Types of biological networks

I. Intra-cellular networks
Some intra-cellular networks:

1. Metabolic networks
2. Transcriptional regulation networks
3. Cell signalling networks
4. Protein-protein interaction networks
5. ...
Metabolic networks

• A network of biochemical reactions in a cell
• Partially experimental, partially reconstructed from genome sequence – see the WIT paper from *Nucleic Acids Research*, 2000
• Available for many organisms, from bacteria to human
• Available on-line:
  – KEGG (Kyoto Encyclopedia of Genes and Genomes) – info on genes, proteins, reactions, pathways for eukaryotes and prokaryotes
  – GeneDB – contains similar info
  – BioCyc, EchoCyc, MetaCyc – more specialized info on particular species
  – WIT, renamed to ERGO
Metabolic networks
Metabolic networks

• Used for studying and modeling *metabolism*:  
  – a set of chemical reactions that happen in living organisms that allows an organism to:  
    • respond to the environment,  
    • grow,  
    • reproduce,  
    • maintain its structure

• Modeling from *network topology* (topic of this class) to predictive toxicology

• Consist of *metabolic pathways*:
  – Series of chemical reactions occurring within a cell,  
  – Catalyzed by *enzymes* -- proteins that regulate chemical reactions  
  – Results in a *metabolic product* to be used/stored in the cell, or  
  – Initiation of another metabolic pathways (called *flux generating step*)
Metabolic networks

Example from Viz4All, University of Maryland, College Park:
http://www.cs.umd.edu/class/spring2006/cmsc838s/viz4all/v4a_vis.html
Metabolic networks
Metabolic networks
Metabolic networks
Metabolic networks

- Metabolic pathways of a cell form a *metabolic network*
- Metabolic pathways include the main chemical, mostly enzyme-dependant reactions needed to keep an organism in *homeostasis*, which is an internal regulation that maintains a stable, constant condition of a living system.
- Directed edges are drawn between *enzymes* (proteins that catalyze (accelerate) chemical reactions) and *substrates* (molecules acted upon by an enzyme).
- Thus, enzymes and substrates correspond to *nodes*, *directed edges* to metabolic reactions in a metabolic network.
Metabolic networks

Further readings:


Transcriptional regulation networks

• Model regulation of *gene expression* (gene regulation)

• *Gene regulation* is a process by which information from genes is turned into *gene products* (*RNA* or *protein*)

• Gene regulation gives a cell control over its structure and function, e.g.:
  - *Cellular differentiation* (a process by which a cell turns into a more specialized cell type)
  - *Morphogenesis* (a process by which an organism develops its shape)
  - Adaptability...
Transcriptional regulation networks

- Genes are the nodes and the edges are directed:

- Transcription factor protein $X$, binds regulatory DNA regions of a gene to regulate (stimulate or repress transcription of a gene) the production rate of protein $Y$
Transcriptional regulation networks

Problem: Stimulation and repression of gene transcription are both represented the same way in the network.

Available for model organisms = non-human species manipulated and studied to get insights into workings of other organisms:

- *Saccharomyces cerevisiae* - baker's yeast (see Milo *et al.*, 2002 from the class readings)
- *Escherichia coli* (see Shen-Orr *et al.*, 2002)
- Sea urchin (see Davidson *et al.*, 2002)
- Fruitfly *D. melanogaster*
The Central Dogma of Molecular Biology

Replication
DNA duplicates

Transcription
RNA synthesis

Translation
Protein synthesis

DNA
RNA polymerase
mRNA
Cytoplasm
protein
Cell signaling networks

• *Cell signaling* is a complex system of communication that governs basic cellular activities and actions, e.g., development, repair, immunity etc.

• The cell converts one kind of signal to another

• Errors in signaling cause serious diseases, e.g., cancer, autoimmune diseases, diabetes etc.

• These networks consist of *signaling pathways*:
  – Ordered sequences of biochemical reactions in a cell
  – Changes induced by *receptor* (protein that receives and responds to a stimulus) activation

• In these networks, proteins are the *nodes* and the *edges* between them are *directed*

• In this course – *systems biology* view of signaling
The entire set of changes induced by receptor activation is called a *signal transduction pathway* / *signaling pathway*
Cell signaling networks

Famous examples (lots of literature on them):

