Post-translational modification bias between organism complexity: Trends observed across diverse lifeforms

Oliver Bonham-Carter,
Ishwor Thapa, Dhundy (Kiran) Bastola (*Advisor*)

School of Interdisciplinary Informatics
University of Nebraska at Omaha,
Omaha, NE, 68182 USA

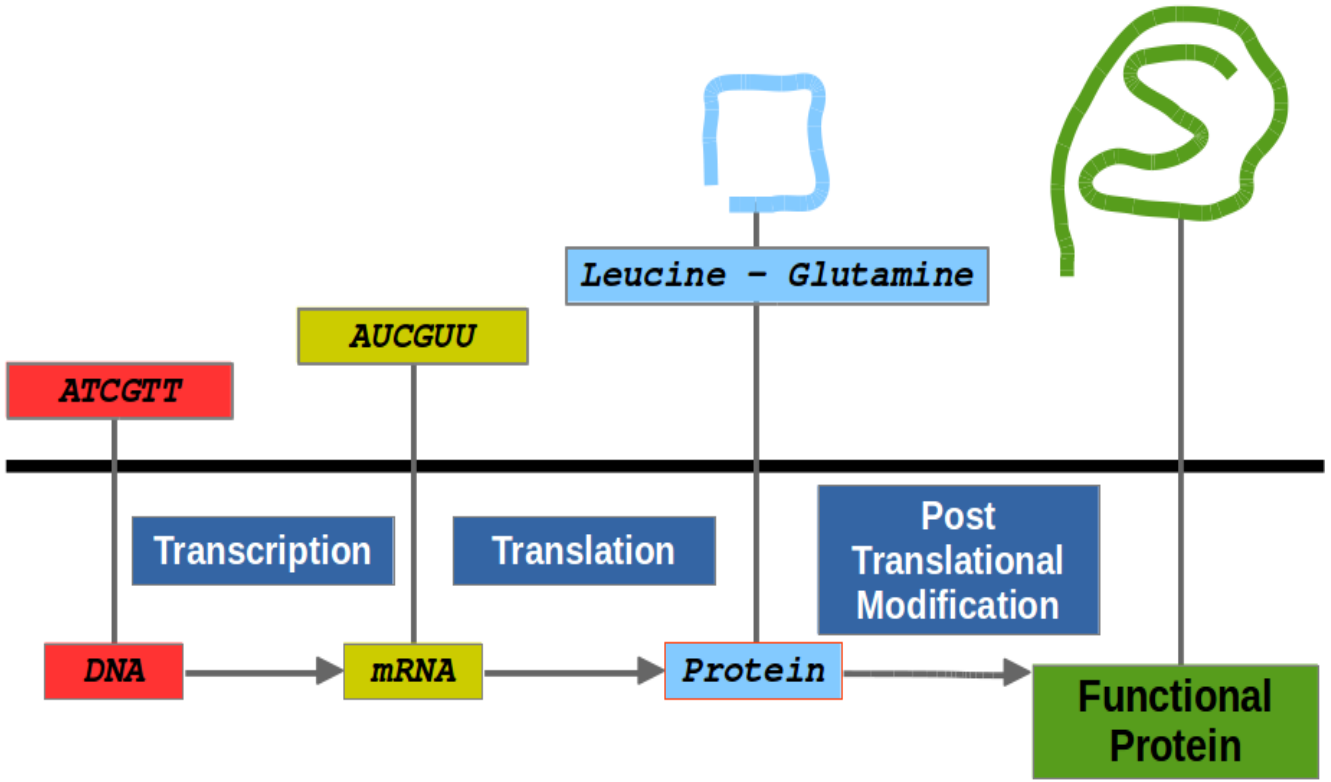
24 April, 2015

A Tailored Customization:
A Short Intro to Post Translational Modifications

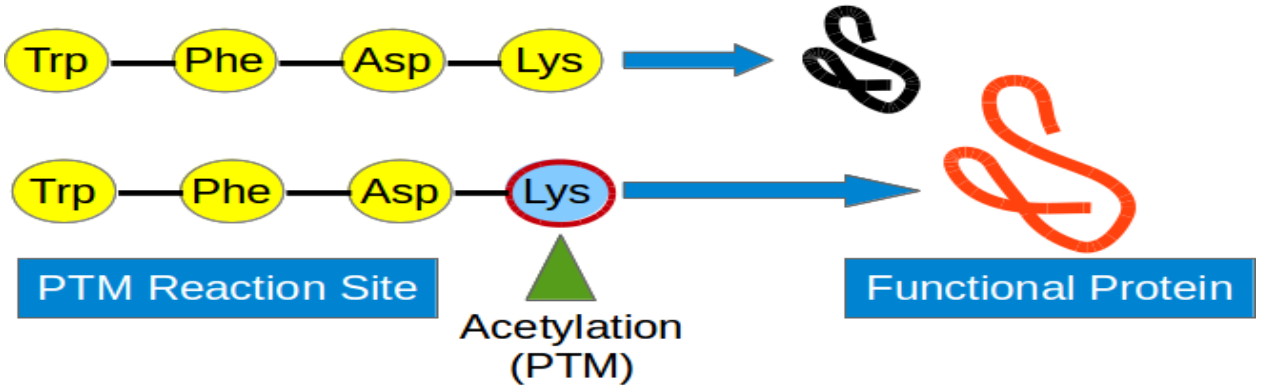
Introduction
Methods
Results
Conclusions
References
Thanks To

