

Novel pipelines to extract differences in proteome dynamics based on health status

Alan A. Cohen

Biological Complexity and Healthy Longevity Associate Professor

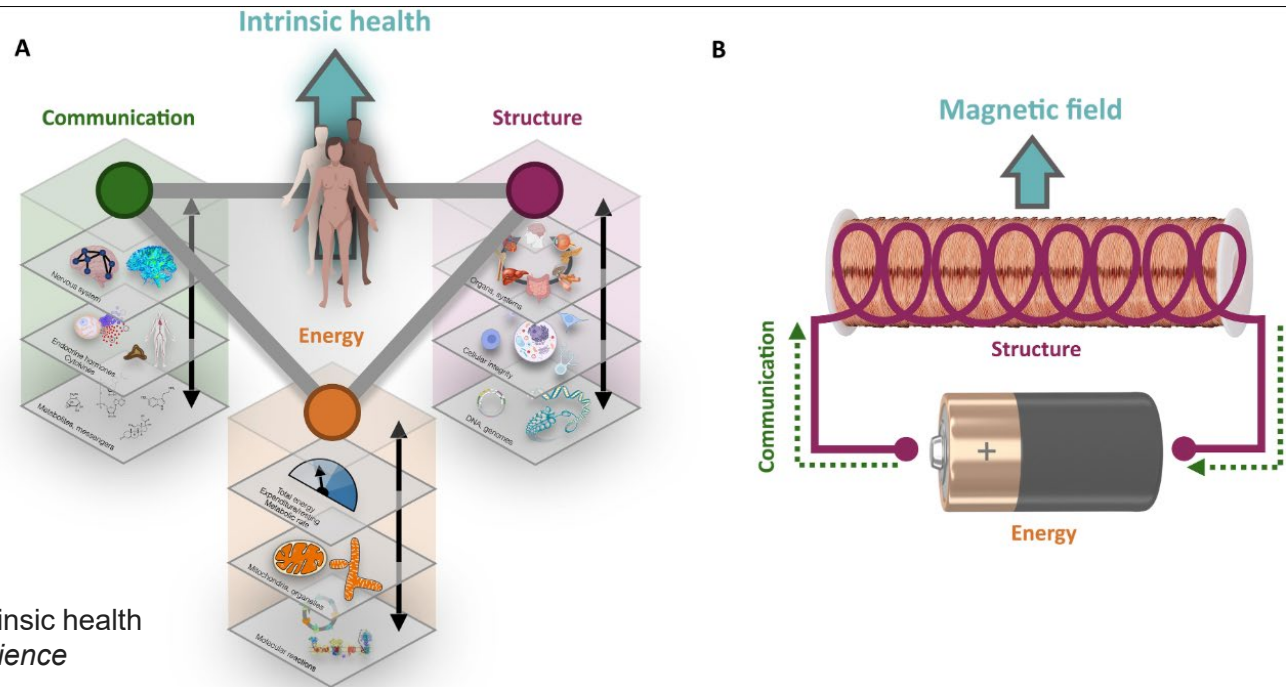
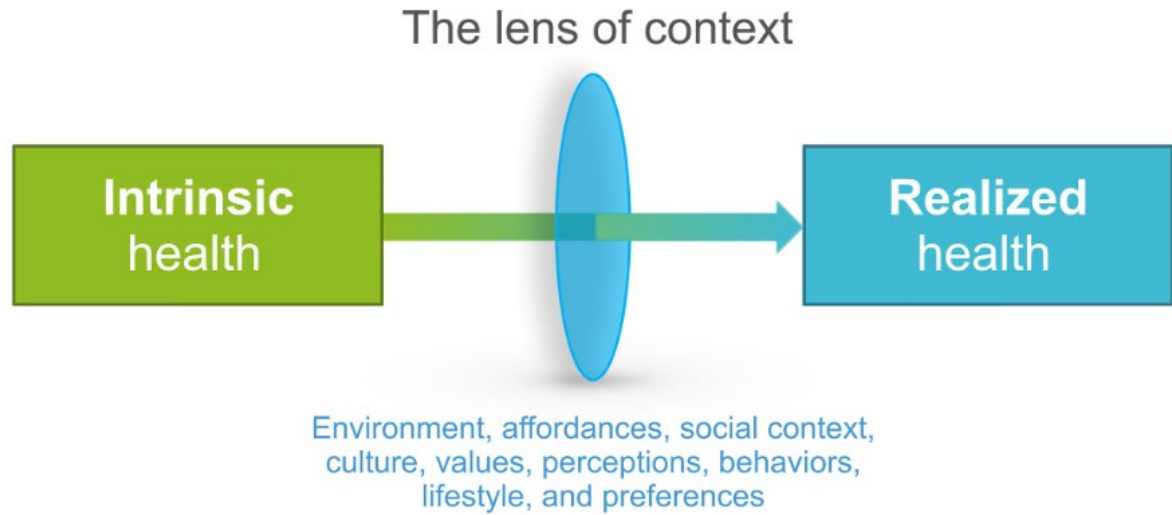
*Jiaying Zhao, Bowen Xu, Tianhui Huang,
Sewanou Honfo, Caroline Trumppff,
Martin Picard & Molei Liu*

