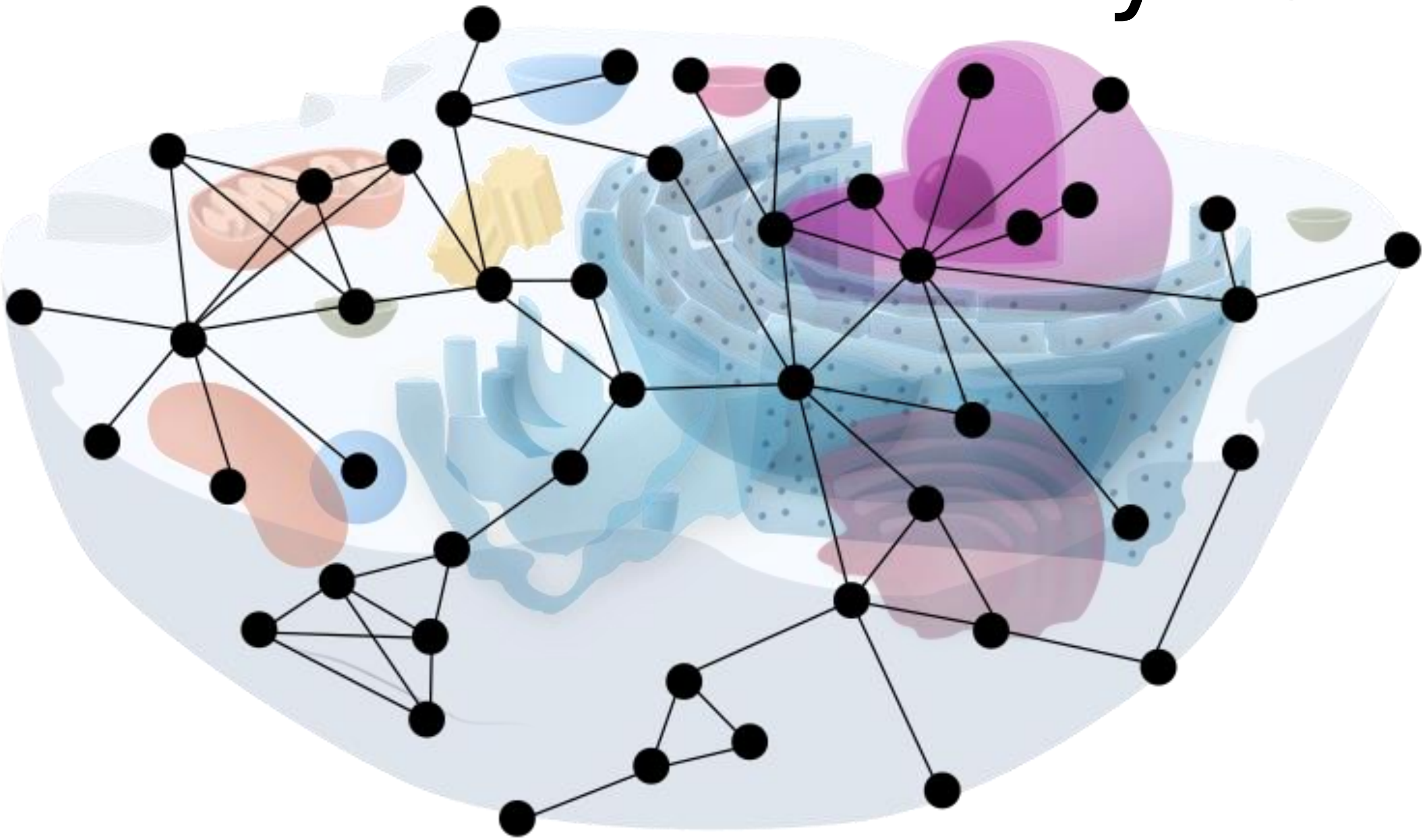
# Networks Across Biological Scales: From molecules to Ecosystems



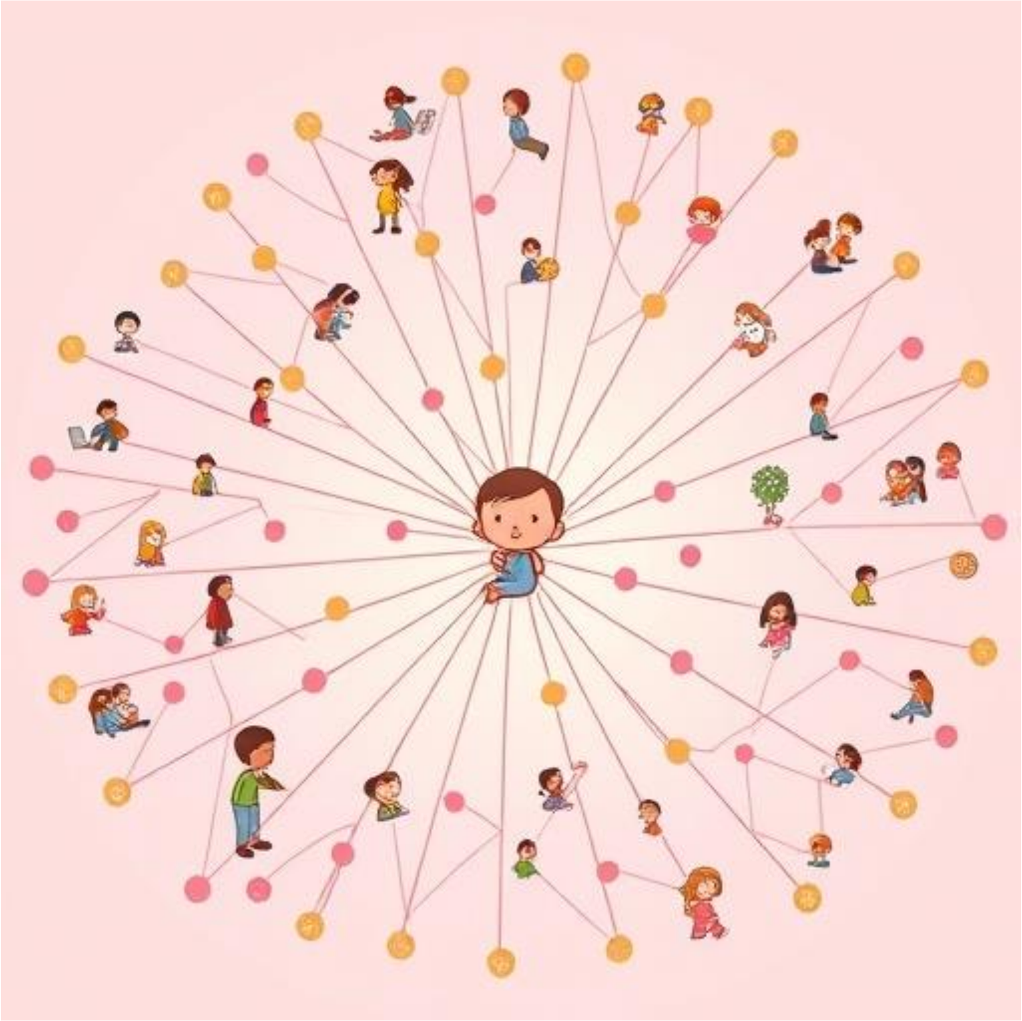
Ankit Agrawal
University of Würzburg
Germany

# Outline agenda

- Biological networks are ubiquitous
- Basic component of a network
- Constructing and modeling networks
- Properties of networks
- Available data for network analysis
- Demonstration: A practical example of network analysis
- Thoughtful experiment of networks
- Conclustion and future directions in biological networks

QR

# Social networks



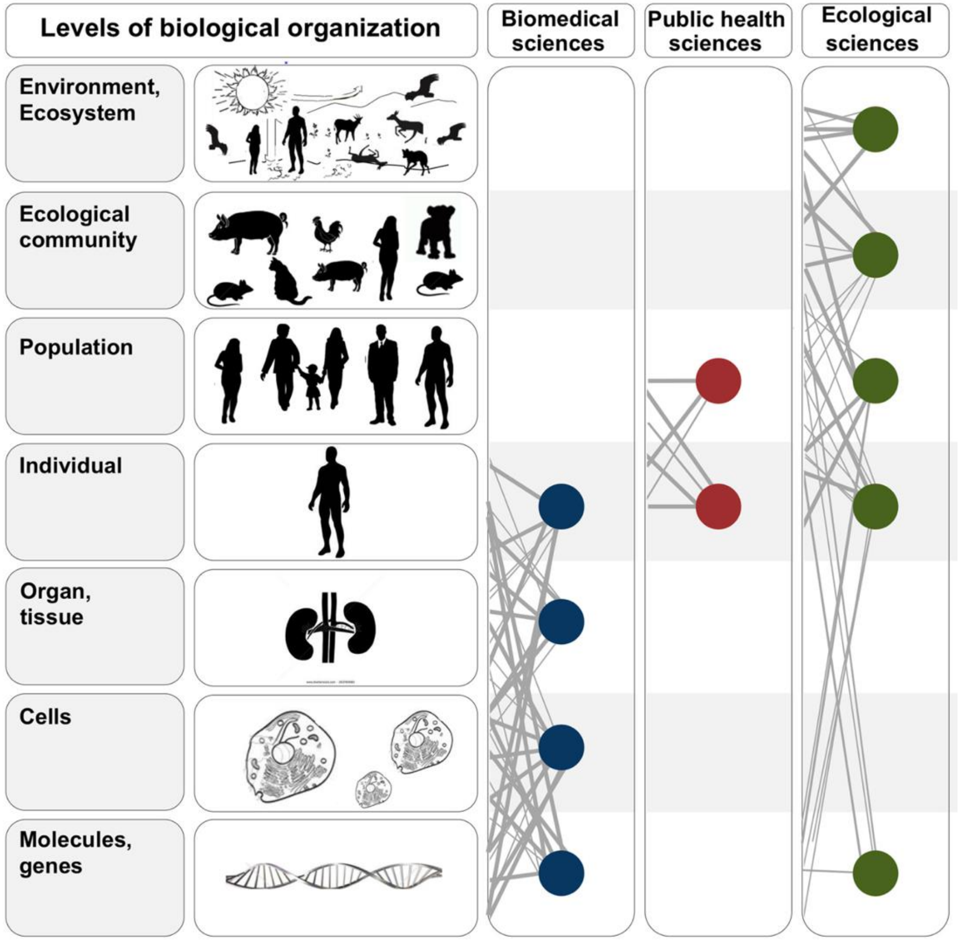
https://deepai.org/machine-learning-model/text2img