PTMs in the Central Dogma of Biology

Introduction
Methods
Results
Conclusions
References
Thanks To



PTMs: a Step in Protein Biosynthesis

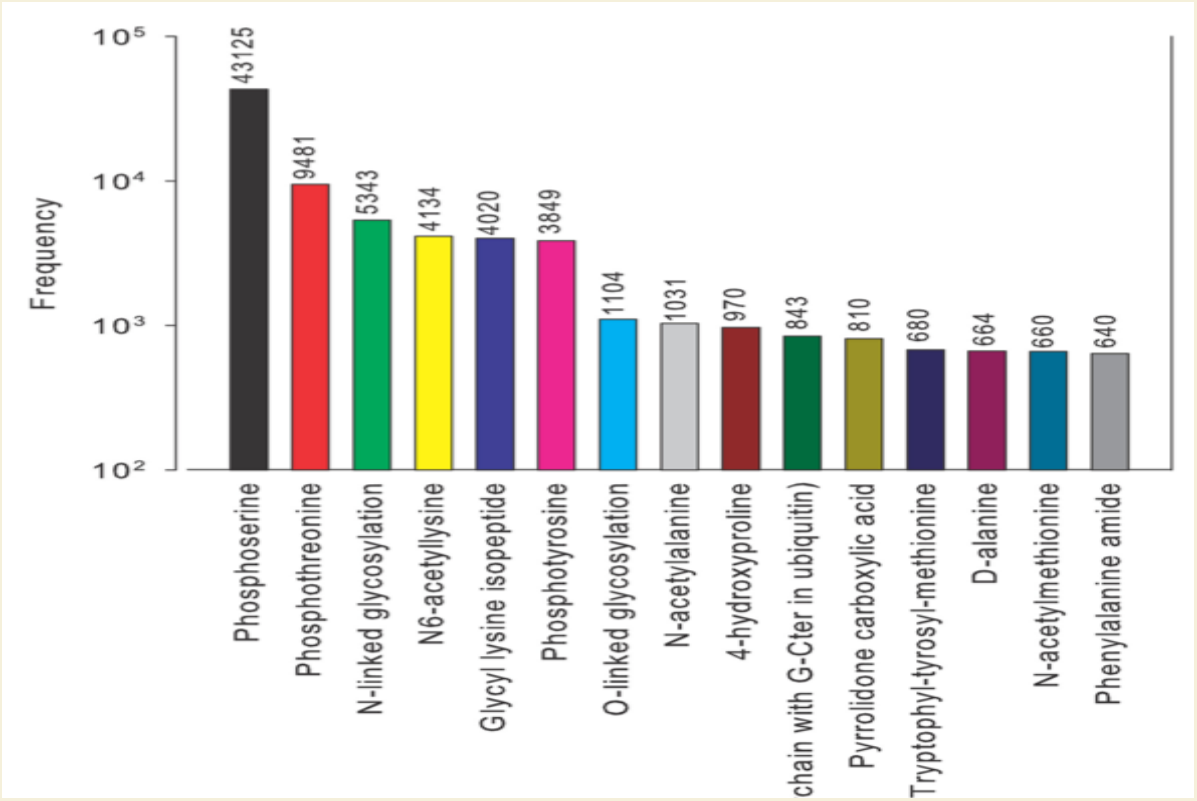


- Protein regulators
- Able to quickly change protein behaviors and functions

Gene expression	acetylation	glycosylation
Protein regulation	phosphorylation	sumoylation

PTMs Are Not Used Consistently Across Organisms

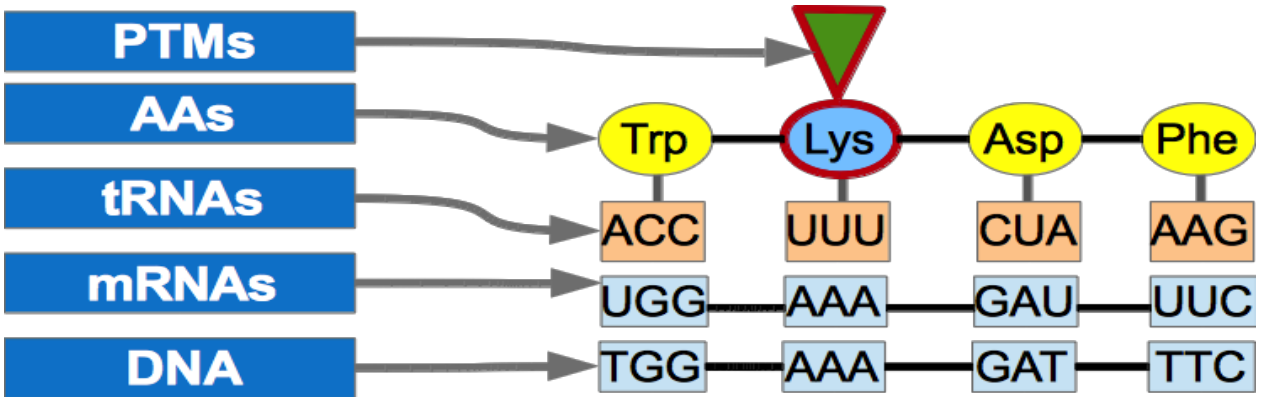
Introduction
Methods
Results
Conclusions
References
Thanks To



- What could alter the way that PTMs are used across organisms?

Khoury *et al.* Proteome-wide post-translational modification statistics. `http://selene.princeton.edu/PTMCuration/`

PTMs Interact With Protein at Reactive Sites



- Reactive Site: a point of reaction between proteins and PTMs (lysine for acetylation).
- Lysine is a specific amino acid that targets PTM interaction
- Could reactive site placement interfere with PTM bias?

Could tRNA Bias Influence PTM Bias?

tRNA Biases Interfere With Amino Acid Biases!

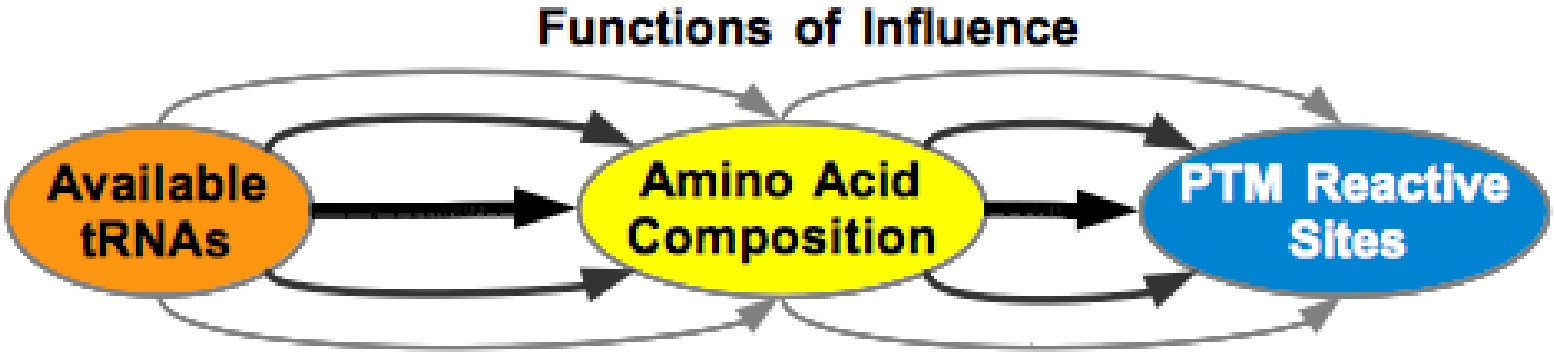
Evidence of a Pathway of Reduction in Bacteria: Reduced Quantities of Restriction Sites Impact tRNA Activity in a Trial Set

Oliver Bonham-Carter, Lotfollah Najjar, Dhundy Bastola
School of Interdisciplinary Informatics
University of Nebraska at Omaha
Omaha, NE, 68182, USA
Email{obonhamcarter, lnajjar, dkbastola}@unomaha.edu

- Bonham-Carter, Oliver, Lotfollah Najjar, and Dhundy Bastola. "**Evidence of a pathway of reduction in bacteria: Reduced quantities of restriction sites impact trna activity in a trial set.**" Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics. ACM, 2013.