*Columbia
University*

*Mailman School
of Public Health*

*Department of
Environmental
Health Sciences*

Context:
Can we
measure
intrinsic
health?



Cohen, A. A., Picard, M., et al. (2025). Intrinsic health as a foundation for a science of health. *Science Advances*, 11(25), eadu8437.

Recent publication:

- Zhao, Jiaying, Bowen Xu, Tianhui Huang, Sewanou Hermann Honfo, Caroline Trumpff, Martin Picard, Alan A. Cohen, and Molei Liu. "Examining saliva proteomic dynamics in mitochondrial diseases from a perspective of intrinsic health." *Scientific Reports* (2025).

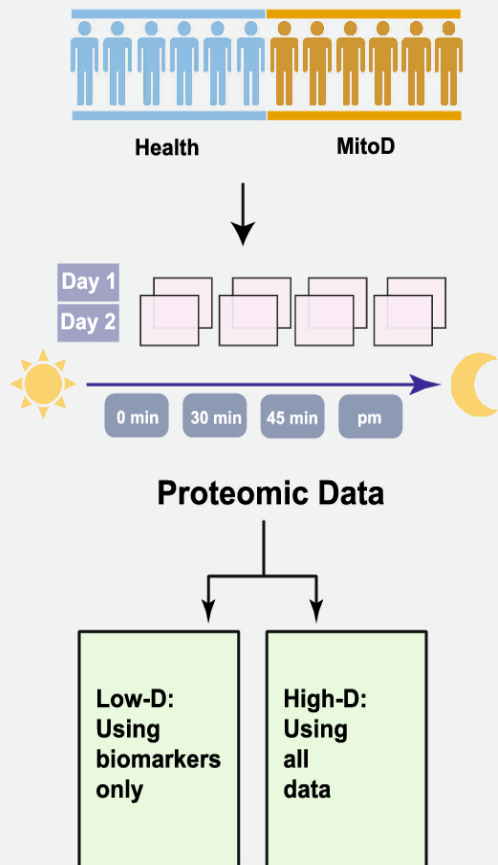
Question:

Is health status encoded in the dynamics of how the proteome responds to a stressor?

