Evidence of post translational modification bias extracted from the tRNA and corresponding amino acid interplay across a set of diverse organisms

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

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

20 September 2014
A Tailored Customization

Introduction

Methods

Results

Conclusions

References

Thanks To
Post Translational Modifications (PTMs) are steps in biosynthesis to create functional proteins.
Protein regulators

Quick protein adaptation:
- Enable protein to overcome stress conditions
- Restore protein to original form after stress has elapsed

Similar kinds of functionality across PTMs:
- Gene expression: acetylation, glycosylation, etc.
- Protein regulating: phosphorylation, sumoylation, etc.
Transfer RNAs (tRNAs): Amino Acid Delivery

- Each Amino Acid (AA) is added by a unique tRNA to protein sequence during *translation*
A Unique tRNA For Each AA

- tRNAs have unique codons to deliver the same AAs
- If a tRNA is unavailable, the delivery of its AAs in protein sequence is impossible
Distribution of tRNAs

- Adapter molecules between the mRNA sequence and peptide sequence (protein)
- Different tRNA compositions across genomes (Codon Usage Database: http://www.kazusa.or.jp/codon/)
PTMs Interact With Protein at Reaction Sites

- Reaction Site: a point of reaction between proteins and PTMs (lysine for acetylation).
- Lysine is a specific amino acid that targets PTM interaction.
PTMs Are Not Used Consistently Across Organisms

A Connection Between tRNAs, AAs and PTMs

Unique tRNAs responsible for delivering specific AAs
- AAs form the specific reaction sites where PTMs interact with protein
- The availability of tRNAs may impact which PTMs are able to interact with protein.
Mt is first and then nuclear processes respond to stresses (heat shock, etc.)

Energy is required to drive the PTM step to adapt proteins

Zhoa et al. “A mitochondrial specific stress response in mammalian cells”
Research Interest

- tRNA usage affects AA composition which makes it possible for PTMs to react with protein
- Stress
  - Mt: Energy makers and first to respond to stress
  - Protein adaptation by PTMs

What is the difference between AA, PTMs and tRNAs of first and second responders (Mt and non-Mt)?
- In terms of AA, tRNA, PTMs: profile Mt and non-Mt proteins.
Global **Snap-Shot**: Diverse Organisms of Mt

<table>
<thead>
<tr>
<th>Organism</th>
<th>N. Seqs.</th>
<th>Top PTMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mustard Plant</td>
<td>707</td>
<td>Glycosylation, Phosphoserine</td>
</tr>
<tr>
<td><em>Arabidopsis thaliana</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nematode worm</td>
<td>199</td>
<td>Glycosylation, Lipidation</td>
</tr>
<tr>
<td><em>Caenorhabditis elegans</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic Dog</td>
<td>60</td>
<td>Glycosylation, Phosphoserine</td>
</tr>
<tr>
<td><em>Canis familiaris</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zebrafish</td>
<td>202</td>
<td>Glycosylation, Phosphoserine</td>
</tr>
<tr>
<td><em>Danio rerio</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td>1027</td>
<td>Phosphoserine, Glycosylation</td>
</tr>
<tr>
<td><em>Homo sapiens</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>House Mouse</td>
<td>973</td>
<td>Phosphoserine, Glycosylation</td>
</tr>
<tr>
<td><em>Mus musculus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>European rabbit</td>
<td>46</td>
<td>Glycosylation, Phosphoserine</td>
</tr>
<tr>
<td><em>Oryctolagus cuniculus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway Rat</td>
<td>571</td>
<td>Glycosylation, Phosphoserine</td>
</tr>
<tr>
<td><em>Rattus norvegicus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakers Yeast</td>
<td>1056</td>
<td>Phosphoserine, Glycosylation</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Diverse organisms of Mt: Observed trends may be global
Data Collection

**DataBase: UniProtKB/Swiss-Prot**
- for each protein by organism:
  - obtain type and count of:
    - PTMs,
    - Reaction Locations (Active Sites),
    - Amino Acids (composition)

**DataBase: Codon Usage**
- for each organism:
  - obtain type and count of:
    - Codon frequencies (*)

(*) Since a tRNA involves a specific codon, the codon frequencies were good approximations of tRNA frequencies

- Collected over mitochondrial (Mt) and nonMt proteins
Three Frequency Equations for Three Questions