A Connection Between Variables
tRNAs, AAs and PTM Reaction Sites → PTMs

Introduction
Methods
Results
Conclusions
References
Thanks To

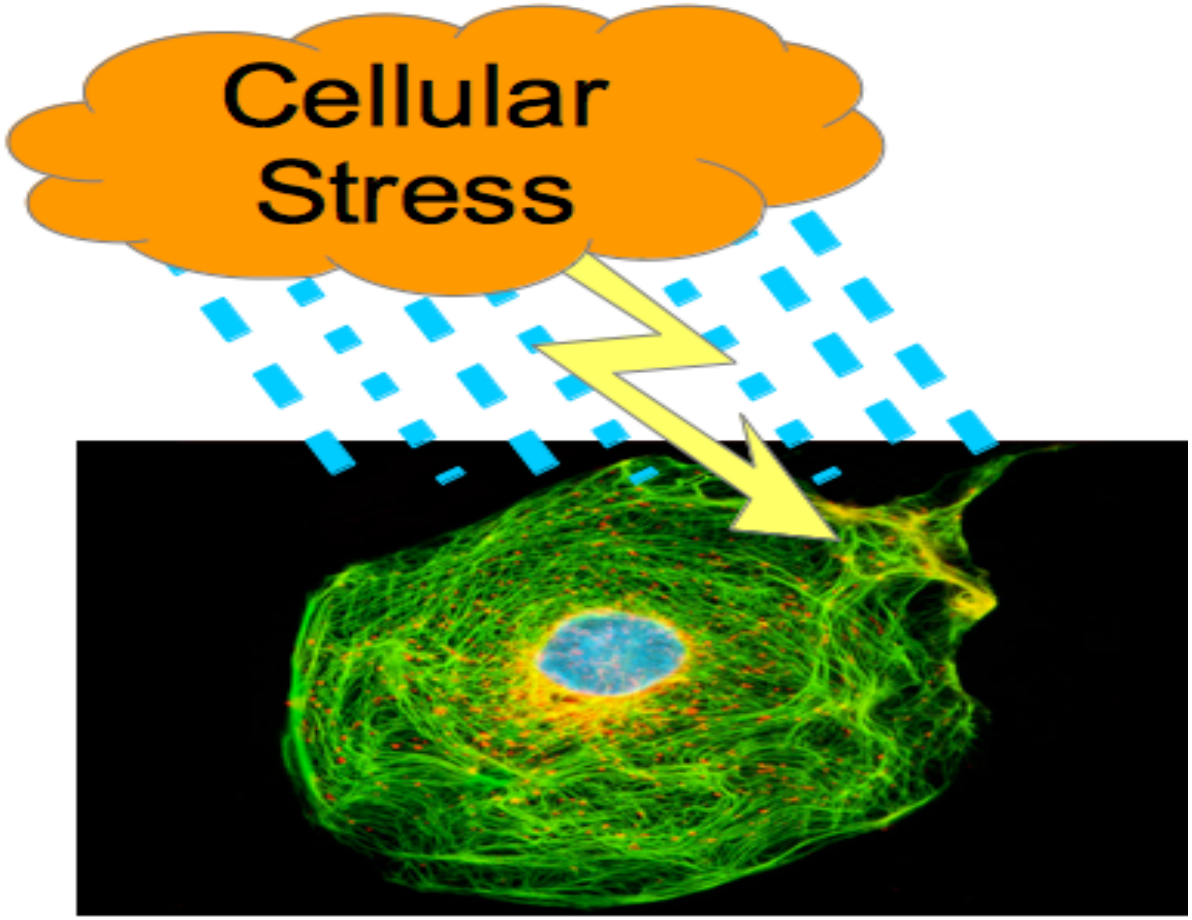


- AAs form the specific reactive sites where PTMs interact with protein

Stresses Influence PTMs

Nor rain, wind, thunder, fire are my daughters. -Shakespeare (King Lear)

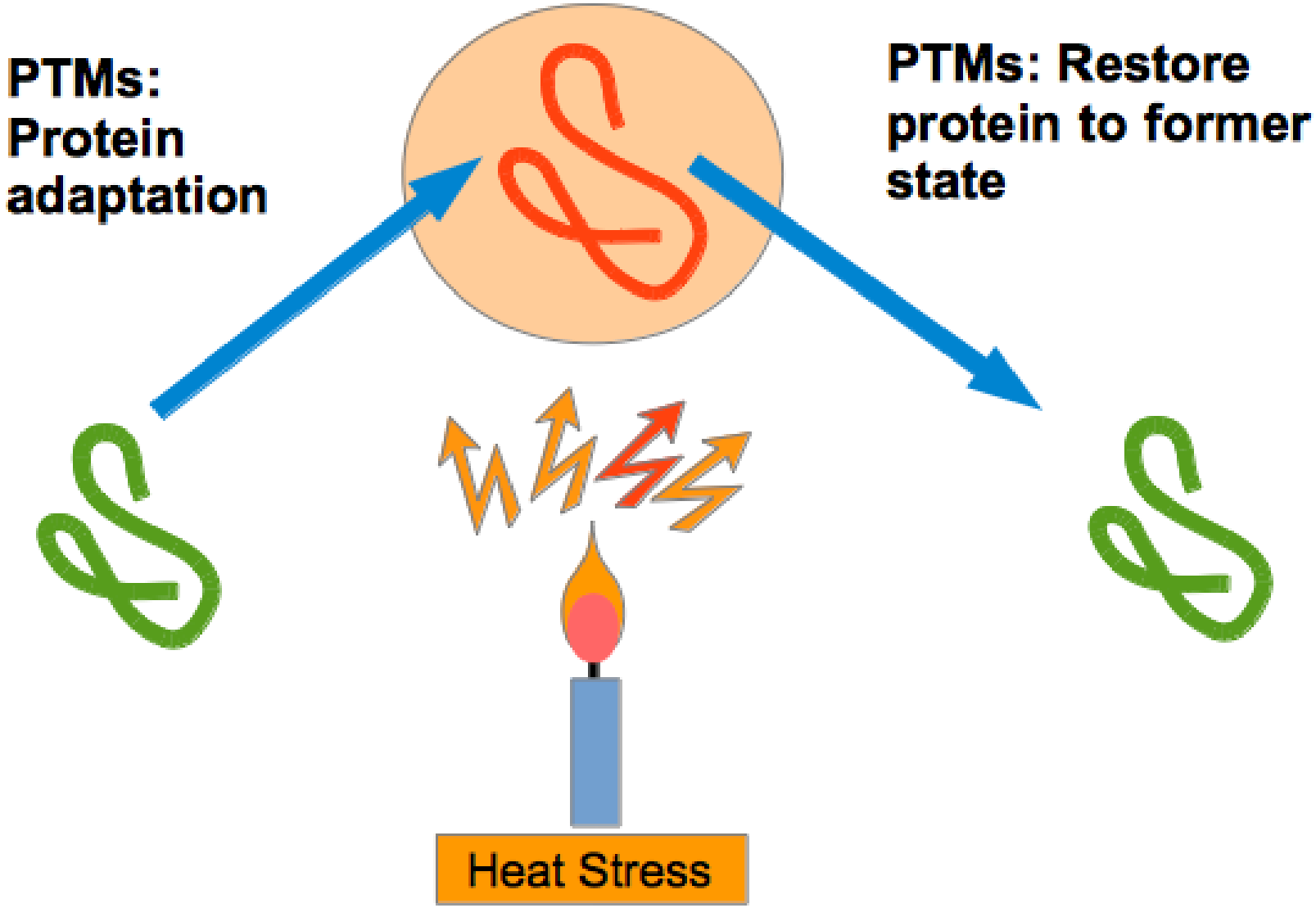
Introduction
Methods
Results
Conclusions
References
Thanks To



- Cellular Stresses: Carbonylation, Free Radicals, Heat Shock, Microgravity, Saline, and others.