- Use data on the awakening response
- Saliva proteomics
 - Awakening, 30 mins, 45 mins, evening
 - Repeated over 2 days
- Small pilot study: six MitoD, six controls
- Use **unsupervised** methods to predict health status (MitoD versus control)

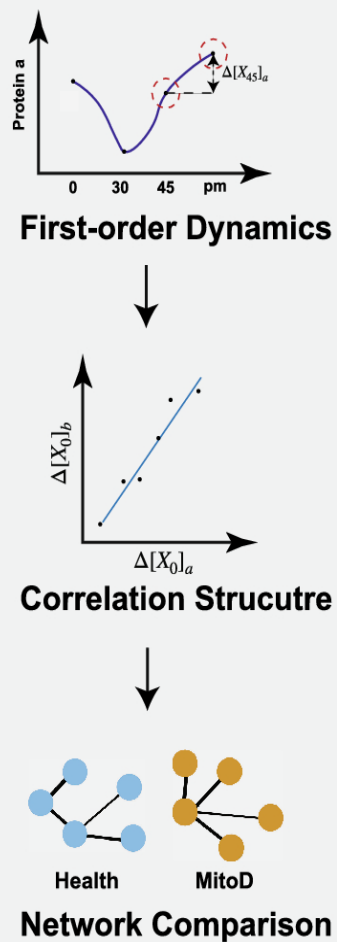
(a)

Data Structure



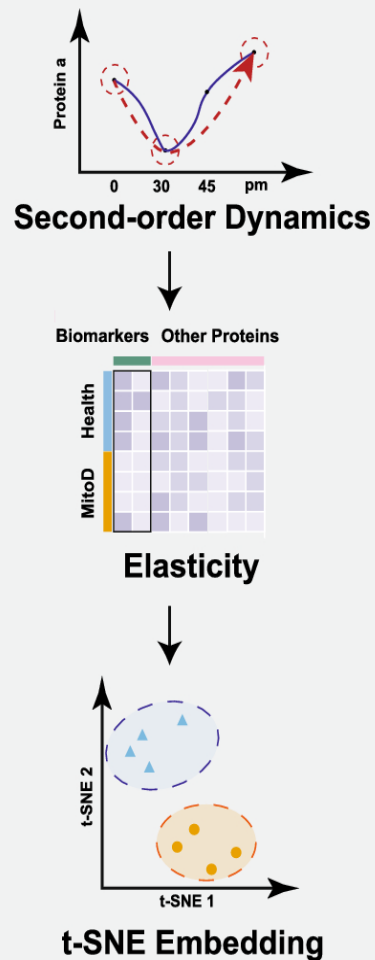
(b)

Delta Clustering



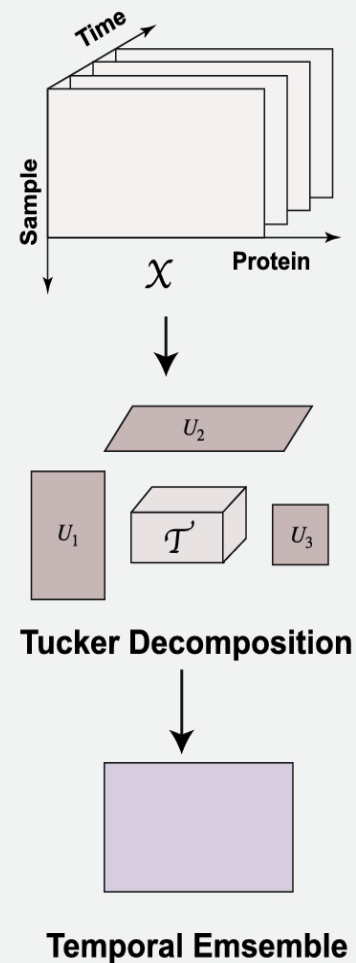
(c)

Proteome Elasticity

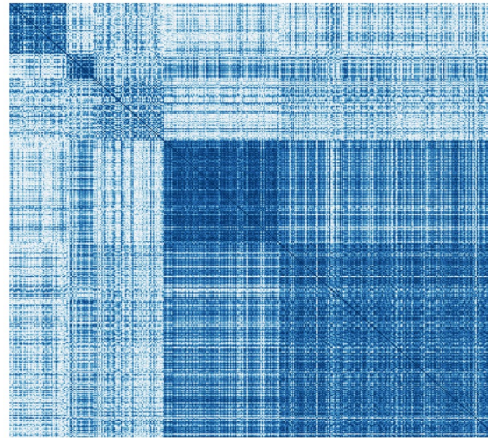


(d)