1. Mitogen-activated protein kinase (MAPK) pathway (originally called “ERK” pathway)
   - MAPK protein is an enzyme, a protein kinase, which can attach phosphate groups to target protein, (such as to a transcription factor)
   E.g.:
     - MYC is an oncogene transcription factor (oncogene is a gene which when mutated or expressed at high levels helps turn a normal into a tumor cell) expressed in a wide range of human cancers
       - MAPK can phosphorilate (attach phosphate group) MYC and so alter gene transcription and cell cycle progression
     - EGFR = “epidermal growth factor receptor”
       - activates MAPK pathway
       - mutations affecting its expression/activity can result in cancer
Cell signaling networks

2. *Hedgehog signaling pathway*
   - One of the key regulators of animal development
   - Conserved from fly to human
   - Establishes basis of fly body plan
   - Important during *embryogenesis* (the process by which the embryo develops) and *metamorphosis* (from larva to pupa to adult)

3. *TGF-beta signaling pathway*
   - The “transforming growth factor” (TGF) signaling pathway
   - Involved in:
     - cell growth,
     - cell differentiation
     - *apoptosis* (programmed cell death)
Protein-protein interaction networks

- A *protein-protein interaction (PPI)* usually refers to a physical interaction, i.e., binding between proteins.

Protein structure:
Protein-protein interaction networks

- Can be other associations of proteins such as functional interactions – e.g., synthetic lethality (see next class)
- PPIs are very important for structure and function of a cell:
  - participate in signal induction
  - play a role in many diseases (e.g., cancer)
  - can be stable interactions to form a protein complex
    (a form of a quaternary protein structure, set of proteins which bind to do a particular function, e.g., ribosome, hemoglobin – illustrated below)
Protein-protein interaction networks

• can be *transient interactions*, brief interactions that modify a protein that can further change PPIs
  – e.g., protein kinases, add a phosphate group to a target protein.
  – A protein can carry another protein (e.g., *nuclear pore importins* – proteins that carry other proteins from cytoplasm to nucleus and vice versa)

• Transient interaction form the *dynamic part of PPI networks*

• Some estimates state that about 70% of interactions is stable and 30% are dynamic
Protein-protein interaction networks

PPI are essential to almost every process in a cell.

Thus, understanding PPIs is crucial for understanding life, disease, development of new drugs (most drugs affect protein-protein interactions).

Methods to detect PPIs

- Biological and computational approaches.
- None are perfect, i.e., high rates of false positives (interactions present in the data sets that are not present in reality) and false negatives (missing interactions)
Protein-protein interaction networks

Methods to detect PPIs (continued):

1. Co-immunoprecipitation (CoIP, “Pull-down”):
   – The protein of interest is isolated with an antibody specific to the protein.
   – Interacting partners (proteins) of the protein, that stick to the protein of interest, are then identified by:
     • *gel electrophoresis techniques*, which separate proteins by mass, e.g.,
       – Western blotting
       – Mass spectrometry (see below)
     • They identify constituents of a protein complex
       – Verifies interactions between suspected interacting partners
       – Not a screening approach
Protein-protein interaction networks

Methods to detect PPIs (continued):

• Problems:
  – “spoke” model (bipartite graph $K_{1,n}$) versus
  – “matrix” model (complete graph, $K_n$) of PPIs in a protein complex?

Schematic:
Protein-protein interaction networks

Problems:

"Spoke" model

V.S.

"Matrix" model

by complete graph:
edges between all
proteins
Protein-protein interaction networks

Methods to detect PPIs (continued):

2. Yeast Two-hybrid screening (Y2H)

- Investigates interaction between artificial (genetically engineered) fusion proteins (created from two or more proteins or peptides), one to a reporter gene (a gene attached to another gene) and the other to a transcription factor.
- If the interaction exists, a reporter gene is transcriptionally activated.
- It happens in yeast nucleus.
- One protein (in PPI) is “bait”, the other is “prey”
Protein-protein interaction networks

- Potential problem:
  - Interest in a particular pathway of, say 15 proteins
  - These 15 proteins are all “bates”
  - There is an order of magnitude more “prays”
  - This imposes a particular structure on the PPI network by experimental design without reflecting the underlying network topology

Thus, a matrix of $n \times n$ needs to be probed, where each bait is also a prey.
Protein-protein interaction networks

- Pros of Y2H:
  - Scalable

- Cons of Y2H:
  - High false positive and negative rate (as high as 50%)
  - Thus, our computational methods must be robust to noise in the data
  - Why so much noise?
    Because Y2H investigates interactions between:
    - over expressed
    - artificial, fusion proteins
    - in the yeast
    - in the yeast’s nucleus

*Each of these steps is noisy.*

E.g.: proteins need to be in their native environment, not in nucleus
Thus, PPIs between membrane proteins are underrepresented.