- Composition of **PTM** content in Mt versus nonMt protein?
  \[ freq(PTM_{(i,j)}) = \frac{\text{count}(PTM_{(i,j)})}{\sum_{i=1}^{N_{(PTMs)}} \text{count}(PTM_{(i,j)})} \]

- Composition of **AA** content in Mt versus nonMt protein?
  \[ freq(\text{aminoAcid}_{(i,j)}) = \frac{\text{count}(\text{aminoAcid}_{(i,j)})}{\sum_{i=1}^{N_{(actSites)}} \text{count}(actSite_{(i,j)})} \]

- Composition of PTM **reaction sites** in Mt versus nonMt protein?
  \[ freq(\text{ReactionSite}_{(i,j)}) = \frac{\text{count}(\text{aminoAcid}_{(i,j)})}{|\sum_{i=1}^{N_{(Proteins)}} Seq_{(i,j)}|} \]

- For \((i, j) = (element[i], organism[j])\)
- \(N = \text{size of set}\)
- \(\text{count()}\) function returns number of an element in set of size \(N\)
Composition of \textbf{PTM} content in Mt versus nonMt protein?

\[ \text{freq}(PTM_{i,j}) = \frac{\text{count}(PTM_{i,j})}{\sum_{i=1}^{N_{PTMs}} \text{count}(PTM_{i,j})} \]
PTMs: Types and Magnitudes Across Organisms
Mt: More significant PTMs; Glycosylation and Phosphoserine predominance

**x-axis:** Organisms, **y-axis:** PTMs, **Grid:** Freqs > 0.1 value threshold.
Composition of **AA** content in Mt versus nonMt protein?

\[ freq(\text{aminoAcid}_{i,j}) = \frac{\text{count}(\text{aminoAcid}_{i,j})}{\sum_{i=1}^{N(\text{actSites})} \text{count}(\text{actSite}_{i,j})} \]
AA: Types and Magnitudes Across Organisms
Mt: Fewer reaction sites, similar freqs: *H. Sapiens* to *R. norvegicus*

$x$-axis: Organisms, $y$-axis: PTMs, Grid: Freqs $> 0.1$ value threshold.
Composition of PTM reaction sites in Mt versus nonMt protein?

\[
\text{freq}(\text{ReactionSite}(i,j)) = \frac{\text{count}(\text{aminoAcid}(i,j))}{|\sum_{i=1}^{N_{\text{Proteins}}} \text{Seq}(i,j)|}
\]

Equation 3 Provides Edges:
Describe relationships between AAs, tRNAs and PTMs
Caenorhabditis elegans: Mt and non-Mt Networks

Mt: Less PTMs, AA

Nodes: PTMs (left) node size is freq magnitude, tRNA (right)

Edges: active site freq, thickness magnitude of active site interactions
Arabidopsis thaliana: Mt and non-Mt Networks

Mt: Less PTMs, AA

Nodes: PTMs (left) node size is freq magnitude, tRNAs (right)
Edges: active site freq, thickness magnitude of active site interactions
**Homo sapiens: Mt and non-Mt Networks**

**Mt:** Less PTMs, AA

Nodes: PTMs (left) node size is freq magnitude, tRNAs (right)

Edges: active site freq, thickness magnitude of active site interactions
### Some of the Conclusions

<table>
<thead>
<tr>
<th>Protein Type</th>
<th>PTMs</th>
<th>Reaction Sites per PTM</th>
<th>Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mt</td>
<td>Few</td>
<td>Few</td>
<td>Sparse and organized</td>
</tr>
<tr>
<td>Non-Mt</td>
<td>Many</td>
<td>Many</td>
<td>Dense, disorganized and messy</td>
</tr>
</tbody>
</table>

Some of the Conclusions

<table>
<thead>
<tr>
<th>Particulars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein Type</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>Mt</td>
</tr>
<tr>
<td>Non-Mt</td>
</tr>
</tbody>
</table>
Some of the Conclusions

Overall

- PTMs reacting with AAs of few tRNAs (ex: Tryptophan W) also reacted with other reaction sites
- Study: Profile of Mt and non-Mt PTMs and reaction sites
- Future work: To study first effects of stress on protein reaction sites by PTMs
References: Cited in this Presentation


- Codon Usage Database: [http://www.kazusa.or.jp/codon/](http://www.kazusa.or.jp/codon/)
We would like to thank the support students, faculty and staff in the UNO-Bioinformatics Core Facility. 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).

Thank You! Questions?

obonhamcarter@unomaha.edu

IS&T Bioinformatics
http://bioinformatics.ist.unomaha.edu/