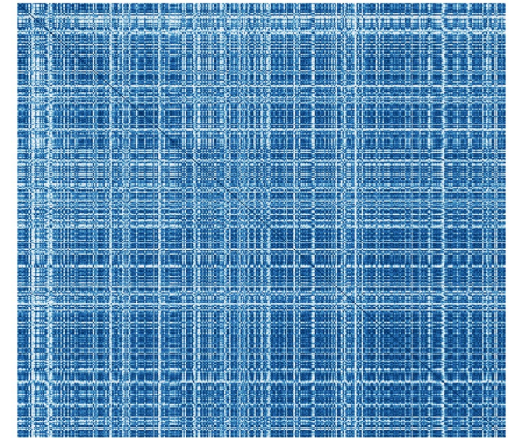
Tensor Decomposition



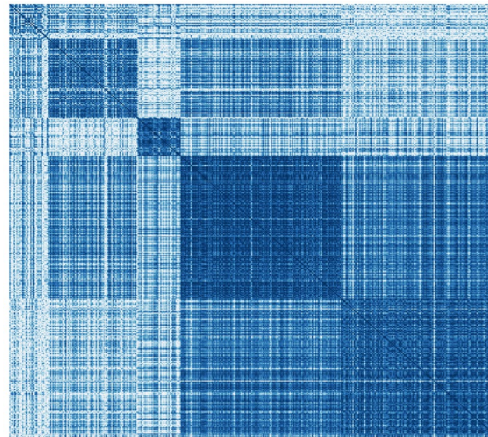
Delta clustering results



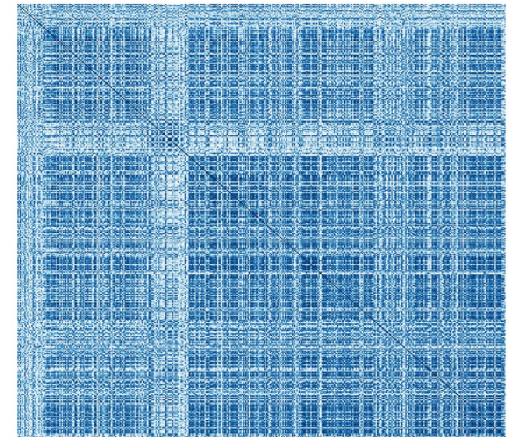
Control



MitoD



Random 6



Random complement

On the 2500 most variable proteins...

Delta clustering results

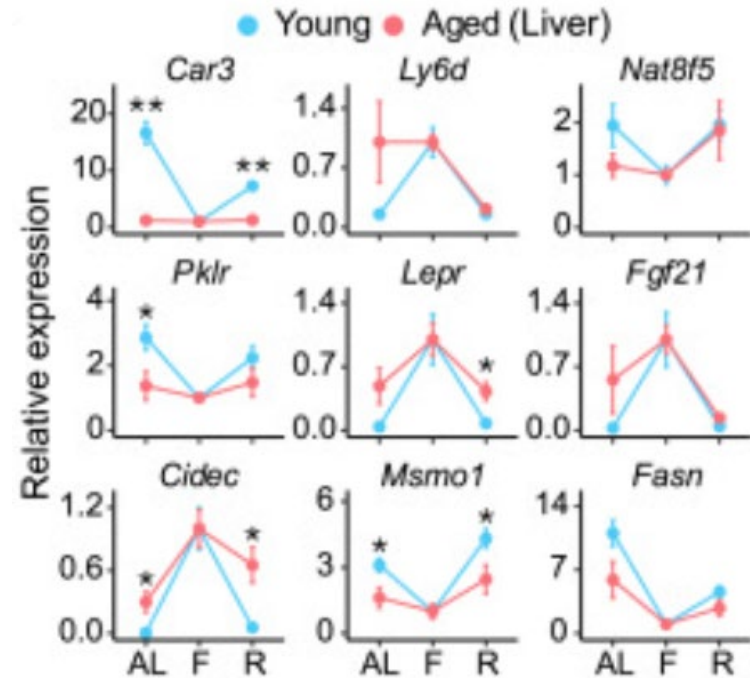
- Across all possible combinations of 6 – 6 divisions of 12 subjects, was the MitoD-Control separation the one that most lost the structure?
- For various parameter values, the p-value for this was always 0.0411
- The p-value using just GDF15 and IL-6 was always well above 0.05

