et al 2021) ii) Colorectal cancer vs control samples (Tanaka et al 2020).
- 2 different types of PTM signal normalization metrics were tested for differential analysis and indicated different sets of top gene hits for the 3 modifications considered i) Phosphorylation ii) Methylation iii) Acetylation.
- Muscle contraction pathways and cardiomyopathy pathway were overrepresented among top hits in the heart dataset.
- CTNNB1, LIMA1 and DCDC2 genes indicated significant differences in PTMs for a colorectal dataset. All of these influence Wnt/β-Catenin Signaling Pathway.
- Statistically significant differences in post-translational modifications for proteins suggest exploration of compounds that can alter these chemical modifications.

Future Directions for Research

- Extend differential analysis to other modifications apart from acetylation, phosphorylation and methylation.
- Explore newer PTM detection tools such as Comet-Recom for PTM identification improvements.
- Investigate impact of PTM on protein structure and protein-protein interactions using alpha fold 3.
- Explore AI and deep learning based approaches for PTM detection.