Assignment 5

Topics: Systems biology, temporal dynamics of gene expression, multi-omics integration, and biological networks

Due date: 3 November 2025 at 11:59pm

Rules:

- You can work in group, but write your own answers
- You can use AI to help, but don't abuse it. Credit AI when used
- The objective of the assignment is to provide you with experience. Explain your work and observations. Don't just paste a screenshot of the result.
- You can contact me to ask for clarification

Credit: GPT-5 was used to aid the design of the assignment

Part A. Simulation of gene expression temporal dynamics

We will use Microsoft Excel (or if you are familiar with some programming languages, feel free to use them) to perform the simulation.

Attach graphs of your simulation result to support your answer whenever possible.

The system contains two genes, A and B, which regulate each other according to the following dynamics:

$$\frac{d[A]}{dt} = k_1 - k_2[A] - k_3[A][B]$$

$$\frac{d[B]}{dt} = k_4[A] - k_5[B]$$

For your first simulation, use the following parameter values:

$$k_1 = 1, k_2 = 0.5, k_3 = 0.01, k_4 = 0.5, k_5 = 0.5$$

1. For Excel, set up your table as follows:

Time (t)	[A]	[B]	d[A]/dt	d[B]/dt
0	0.5	0.1		
1				
...				

2. Next, calculate **d[A]/dt** and **d[B]/dt** using the value of current [A], [B], and other parameters.

Time (t)	[A]	[B]	d[A]/dt	d[B]/dt
0	0.5	0.1	0.7475	0.2
1				
...				

3. Next, update the values of [A] and [B] for the next time step.

Time (t)	[A]	[B]	d[A]/dt	d[B]/dt
0	0.5	0.1	0.7475	0.2
1	1.2475	0.3		
...				

4. Repeat the process until you observe that [A] and [B] converged to an equilibrium. What is the behavior of this system at equilibrium?
5. How would you change the system parameters (*k*'s) to achieve the following:
    a. Higher [A] at equilibrium
    b. Significantly more [A] than [B] at equilibrium
6. How many equilibria does this system have? If you cannot prove it analytically, you can try changing the initial [A] and [B] to explore the system.
7. If the system is changed to:

$$\frac{d[A]}{dt} = k_1 - k_2[A] + k_3[A][B]$$

so that the A-B complex now activates A's transcription. Using the same parameter values as specified above, can you identify a new behavior of the system?
Hint: Try large values for the initial [A] and [B].

Part B. Literature Analysis (Systems Biology)

Study this landmark paper from 2001, science.292.5518.929, "Integrated Genomic and Proteomic Analyses of a Systematically Perturbed Metabolic Network" by Ideker, T. et al., and answer the following questions:

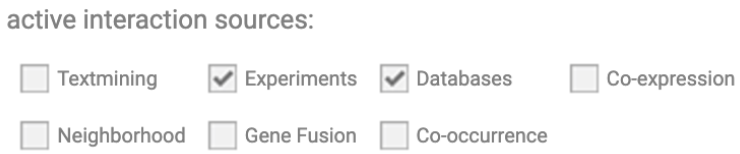
- What are the different types of biological data combined together in this study?
- How did the authors combined multiple data types to understand the biology of the galactose metabolism pathway?

- Interpret Figure 2B. How well did the authors' model predict the effect of gene deletion? Are there disagreements?

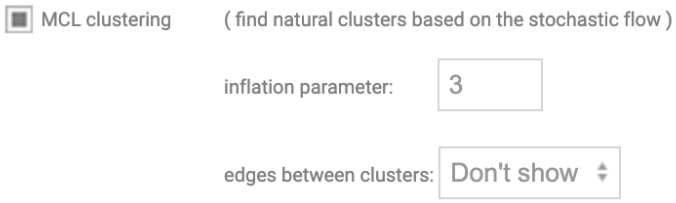
Part C. Exploring Biological Networks

Suppose you are interested in proteins and molecular interactions involved in Parkinson's Disease.

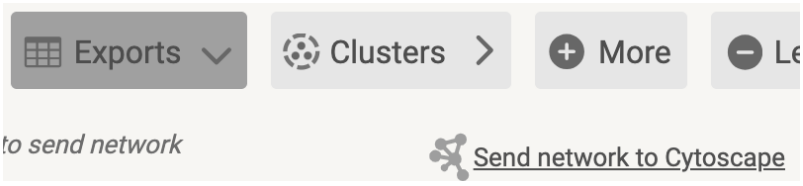
1. Go to STRING web tool at https://string-db.org/
2. Search for protein interaction network involved in Parkinson's Disease through their **Pathway / Process / Disease** tab
3. Remove predicted interactions through **Settings**. Retain only **Experiments** and **Databases** as shown.



4. Cluster proteins into broad functional groups using the **Clusters** tool as shown.



5. Report the result. What are the major protein functions found? How would you interpret their relationship to Parkinson's Disease.
6. Explore the **Analysis** tool.
    a. What is the average clustering coefficient of this network? Is it high or low?
    b. What other diseases share related proteins with Parkinson's Disease?
7. Set up **Cytoscape** (https://cytoscape.org/) on your computer. Install **stringApp** to ensure that Cytoscape can communicate with STRING website.
8. Keep **Cytoscape** open, go back to **STRING**, enter the **Export** tab and click on **send network to Cytoscape**. You may need to update **Cytoscape** and **stringApp**.



9. Run the **NetworkAnalyzer** app on Cytoscape to calculate topological properties of the network (treat the network as undirected) and answer the following questions:
    a. Which proteins have the highest **betweenness** scores? Does it make sense for them to have high betweenness scores based on their locations in the network?
    b. Which proteins have the highest **closeness** scores? Does it make sense for them to have high closeness scores based on their locations in the network?