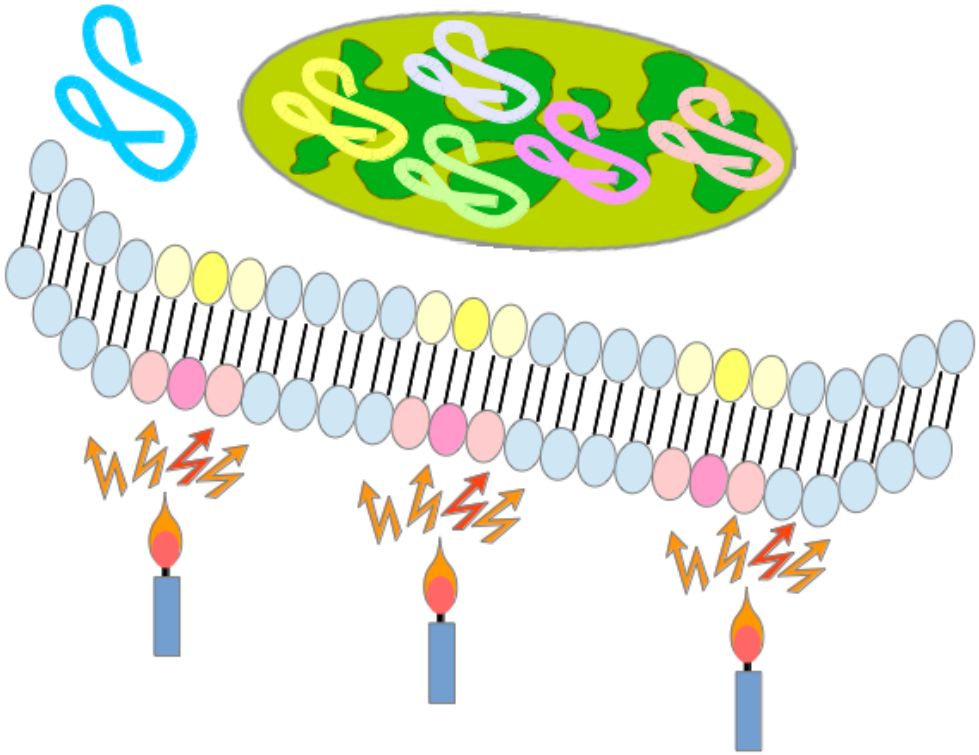
PTMs: Quick Adaptations to Env. Stresses

Introduction
Methods
Results
Conclusions
References
Thanks To



Could a Bias Come From Stresses

Introduction
Methods
Results
Conclusions
References
Thanks To



- Mitochondria (Mt): Cellular Energy Producers
- Mt proteins: Likely evolved with different stress types to influence tRNA distributions
- Is there bias observed in Mt PTM composition?

Research Interests

Introduction
Methods
Results
Conclusions
References
Thanks To

One of Our Paper's Hypotheses...

- Mt proteins are likely protected from main stream cellular stresses.
- Are there fewer PTMs available to Mt for stress response?
- Can biases be observed from across organisms (in both Mt and non-Mt)?

Data: Organisms

Introduction
Methods
Results
Conclusions
References
Thanks To

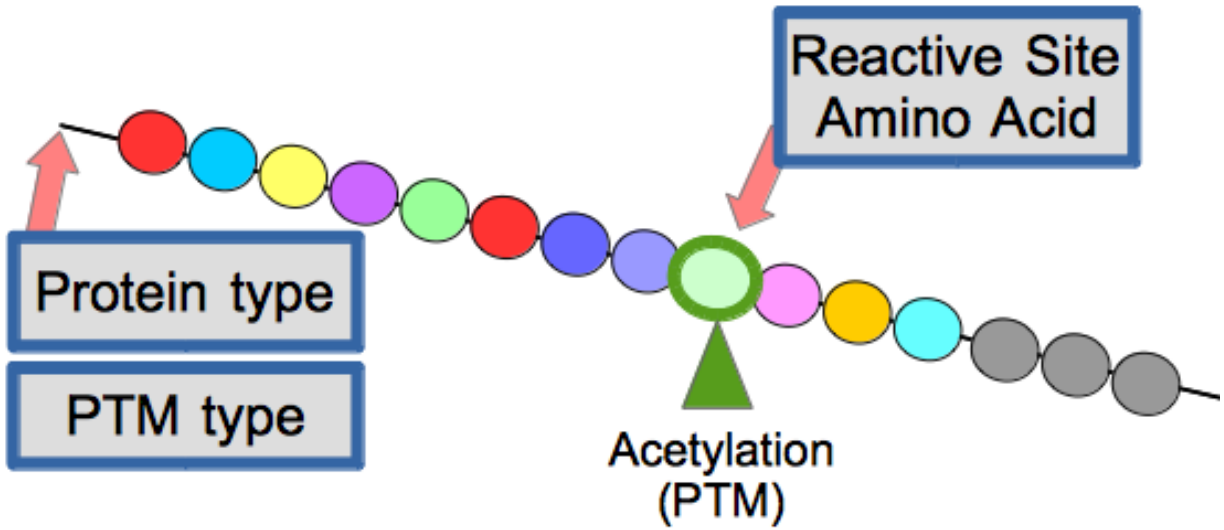
Table : The proteomes came from these organisms for our study.

No.	Organism	Commonly	Top PTM
1	Arabidopsis thaliana	Plant	Glycos
2	Caenorhabditis elegans	Worm	Glycos
3	Canis familiaris	Dog	Glycos
4	Danio rerio	Fish	Glycos
5	Homo sapiens	Human	Phospho
6	Mus musculus	Mouse	PhosPho
7	Oryctolagus cuniculus	Rabbit	Glycos
8	Rattus norvegicus	Rat	Glycos
9	Saccharomyces cerevisiae	Yeast	Phospho

Data: Frequencies of Elements

DataBase: Uniprot

```
for each protein by organism:
 obtain type and count of:
   PTMs,
   Reaction Locations (Active Sites),
   Amino Acids (composition)
```



- Collected over Mt and nonMt proteins

Three Frequency Equations for Three Questions

Introduction
Methods
Results
Conclusions
References
Thanks To

- Composition of **PTM** content in Mt versus nonMt protein?
  - $freq(PTM_{(i,j)}) = \frac{count(PTM_{(i,j)})}{\sum_{i=1}^{N_{(PTMs)}} count(PTM_{(i,j)})}$
- Composition of **AA** content in Mt versus nonMt protein?
  - $freq(aminoAcid_{(i,j)}) = \frac{count(aminoAcid_{(i,j)})}{\sum_{i=1}^{N_{(reactiveSites)}} count(reactiveSite_{(i,j)})}$
- Composition of PTM **reactive sites** in Mt versus nonMt protein?
  - $freq(ReactiveSite_{(i,j)}) = \frac{count(aminoAcid_{(i,j)})}{|\sum_{i=1}^{N_{Proteins}} Seq_{(i,j)}|}$

- For $(i, j) = (element[i], organism[j])$
- $N =$ size of set
- $count()$ function returns number of an element in set of size N