Prompt

Create a cartoon-style illustration showing the evolution of a person’s social network from infancy to adulthood. Start with a baby in the center who first connects with their parents and siblings. As the child grows, add connections to extended family (grandparents, aunts, uncles), followed by friends from school, teachers, and neighbors. Then, show young adult connections branching out to university friends, mentors, and colleagues. Make the network web grow larger and more interconnected with each stage of life, symbolizing social expansion

# Hierarchy of biological organization, from molecules and genes to ecosystems

Concentrates on the structural and topological organization of biological systems




Vanessa O. Ezenwa Plos Pathogens 2015

Focuses on the holistic understanding of biological systems, emphasizing how components interact to produce emergent behaviors

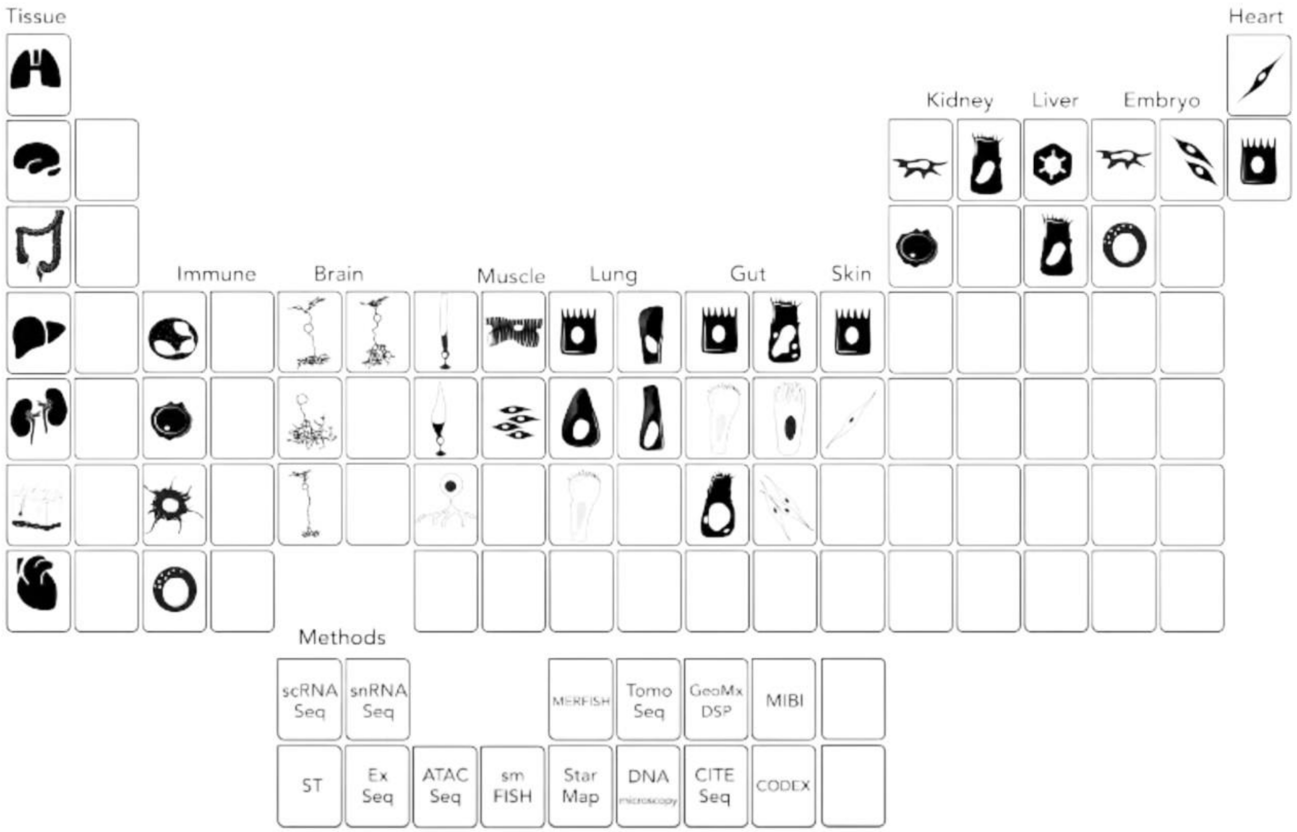
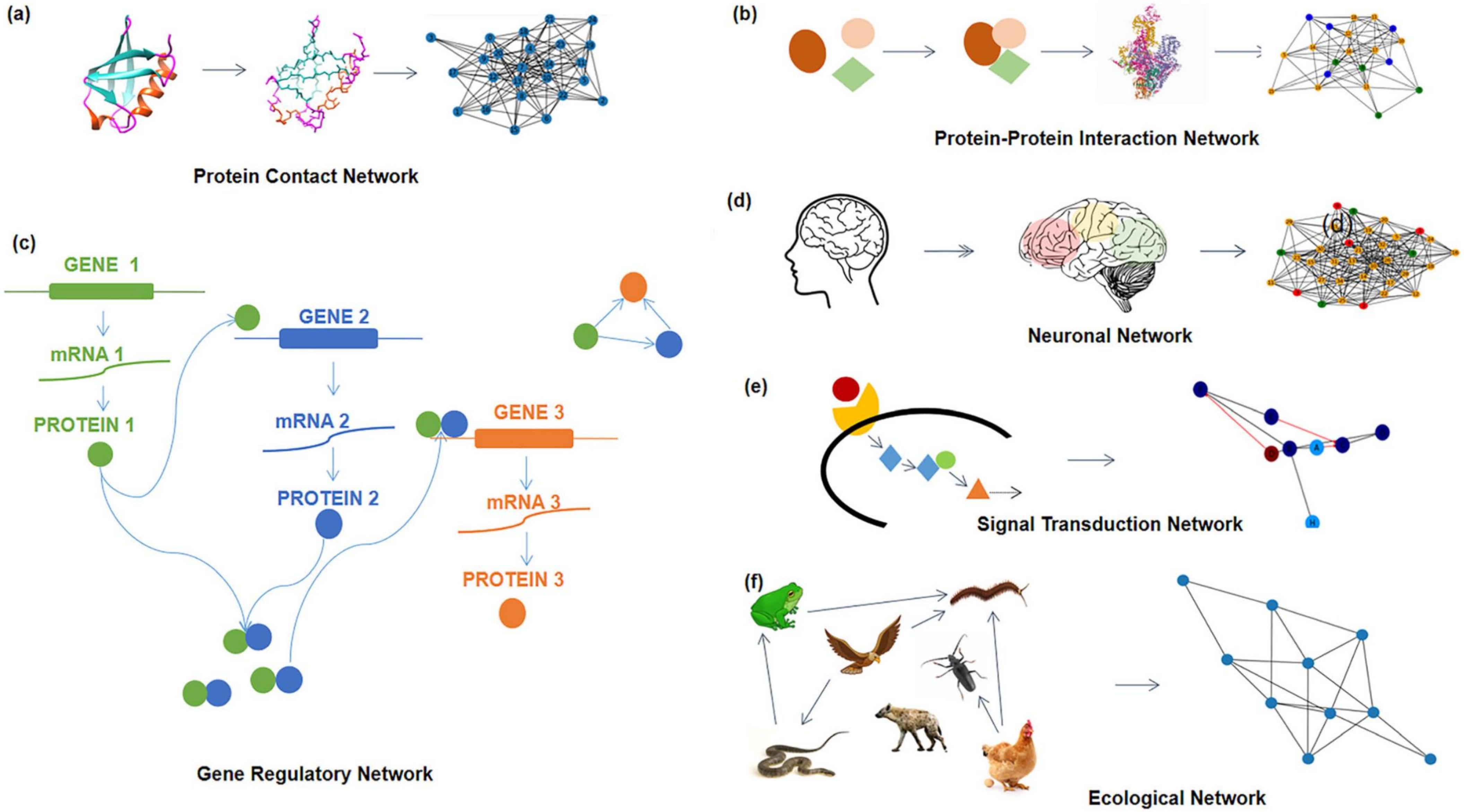