elasticity = sign × amplitude × ratio

$$\text{sign} = \begin{cases} 0, & \text{if } (\mathbf{X}_{30,sj} - \mathbf{X}_{0,sj}) \cdot (\mathbf{X}_{pm,sj} - \mathbf{X}_{30,sj}) \geq 0 \\ 1, & \text{if } (\mathbf{X}_{30,sj} - \mathbf{X}_{0,sj}) \cdot (\mathbf{X}_{pm,sj} - \mathbf{X}_{30,sj}) < 0 \end{cases}$$

$$\text{amplitude} = |\mathbf{X}_{30,sj} - \mathbf{X}_{0,sj}| + |\mathbf{X}_{pm,sj} - \mathbf{X}_{30,sj}|$$

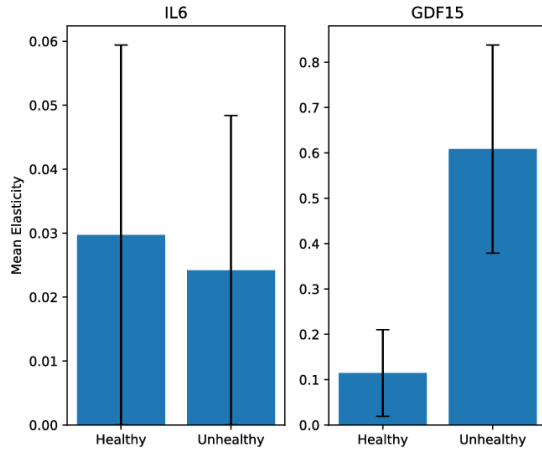
$$\text{ratio} = \frac{\min\{|\mathbf{X}_{30,sj} - \mathbf{X}_{0,sj}|, |\mathbf{X}_{pm,sj} - \mathbf{X}_{30,sj}|\}}{\max\{|\mathbf{X}_{30,sj} - \mathbf{X}_{0,sj}|, |\mathbf{X}_{pm,sj} - \mathbf{X}_{30,sj}|\}}$$