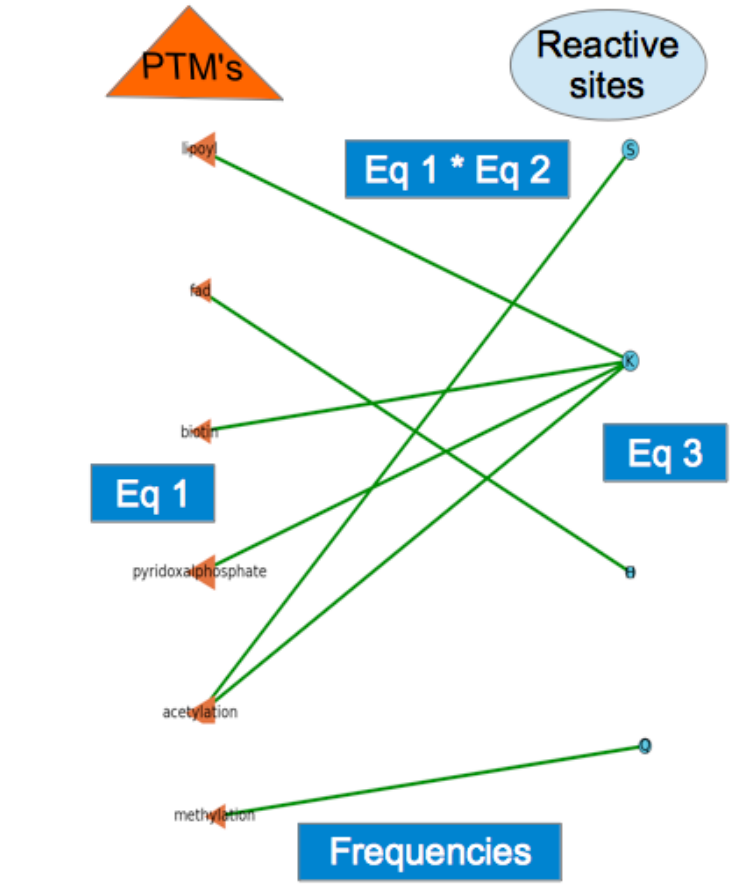
A Network Comparison

Introduction
Methods
Results
Conclusions
References
Thanks To

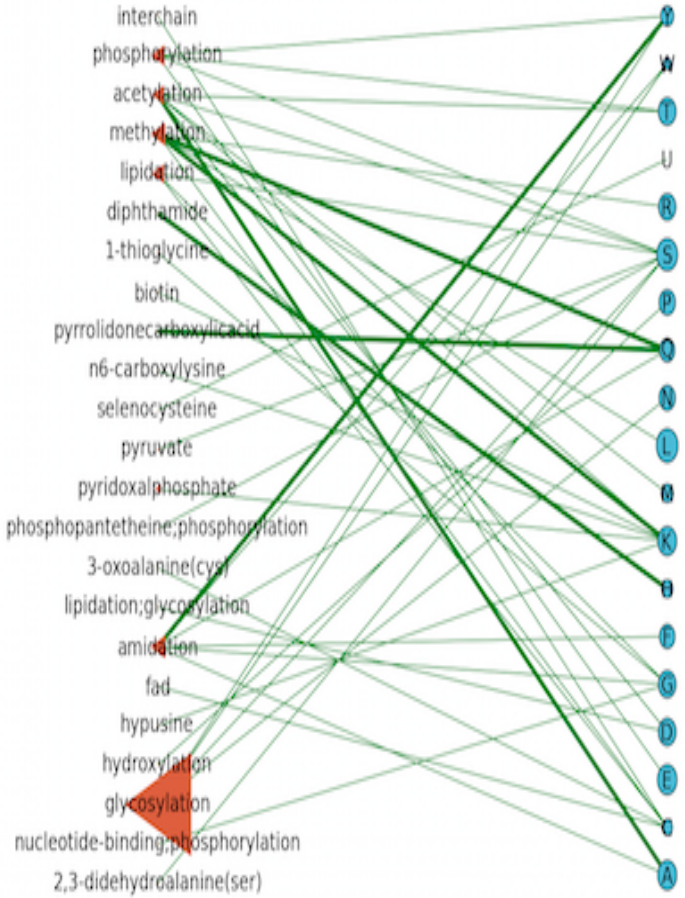
- How are the PTM distributions different between the Mt and non-Mt proteomes?

Mt and non-Mt Networks

Caenorhabditis elegans (Worm)



Mt

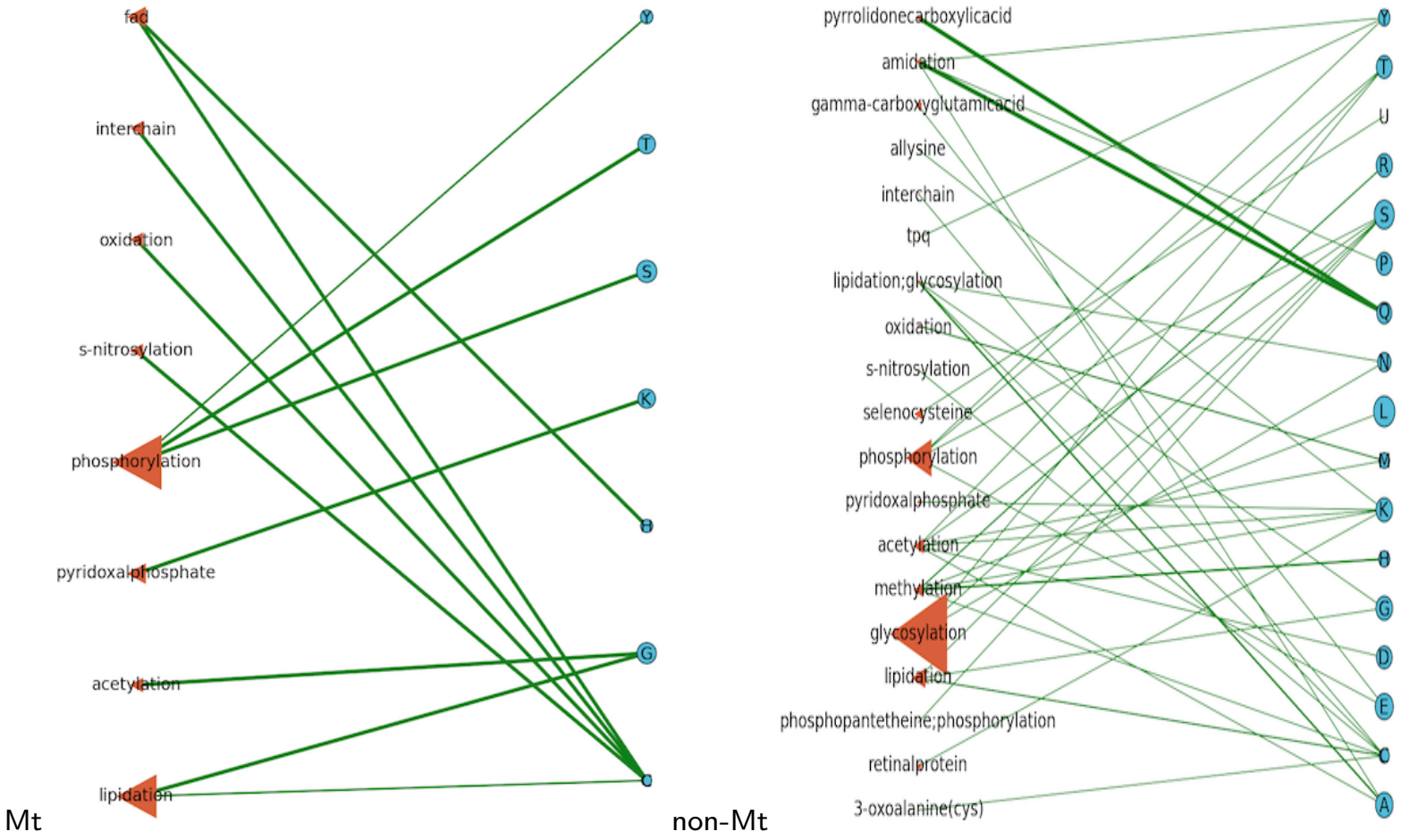


non-Mt

- Nodes: **PTMs** (left) node size is freq magnitude, **reactive site** (right)
- Edges: reactive site freq, thickness magnitude of reactive site interactions