Image credit: Fabian Theis Lecture

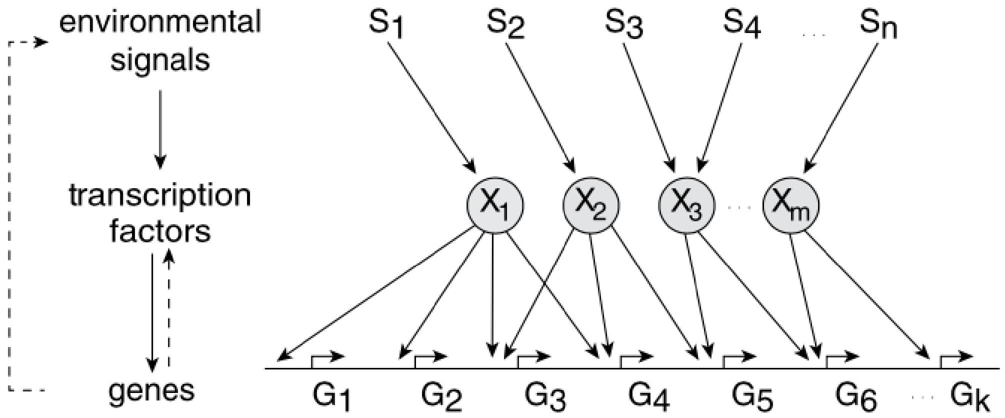
# Types of Biological Network



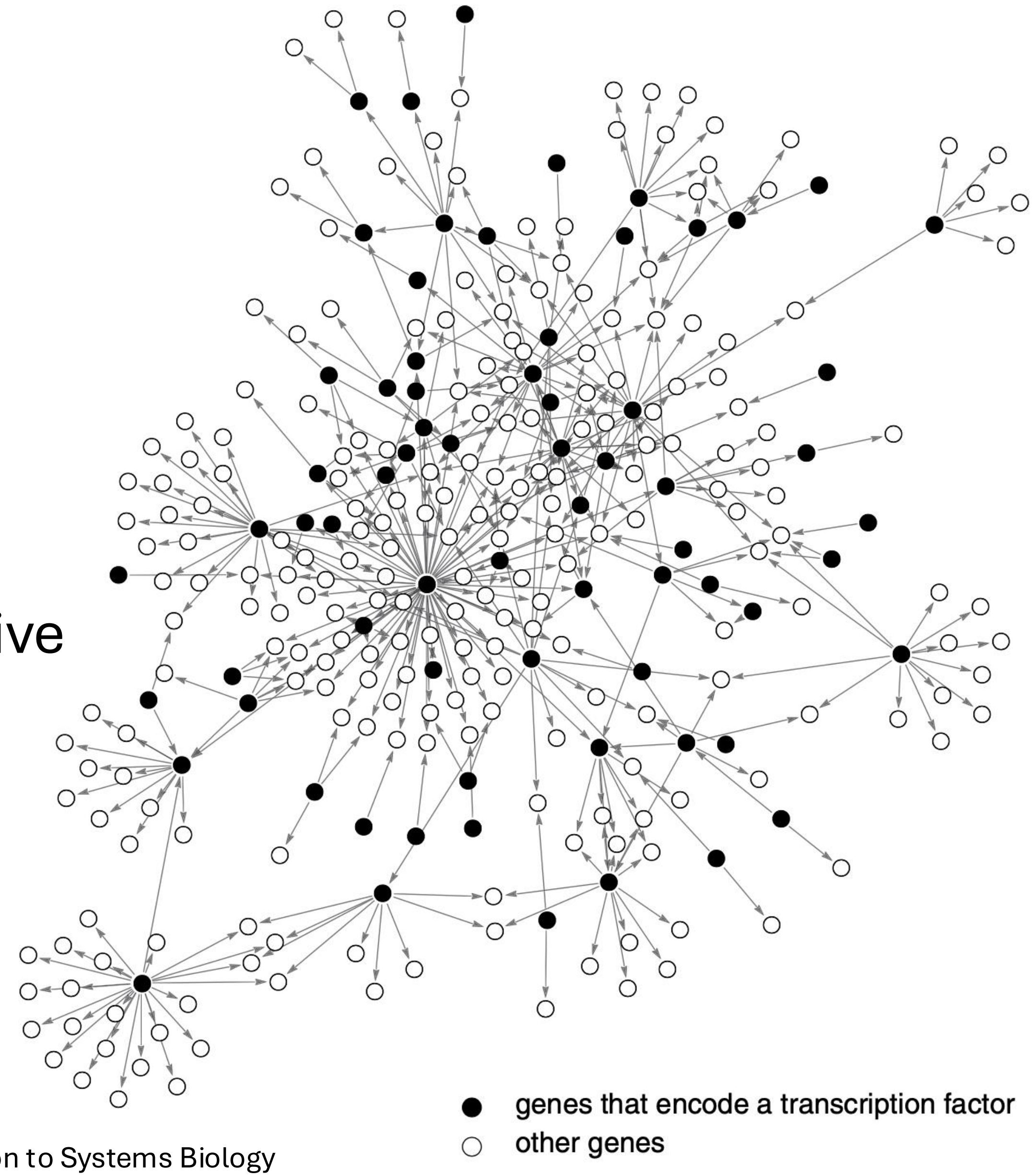

S Chaudhuri and A Srivastava J Biosci 2022

# Cells are living machines

- Cells sense many different signals
  - Temperature and pressure
  - Damage
  - Signal from other cell
  - Nutrients and chemicals
- Cells respond to these signals by producing appropriate proteins
- These special protein called TF
- TF are designed to transit rapidly between active and inactive states




**Transcription network**

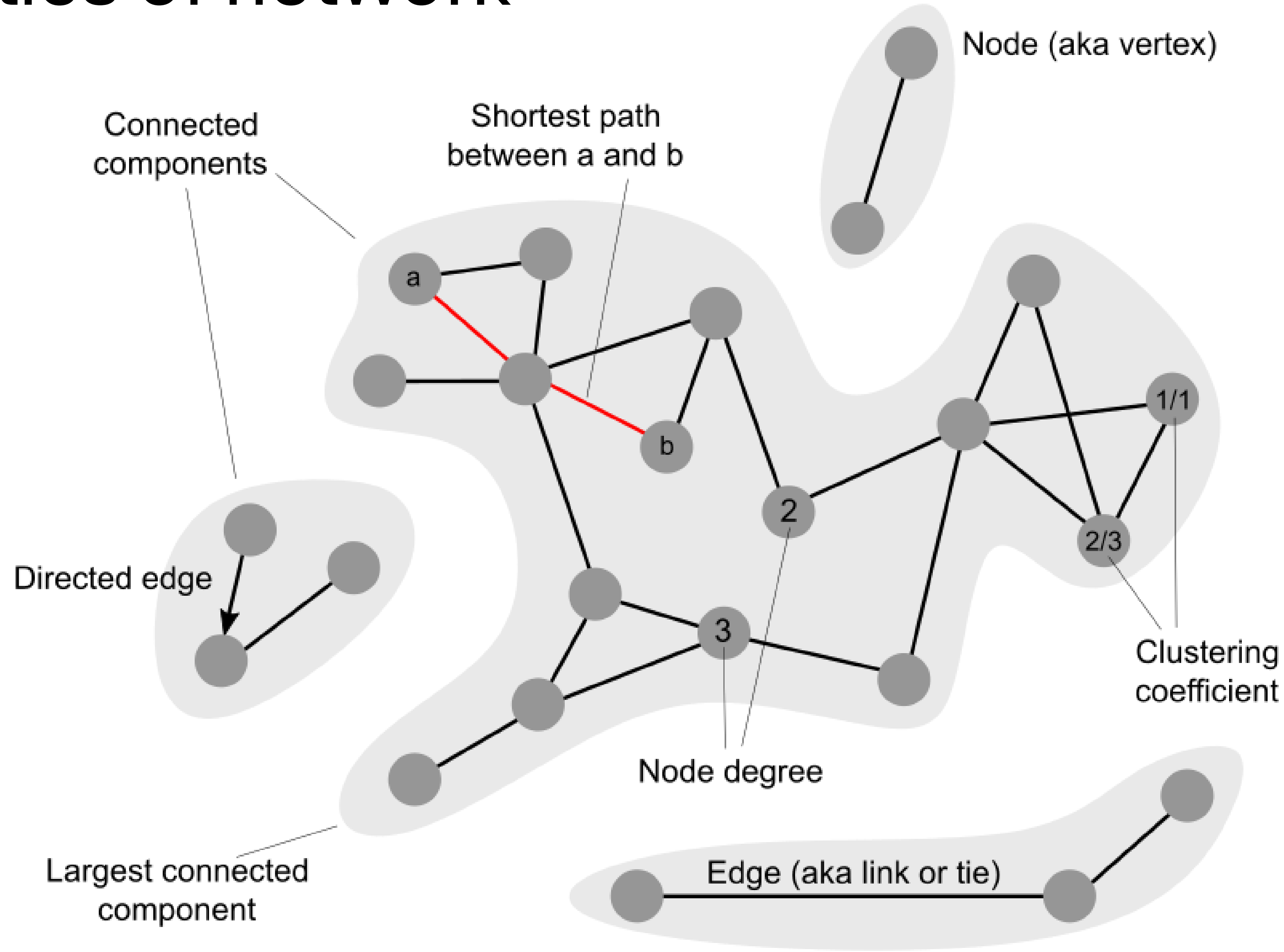



Uri Alon An Introduction to Systems Biology
Design Principles of Biological Circuits-CRC Press