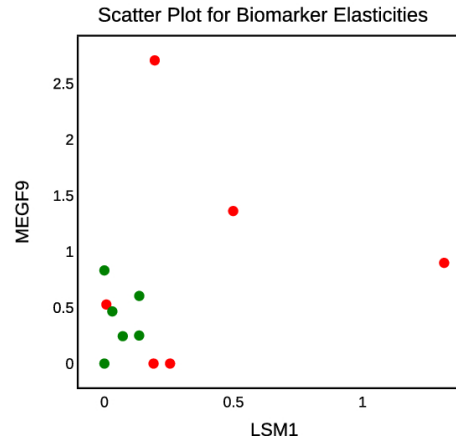
Elasticity

From Zhou et al. Cell Metabolism 2023

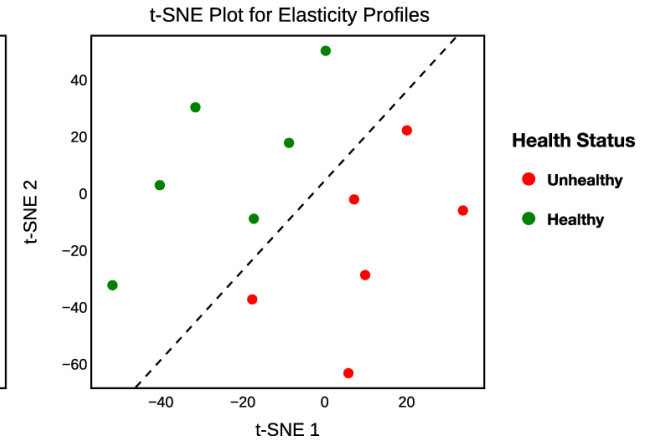
(a)



(b)



(c)

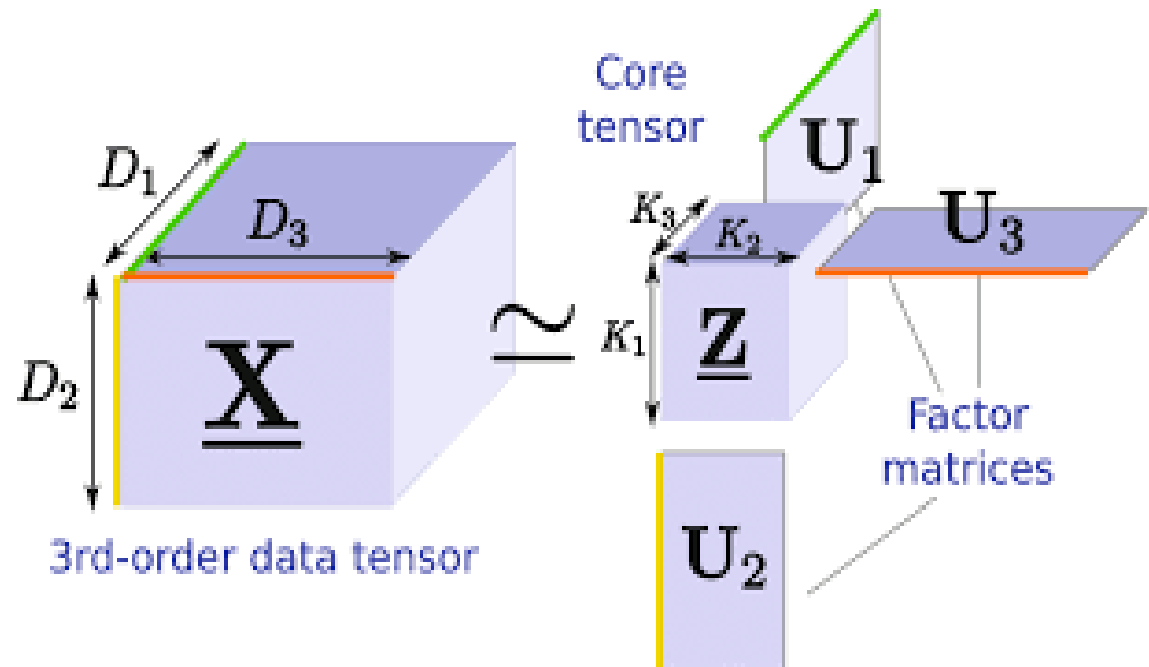


Elasticity

T-SNE of elasticities of all proteins (To, T₃₀, PM)

Part I: Representation learning and visualization

- Method
 - Tensor decomposition (e.g., Tucker decomposition): time * individual * protein.