Mt and non-Mt Networks

Danio rerio (Zebra fish)

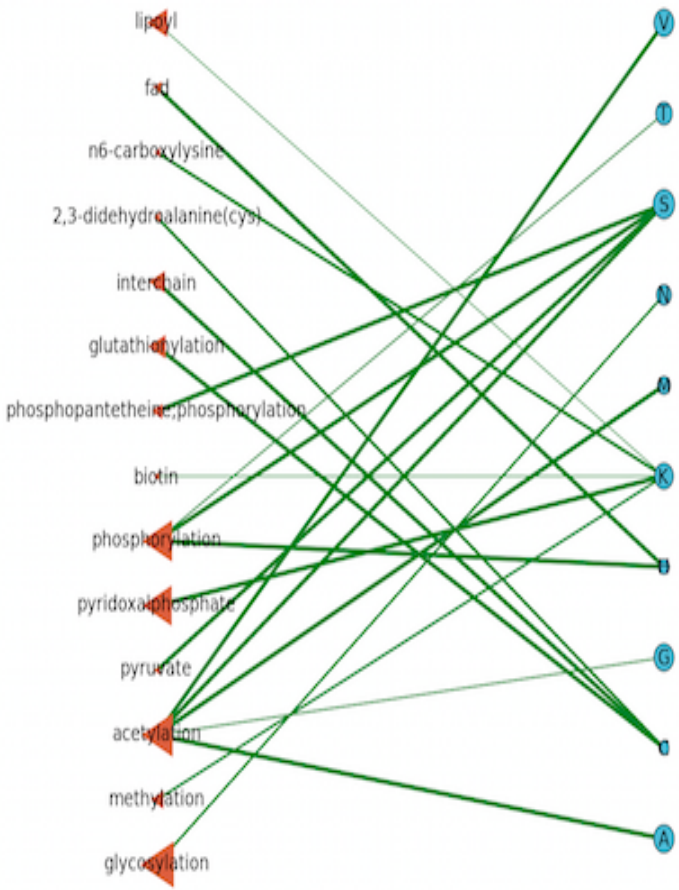


- Nodes: **PTMs** (left) node size is freq magnitude, **reactive site** (right)
- Edges: reactive site freq, thickness magnitude of reactive site interactions

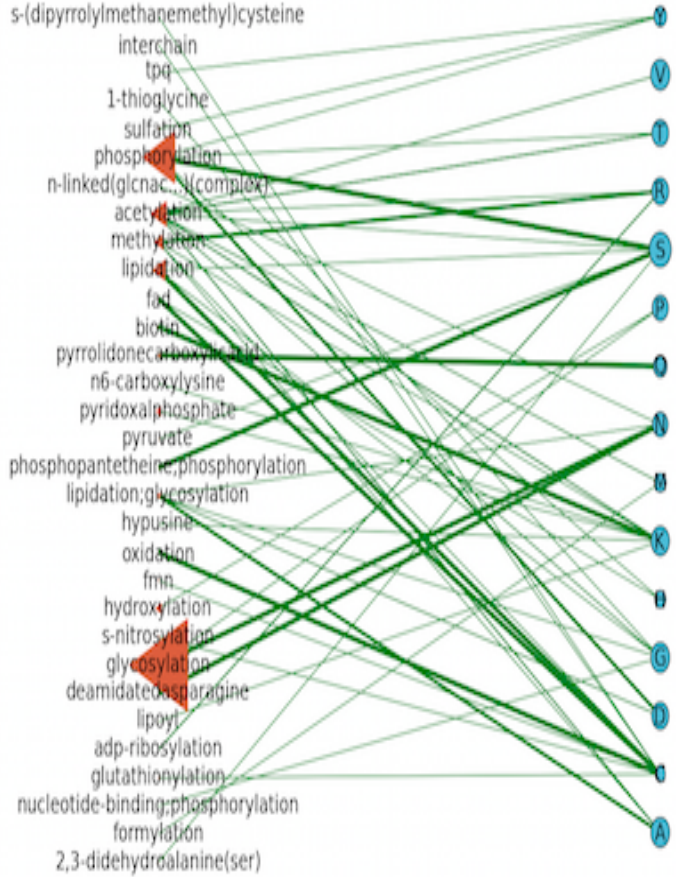
Mt and non-Mt Networks

Arabidopsis thaliana (plant)