# Outline agenda

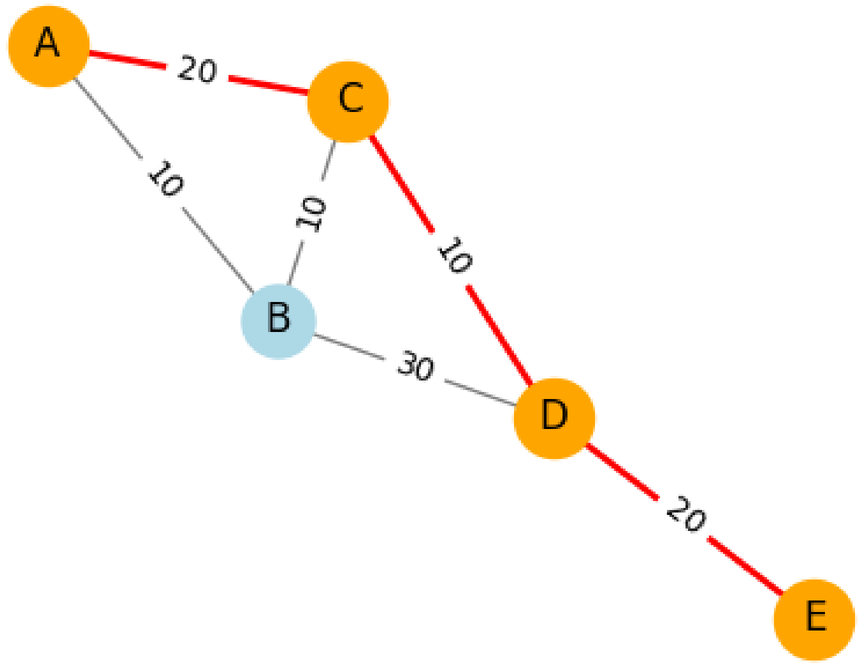
- Biological networks are ubiquitous
- **Basic component of a network**
- Constructing and modeling networks
- Properties of networks
- Available data for network analysis
- Demonstration: A practical example of network analysis
- Thoughtful experiment of networks
- Conclustion and future directions in biological networks

# Properties of network

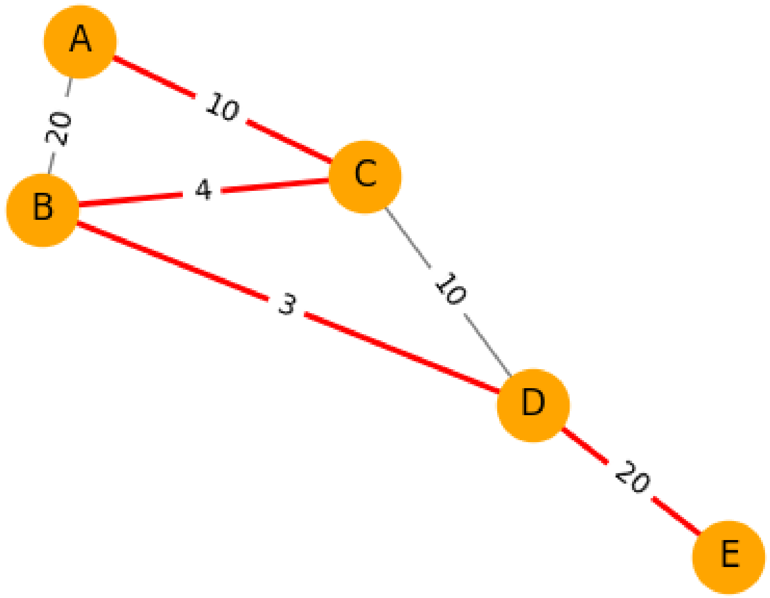



https://cbdm-01.zdv.uni-mainz.de/~galanisl/netbio/

# Shortest path in weighted graph

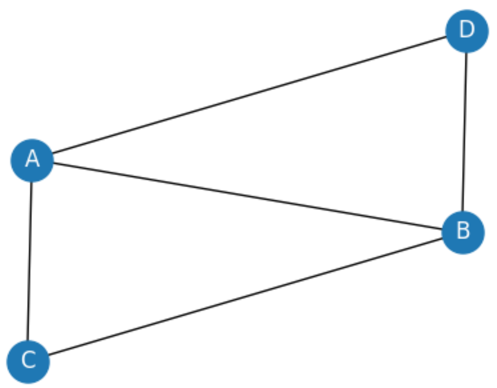



- Shortest Path:
  ['A', 'C', 'D', 'E']
- Shortest Path Length: 50




- Shortest Path:
  ['A', 'C', 'B', 'D', 'E']
- Shortest Path Length: 37

# Centrality

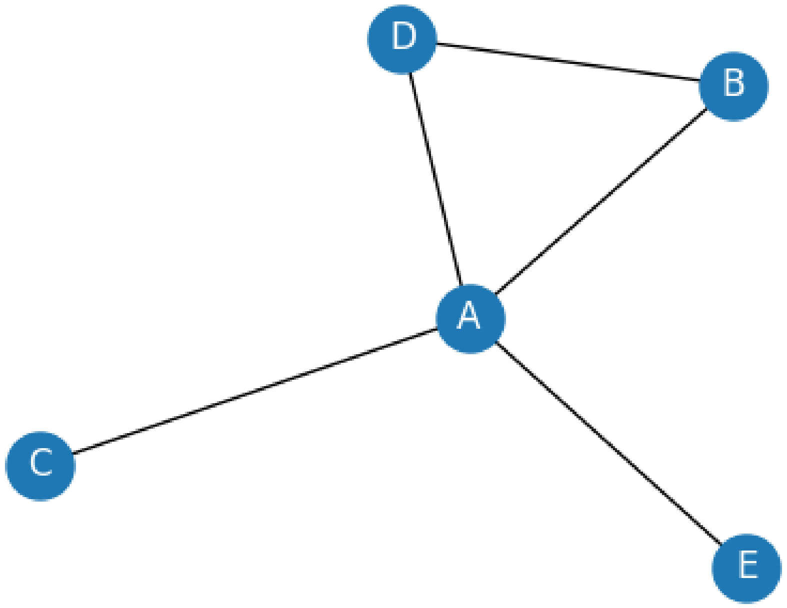



- Degree: number of nearest neighbors / (n-1)

n is maximal possible degree
{'A': 1.0, 'B': 1.0, 'C': 0.66, 'D': 0.66}

- Closeness: centrality of a node u is the reciprocal of the average shortest path distance to u over all (n-1) reachable nodes.

$$C(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)}$$

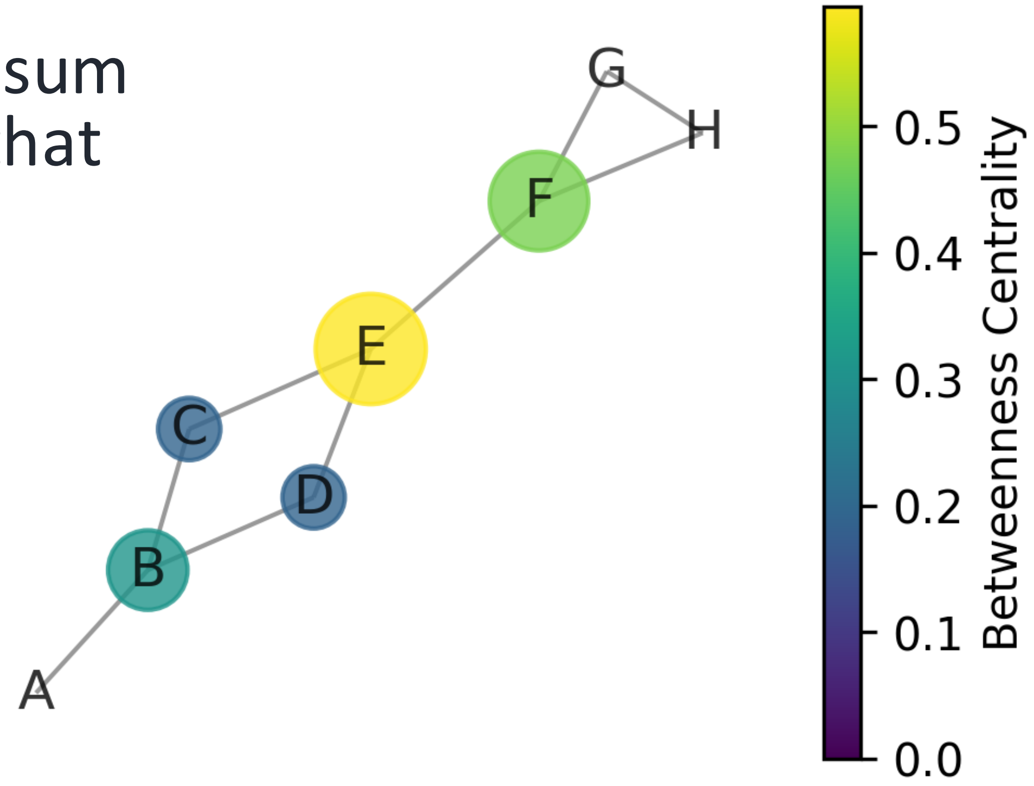



A = 1.0 = 4 / (1+1+1+1) = 4/4
B = 0.66 = 4 / (1+1+2+2) = 4/6
C = 0.57 = 4 / (1+2+2+2) = 4/7
D = 0.66 same as B
E = 0.57 same as C

# Betweeness Centrality

- Betweenness centrality of a node v is the sum of the fraction of all-pairs shortest paths that passes through v.

