Differential Flux Balance Analysis of Quantitative Proteomic Data on Protein Interaction Networks

Biaobin Jiang$^1$* David F. Gleich$^2$ Michael Gribskov$^{1,2}$

$^1$Department of Biological Sciences,
$^2$Department of Computer Science, Purdue University

*Presenter, bjiang@purdue.edu

December 16, 2015
Outline

1 Introduction
   - Background
   - Problem Statement
   - Motivation

2 Methods
   - Flux Balance Analysis
   - Data: Colon Cancer

3 Results
   - Differential Analysis
   - Ego-centric Network Analysis
   - Test of Robustness

4 Conclusion and Acknowledgement
Background: Protein Interactome

- Proteins are basic functional units in biological systems.
- Proteins carry out biological functions through protein-protein interactions.

Problem Statement

**Ultimate Question:**
How does a copy of proteins chooses its binding partners?

- abundance;
- binding affinity;
- stoichiometry;
- post-translational modification;
- etc.
Problem Statement

**Feasible Question:** How does a copy of proteins chooses its binding partners *given proteome-wide abundance?*
Motivation

1. **Mutations** on protein binding sites rewire cancer signaling;
   (Creixell *et al.*, Cell, 163(1), 2015)

2. Protein interaction dynamics can be **experimentally** quantified by AP-SWATH in a small scale;
   (Collins *et al.*, Nature Methods, 10(12), 2013)

3. **Flux balance analysis** is widely used to reconstruct genome-wide biochemical reaction network.

**Aim**: To use **flux balance analysis** to estimate protein interaction dynamics caused by cancer-related **mutations**, which cannot be done **experimentally** in proteome-wide scale.
Flux Balance Analysis Model

Notations:
- \( x \in \mathbb{R}^n \), protein flux for each interaction;
- \( b \in \mathbb{R}^m \), input protein absolute copies;
- \( A \in \mathbb{R}^{m \times n} \), the incidence matrix of protein-interaction graph, where
  \( A(i, k) = A(j, k) = 1 \) if protein \( i \) interacts with protein \( j \);
- \( c \) is an all-one vector.

Goal: to maximize the total flux in the whole protein-protein interaction network. The flux balance analysis can be formulated as a linear programming model,

\[
\begin{align*}
\text{maximize} & \quad c^T x \\
\text{subject to} & \quad Ax \leq b \\
& \quad x \geq 0.
\end{align*}
\]
Model Illustration

Note: Usually, the number of interactions is much larger than the number of proteins (that means $A$ is a fat matrix).
Data from Quantitative Proteomics in Colon Cancer

- **Estimates of Protein Absolute Copy Number:**

- **Network Data:** BioGRID
Results (1): Differential analysis, abundance vs. flux

- Does an up-regulated (or down-regulated) protein definitely up-regulate (down-regulate) all of its interaction fluxes?

**Conclusion**: Not necessarily true.
Results (2): Ego-centric network analysis

- Test whether cancer-related proteins have higher fluctuation in protein flux than cancer-unrelated ones.
- Examine the rankings of 3 metrics of the 18 cancer-related proteins versus other cancer-unrelated ones.
Results (3): Test of robustness against network noise

- Test whether the noise of network data affects the result (2) by:
  - randomly removing interactions (false positive);
  - randomly shuffling protein labels (random control); and
  - randomly adding interactions (false negative).

![Graph showing AUROC vs. Percentage of Altered Edges]

- AUROC values range from 0.5000 to 0.7327.
- The graph compares AUROC values at different percentages of altered edges (from -30% to 30%).
Take-home messages:

- Up/down-regulation of a protein abundance does not necessarily lead to up/down-regulation of its interaction fluxes.

- Cancer-related proteins, affected by genetic mutations, tend to have higher fluctuation in the fluxes of their ego-centric networks.

- Differential flux balance analysis are robust to noisy network data.
Future Plan

- Test the model using a large number of data sets (NCI-60 proteomes);
- Extend the model from interaction-wise to pathway-wise dynamics;
- Analyze flux fluctuation of crosstalk between pathways.

Image resource: wzw.tum.de/proteomics/nci60/.
Acknowledgement

I would like to thank,

- my advisor Dr. Gribskov, Purdue Biological Sciences Department;
- my mentor Dr. Gleich, Purdue Computer Science Department; and
- all my labmates.
The End