Introduction
Methods
Results
Conclusions
References
Thanks To



Mt

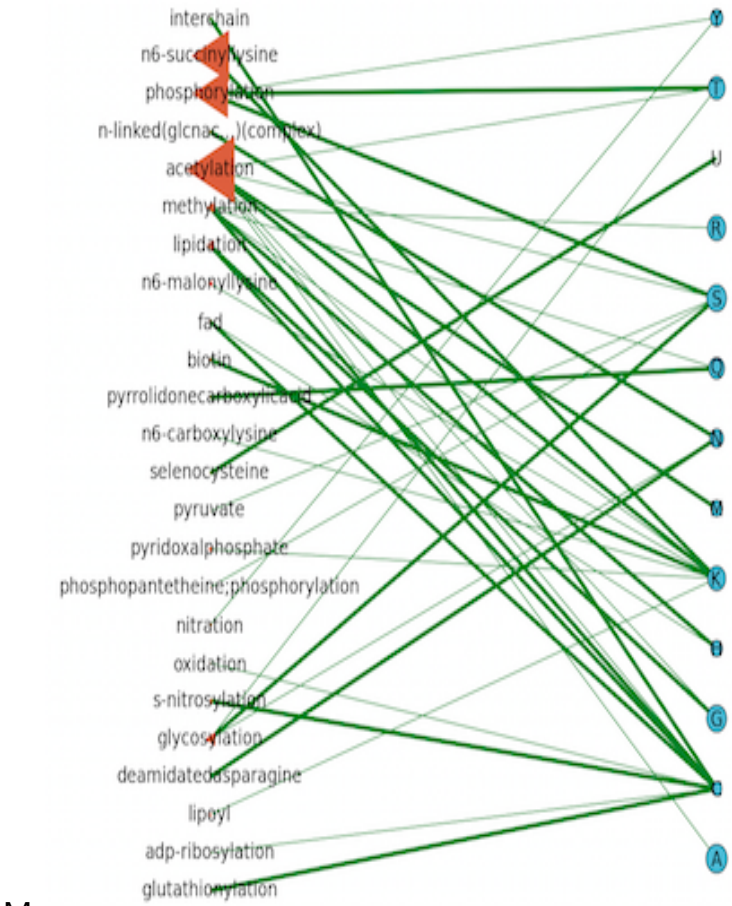


non-Mt

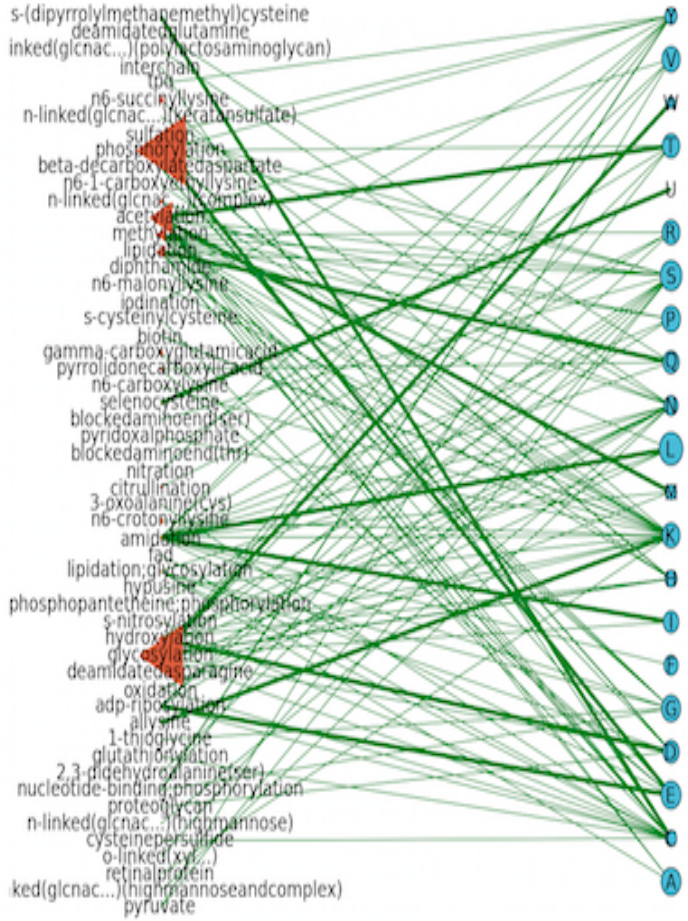
- Nodes: **PTMs** (left) node size is freq magnitude, **reactive site** (right)
- Edges: reactive site freq, thickness magnitude of reactive site interactions

Mt and non-Mt Networks

Homo sapiens (Human)



Mt



non-Mt

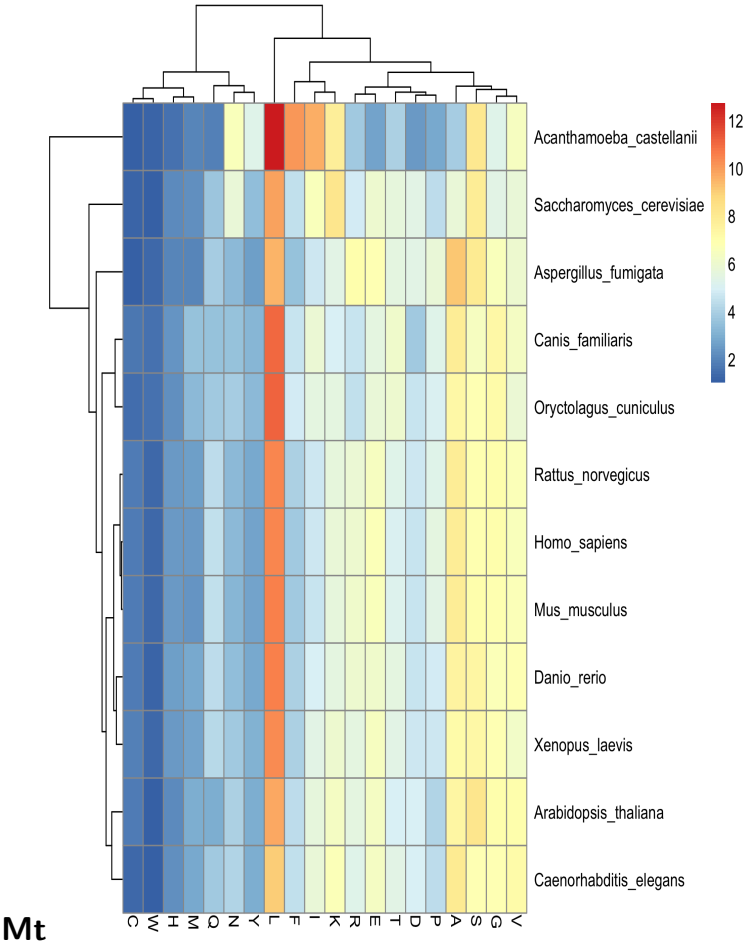
- Nodes: **PTMs** (left) node size is freq magnitude, **reactive site** (right)
- Edges: reactive site freq, thickness magnitude of reactive site interactions

A Heatmap Comparison

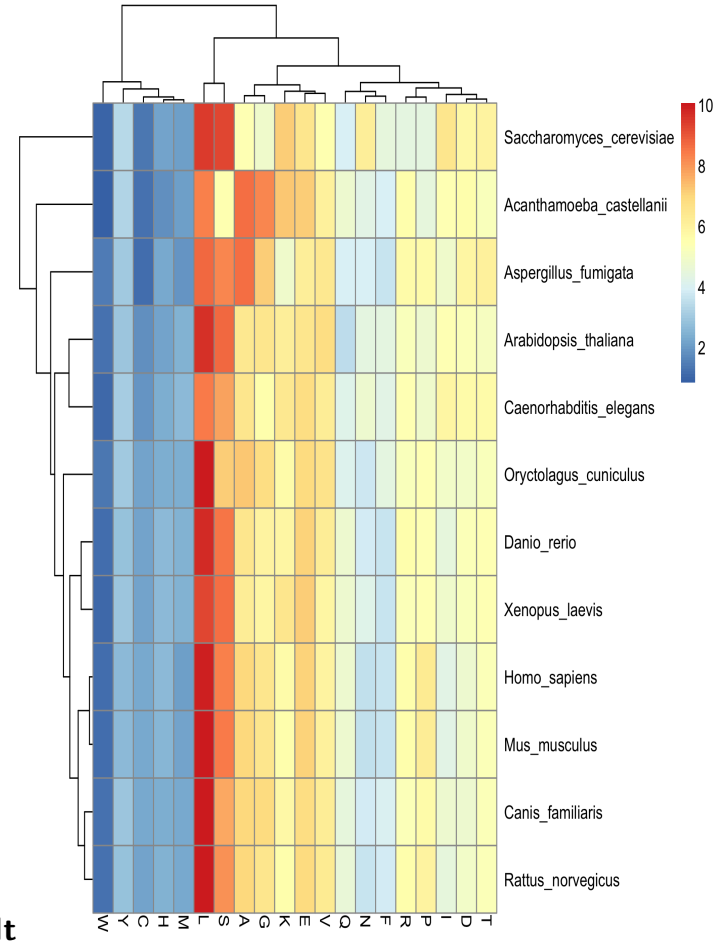
Introduction
Methods
Results
Conclusions
References
Thanks To

- How are the **amino acid** distributions different between the Mt and non-Mt proteomes?
  - $freq(aminoAcid_{(i,j)}) = \frac{count(aminoAcid_{(i,j)})}{\sum_{i=1}^{N_{(reactiveSites)}} count(reactiveSite_{(i,j)})}$