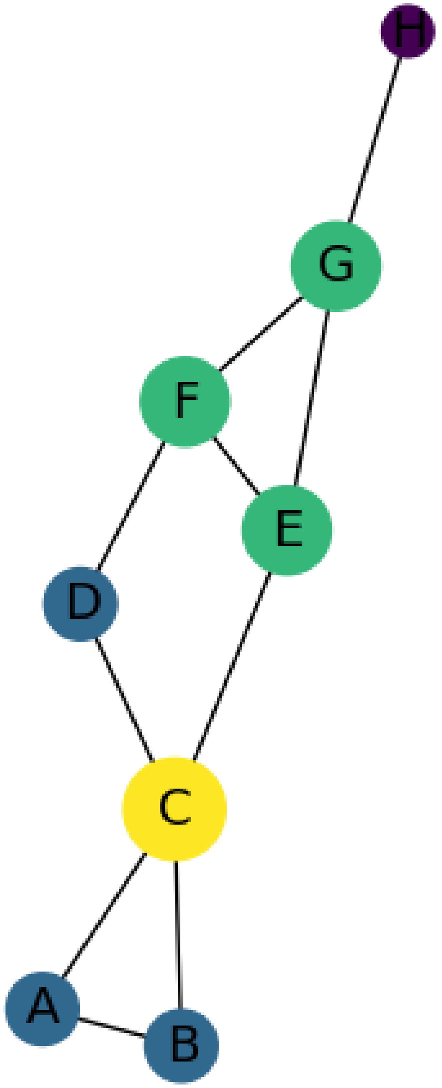
$$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$




A = 0.0
B = 6.5
C = 4
D = 4
E = 12.5
F = 10
G = 0.0
H = 0.0

Node E = (AF+AG+AH+BF+BG+BH+
CF+CG+CH+DF+DG+DH+CD)
=(2/2+ 2/2+ 2/2+ 2/2+ 2/2+ 2/2+
1/1+ 1/1 + 1/1+ 1/1 +1/1+1/1+1/2)

# Clustering coefficient

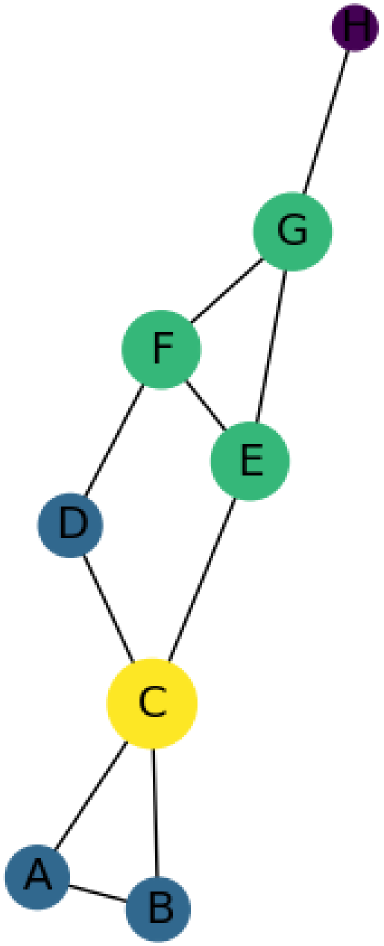



Clustering coefficient is a measure of the degree to which nodes in a graph tend to cluster together.

Neighbors of a given node are connected to each other.

| Node | # of neigh | # possible conne | clustering coef |
|---|---|---|---|
| A | 2 | 1 | (1/1)=1 |
| B | 2 | 1 | (1/1)=1 |
| C | 4 | 6 | (1/6)=0.167 |
| D | 2 | 1 | (0/1)=0 |
| E | 3 | 3 | (1/3)=0.33 |
| F | 3 | 3 | (1/3)=0.33 |
| G | 3 | 3 | (1/3)=0.33 |
| | | Avg coef. | = 0.40 |

# Adjacency matrix




| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| A | | 1 | 1 | | | | | |
| B | 1 | | 1 | | | | | |
| C | 1 | 1 | | 1 | 1 | | | |
| D | | | 1 | | | 1 | | |
| E | | | 1 | | | 1 | 1 | |
| F | | | | 1 | 1 | | 1 | |
| G | | | | | 1 | 1 | | 1 |
| H | | | | | | | 1 | |

# Network modularity

- Modularity measures the strength of division of a network into modules (also called clusters or communities).
- Values
  - closer to 1 indicate strong community structures.
  - near 0 indicate no significant community structure.
  - Negative values means worse-than-random divsions.

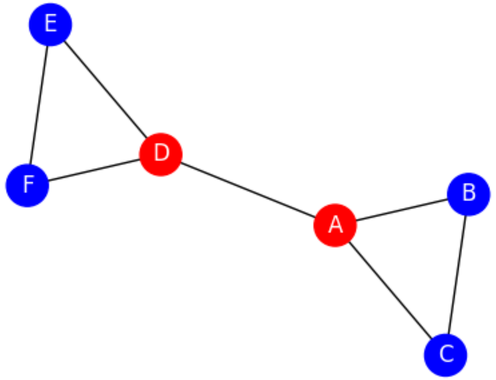
$$Q = \frac{1}{2m} \sum_{ij} \left( A_{ij} - \gamma \frac{k_i k_j}{2m} \right) \delta(c_i, c_j)$$

- m is the number of edges (or sum of all edge weights)
- A is adjacency matrix of graph G
- $k_i$ is the (weighted) degree of noid i
- γ is the resolution parameter
- δ ($c_i$, $c_j$) is 1 if i and j are in same cluster else 0.

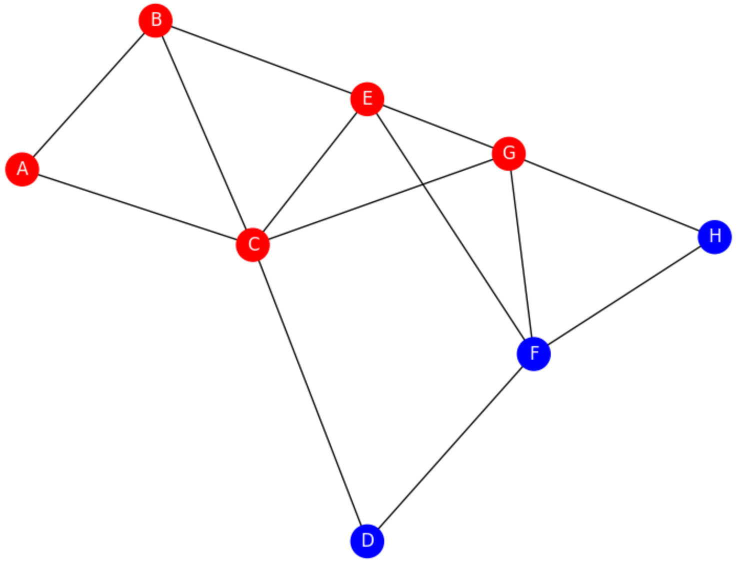
https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.quality.modularity.html

# Examples

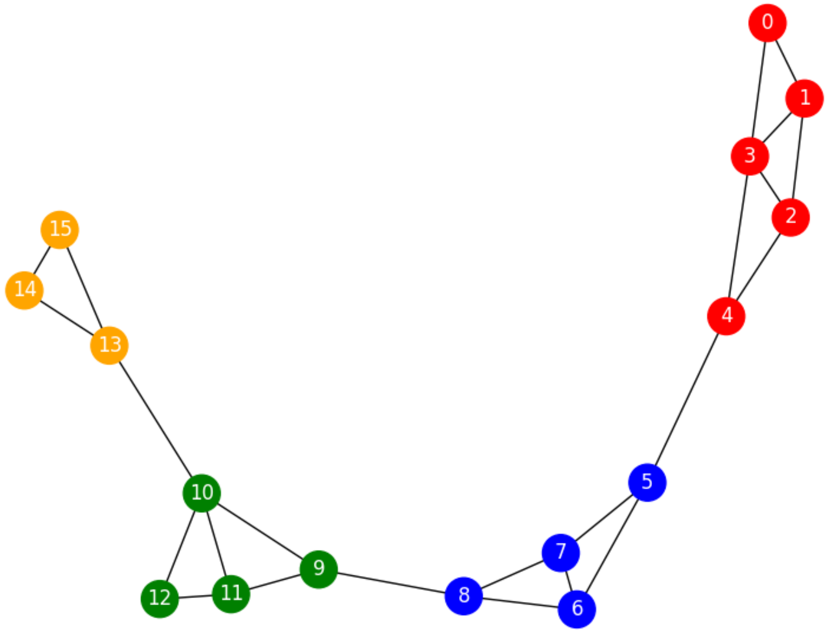
```
mod = nx.community.modularity(G,
nx.community.label_propagation_communities(G))
```