Tensor decomposition results

	1st factor	2nd factor	3rd factor
0	0.75	0.66	-0.07
30	0.44	-0.53	-0.21
45	0.39	-0.35	0.81
pm	0.31	-0.41	-0.54

Interpretation:

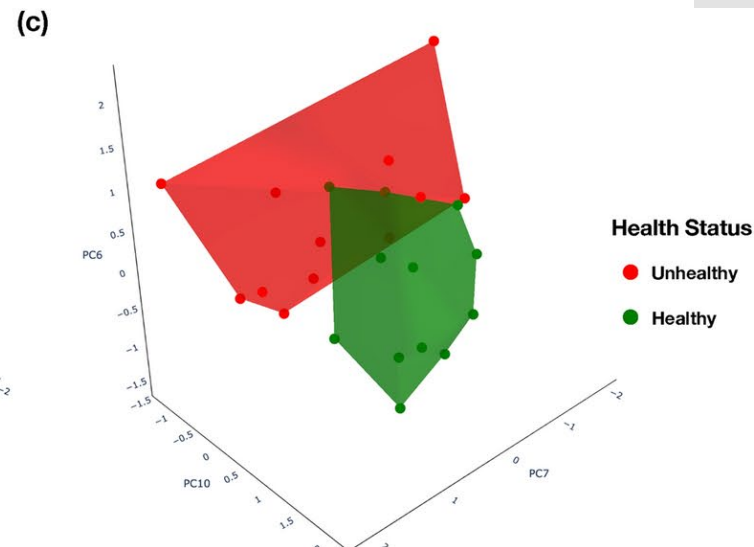
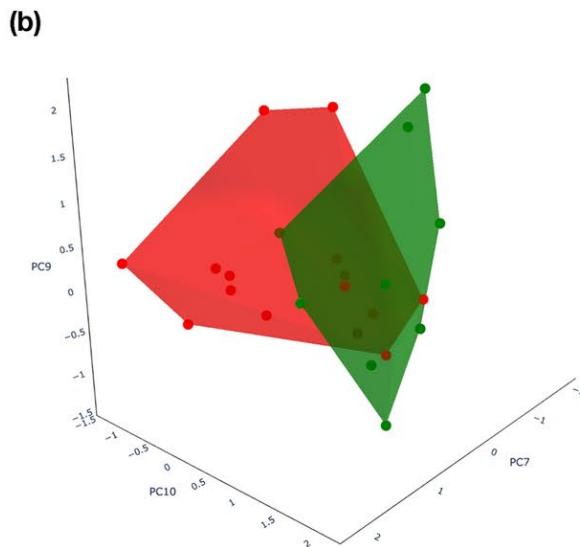
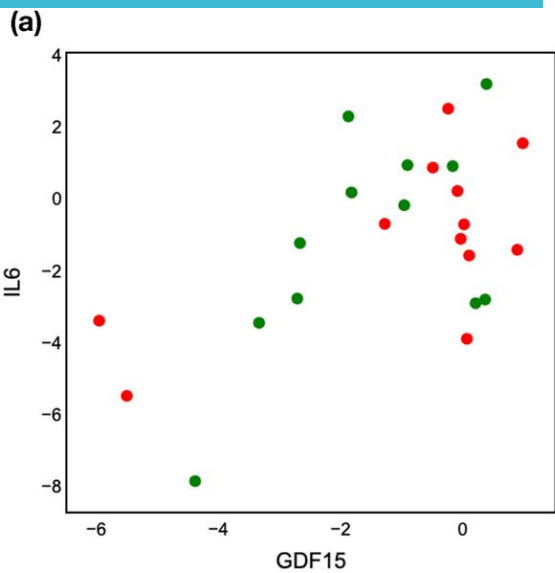
1st:
weighted
sum

2nd:
0 - mean(30, 45,
pm),
with weights

3rd:
45 - mean(30,
pm),
more weights on
pm

*Aggregate all time points:
weighted sum of values in
0, 30, 45, pm*

Tensor decomposition results



Implications

- The distinction MitoD-Control is *inherent* in the dynamics of the proteome
- Multiple pipelines to detect these differences can be developed
- The dynamics of the whole proteome perform better than on one or a few key molecules – it's the ensemble, not a specific mechanism
- Next steps:
 - Validate in other data sets with other stressors and health distinctions
 - Develop better pipelines
 - Find general signals of intrinsic health

Questions?

- Alan.Cohen@Columbia.edu