Composition of All AA's Across Organisms



Mt



non-Mt

- Amino acids frequencies are similar across related organisms
- Dark blue values are very close to zero

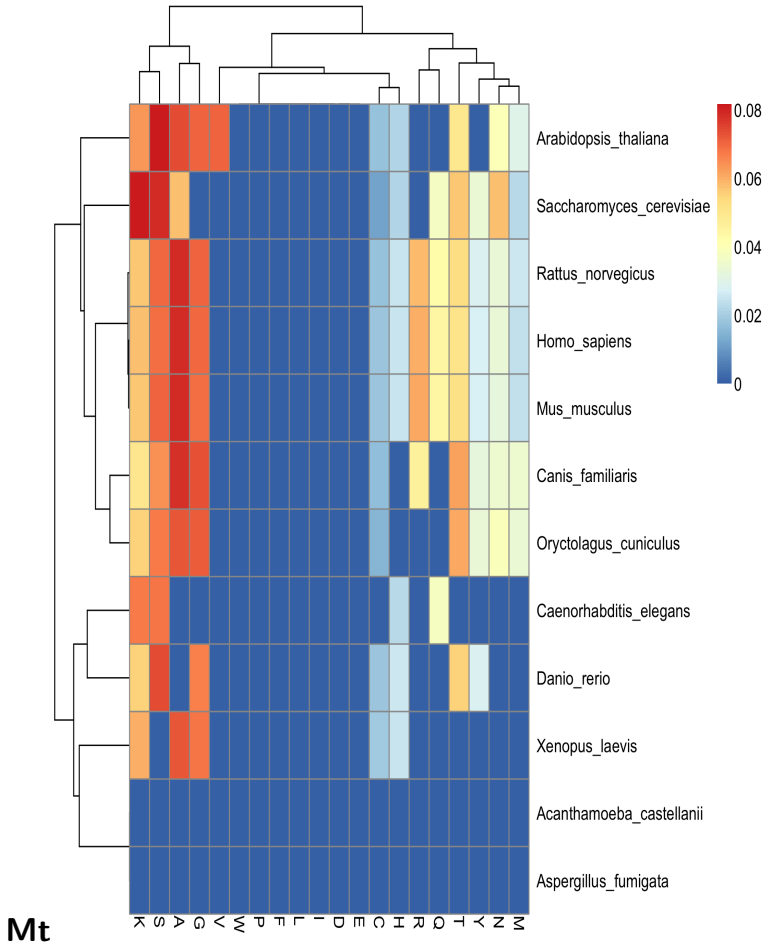
A Heatmap Comparison

Introduction
Methods
Results
Conclusions
References
Thanks To

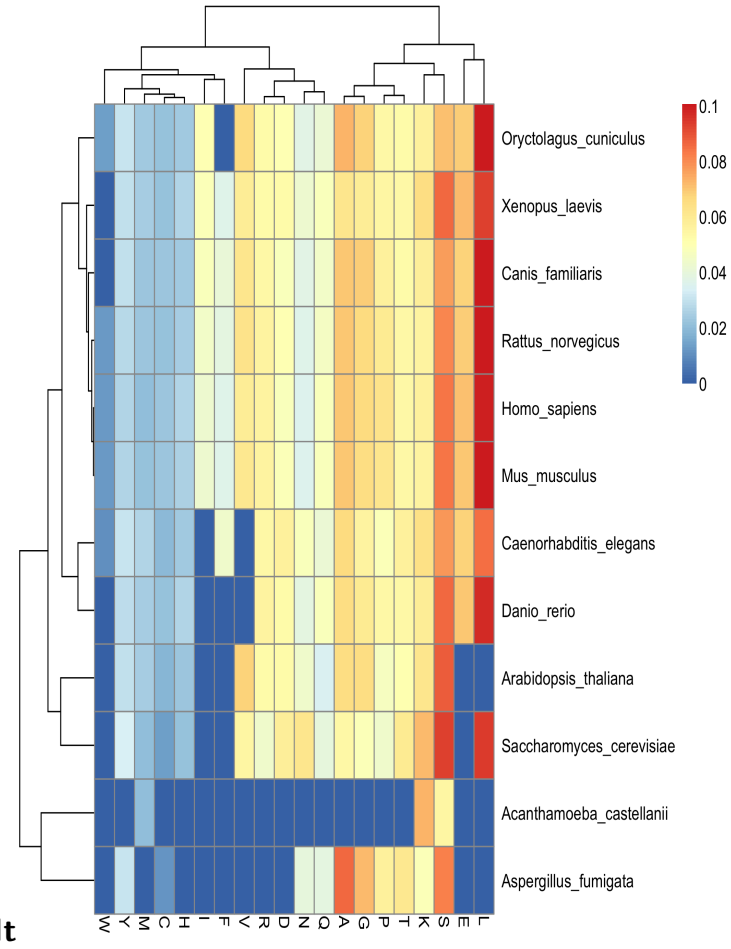
- How are the **PTM Reactive Site** distributions different between the Mt and non-Mt proteomes?
  - $freq(ReactiveSite_{(i,j)}) = \frac{count(aminoAcid_{(i,j)})}{|\sum_{i=1}^{N_{Proteins}} Seq_{(i,j)}|}$

Composition of All PTM RSs Across Organisms

Introduction
Methods
Results
Conclusions
References
Thanks To



Mt



non-Mt

- Fewer reactive sites in Mt
- Dark blue values are very close to zero

Some of the Conclusions

Network Particulars

Protein Type	PTMs	Reactive Sites per PTM	Networks
Mt	Few	Few	Sparse and organized
Non-Mt	Many	Many	Dense, disorganized and messy

- **From Hypothesis**: Mt proteins have fewer PTMs and associated reactive sites than non-Mt proteins
- Future work: To study first effects of stress on protein reactive sites by observing PTM *first-responders*.

References: Cited in this Presentation

Introduction
Methods
Results
Conclusions
References
Thanks To

- Bonham-Carter, Oliver, Ishwor Thapa, and Dhundy Bastola. **"Evidence of post translational modification bias extracted from the tRNA and corresponding amino acid interplay across a set of diverse organisms**." Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics. ACM, 2014.

- Bonham-Carter, Oliver, Lotfollah Najjar, and Dhundy Bastola. "**Evidence of a pathway of reduction in bacteria: Reduced quantities of restriction sites impact trna activity in a trial set.**" Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics. ACM, 2013.

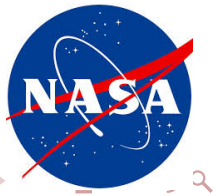
Thanks to: Dr. Hesham Ali, Dr. Dhundy (Kiran) Bastola, Dr. Kate Cooper, Dr. Lotfollah Najjar, Dr. Sanjukta Bhowmick, Jasjit Banwait, Julia Warnke, Ishwor Thapa, Jay Pedersen, Sean West, Scott McGrath, Vladimir Ufimtsev and Kritika Karri.

This project has been funded by the grants from the National Center for Research Resources (5P20RR016469), the National Institute for General Medical Science (NIGMS) (8P20GM103427).

This study also funded by the NASA Nebraska Space Grant (2014-2015).

Thank You! Questions?

obonhamcarter@unomaha.edu

IS&T Bioinformatics

http://bioinformatics.ist.unomaha.edu/