Mod=-0.08



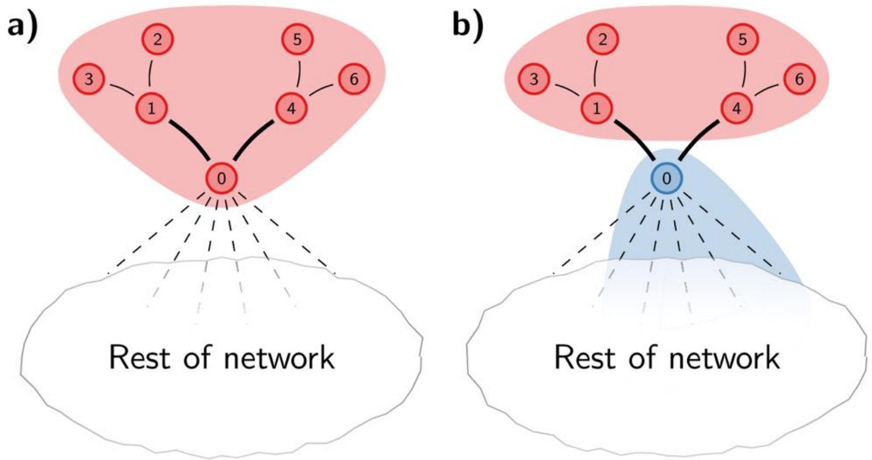

Mod=0.12




Mod=0.60

# Application of modularity

- Louvain clustering 2008
  - Start with a single partition where all nodes are in their own community
  - Iterat repeatedly
    - Local node movement: Move nodes to communities to maximize modularity
    - Network aggregation: Aggregate communities into new nodes within network
- Leiden clustering 2019
  - Faster and address an issue where communities may become internally disconnected but remain as one community

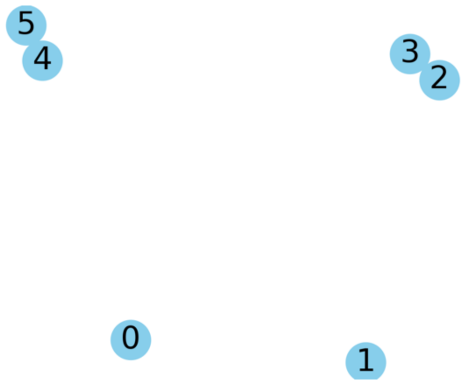



V. A. traag et al Scientific Reports 2019

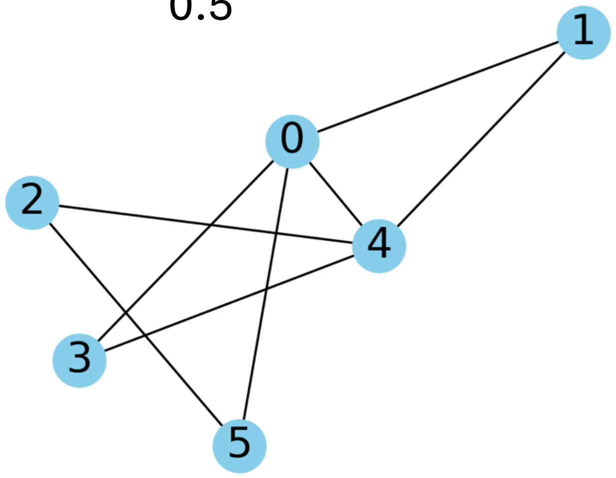
# Random graph

- A graph is constructed by connected nodes randomly.
- Each edge is included in the graph with probability p, independently from every other edge.
- Equivalently, the probability for generating each graph that n nodes and M edges is

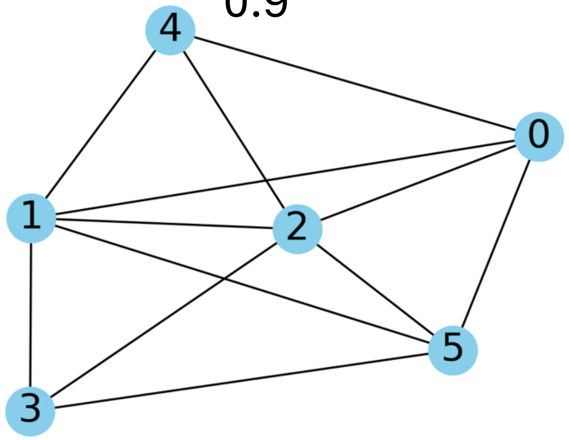
Prob=0.1




0.5




0.9




$$\text{Expected edges} = p \cdot \binom{n}{2}$$

(6 Choose 2)

\# of edges 1.5

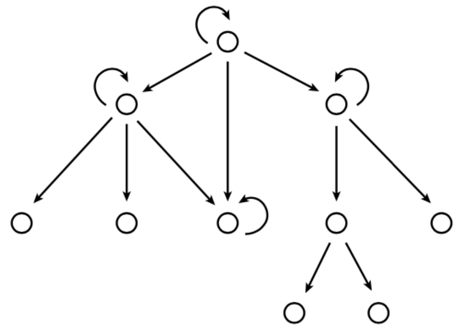
7.5

13.5

# Network motif

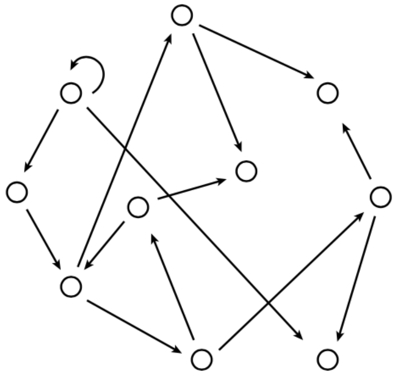
- Patterns that occur in the real network significantly more often than in randomized networks with the same characteristics (number of nodes, number of edges) are called network motifs.
- Edges in the network motifs must be constantly selected in order to survive randomization forces in unexpected high amounts.

‘Real’ network



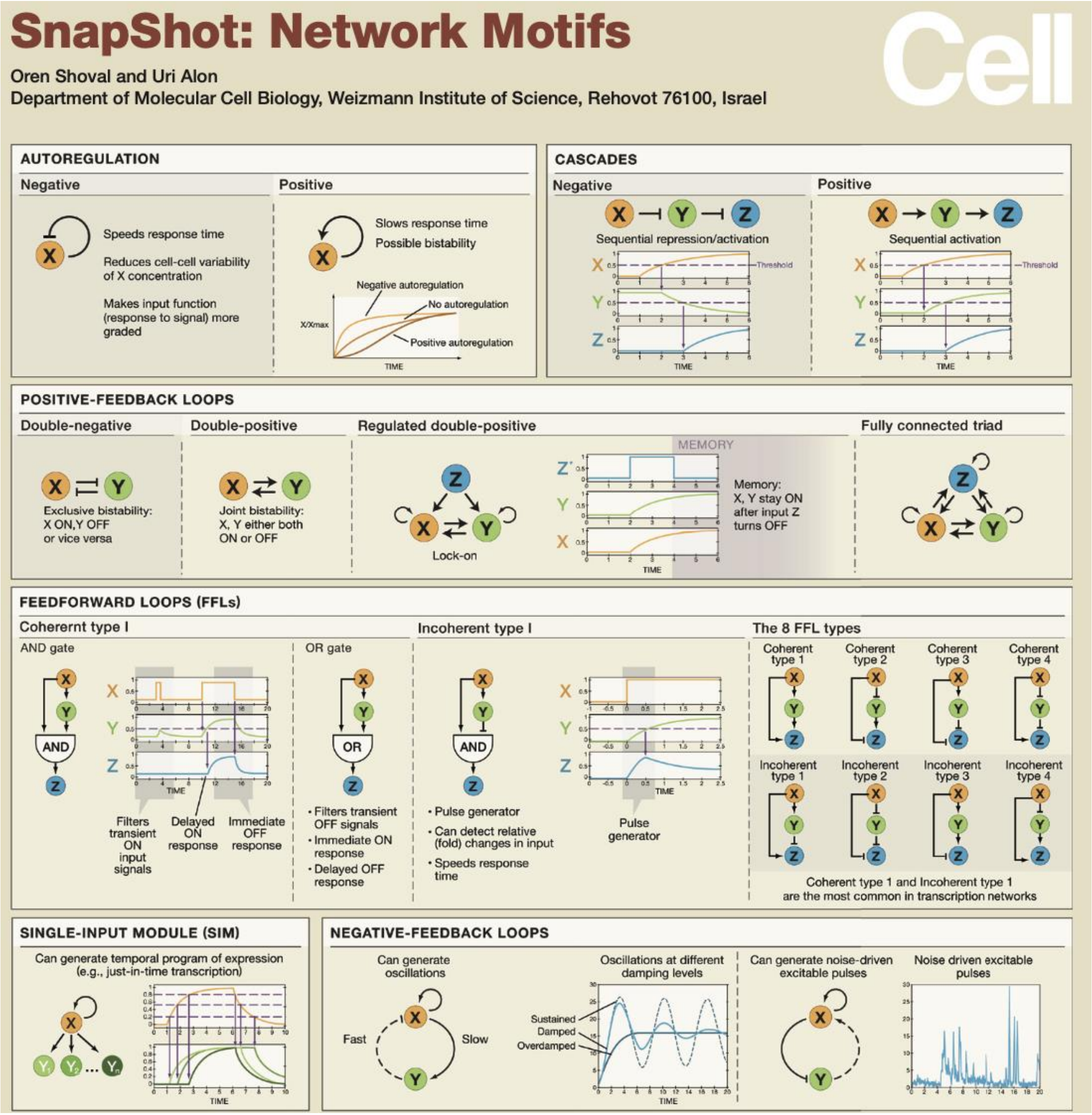
N=10 nodes
A=14 arrows
$N_{self}$=4 self-arrows

Randomized network
(Erdos - Renyi)



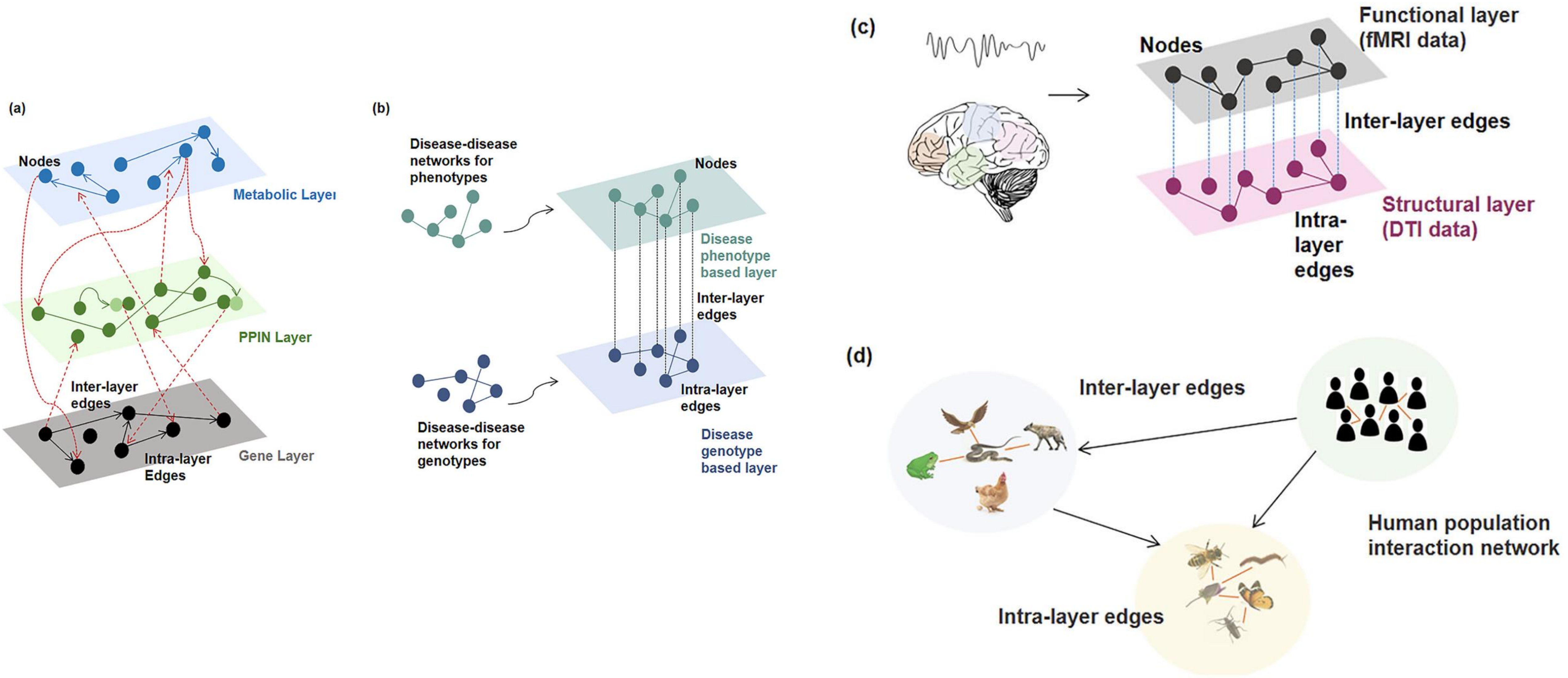
N=10 nodes
A=14 arrows
$N_{self}$=1 self-arrow

# Graph Motif Connections to Systems Biology



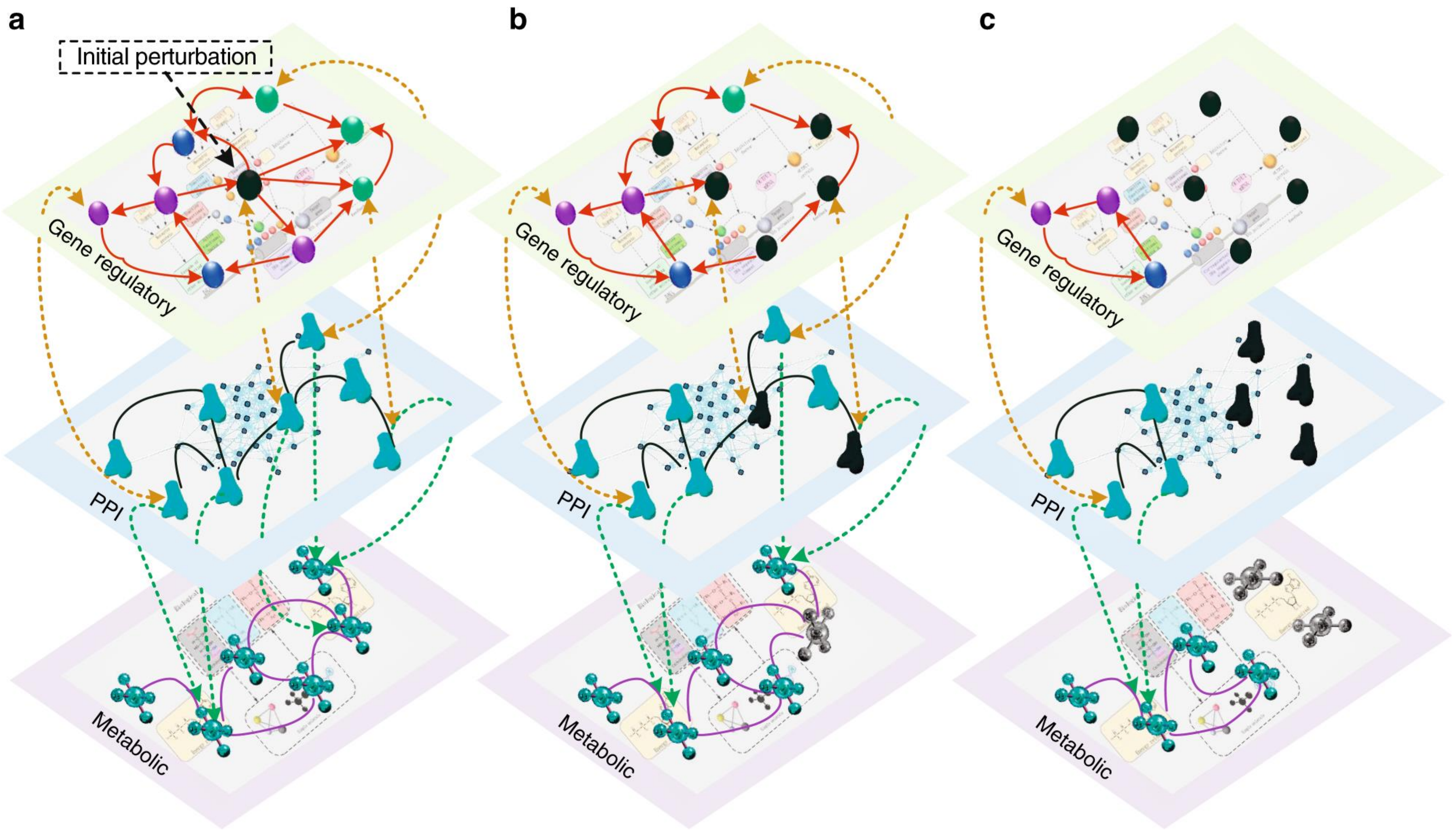

Oren Shoval and Uri Alon, Cell 2010

# Multilayer networks




S Chaudhuri and A Srivastava J Biosci 2022

# Cascading failure process in multilayer biological molecular networks




Xueming Liu et. al. Nat. Comm. 2020

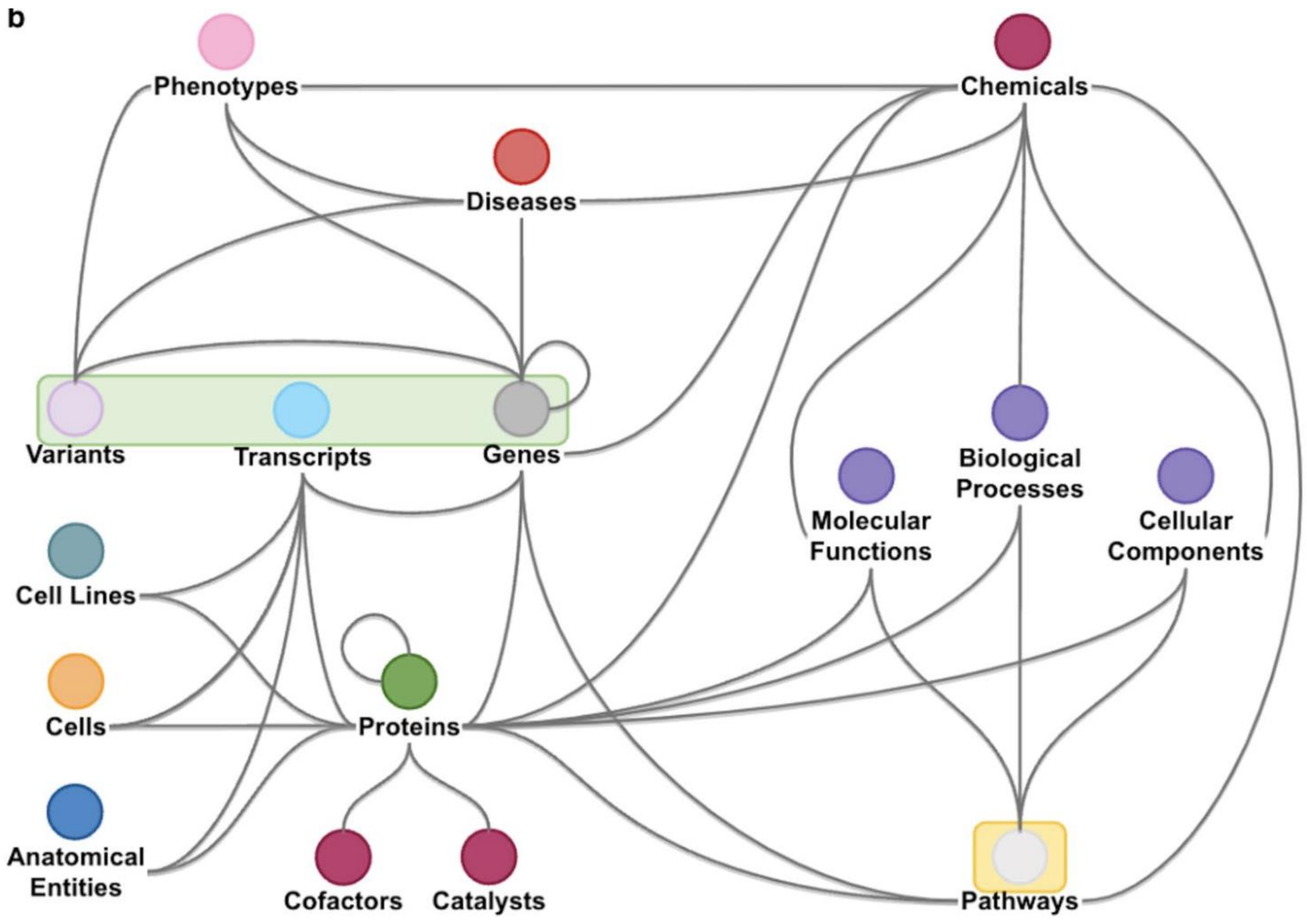
# Neural network

# Robustness and resilience of networks

# Famous algorithms related to graphs

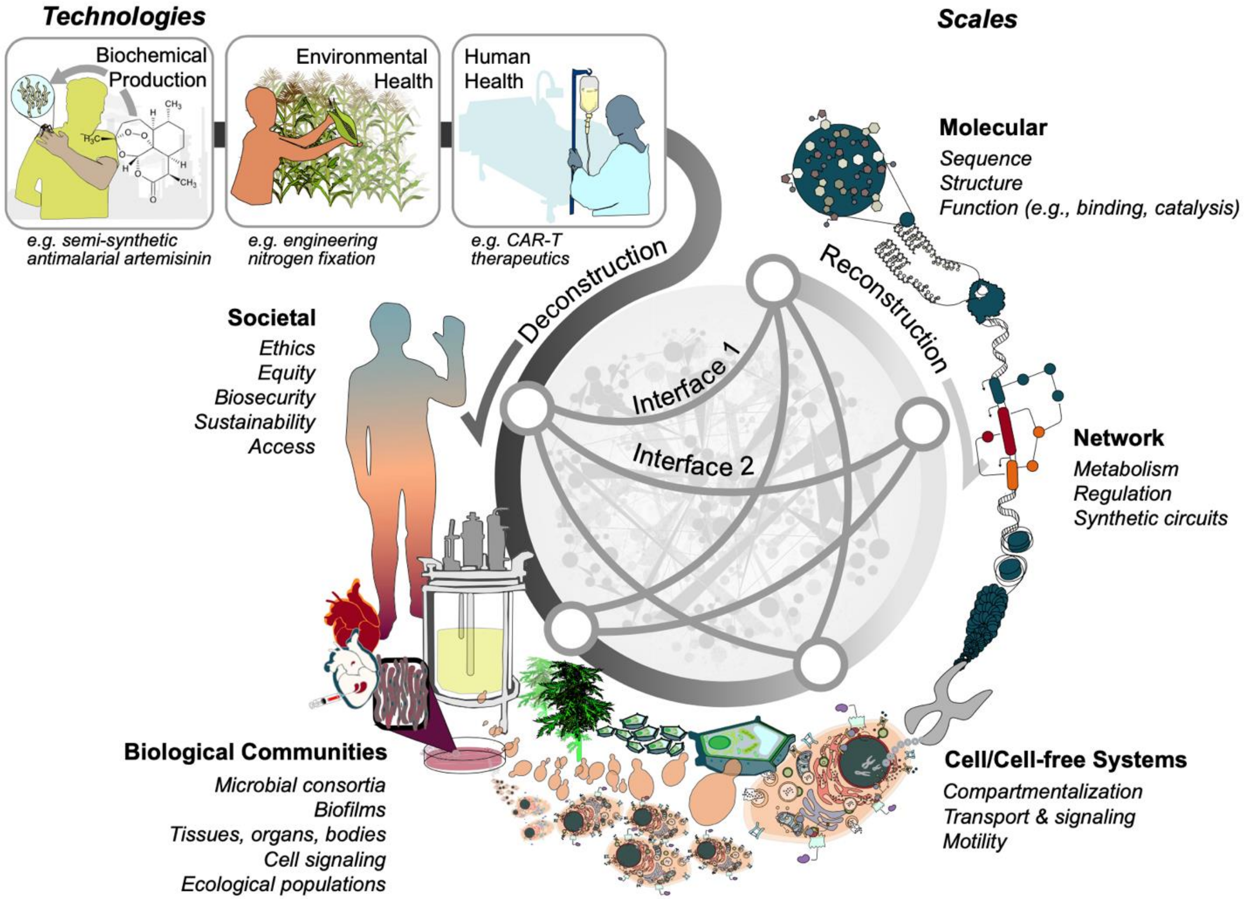
- Dijkstra’s (shortest path problem)
- PageRank
- Leiden community detection
- Kd-tree
- Random Forest
- Causal
- Markov clustering
- Search

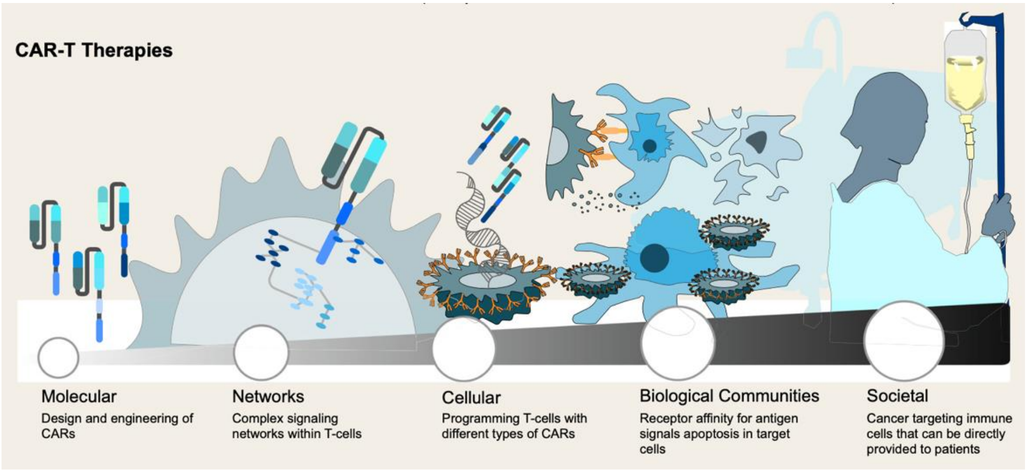
# Knowledge graph of a disease




Tiffany J. Callahan et al Scientific Data 2024

# Deconstructing synthetic technologies across scales




Ashty S. Karim et al Nat Comm 2024