NUTRIOME

Functional analysis

Part 2: Network analysis

Martina Summer-Kutmon

martina.kutmon@maastrichtuniversity.nl

NUTRIOME Workshop 1

Maastricht Centre for Systems Biology (MaCSBio)

30 May 2024

ORCID: 0000-0002-7699-8191

Maastricht University

Funded by the European Union

Marie Skłodowska-Curie Actions

macsbio
MAASTRICHT CENTRE FOR SYSTEMS BIOLOGY

Current knowledge level

NUTRIOME

Note

- This is a short introduction into the basics of network biology.

- We focus on how to explore the network and about different network resources for protein-protein interactions.

NUTRIOME

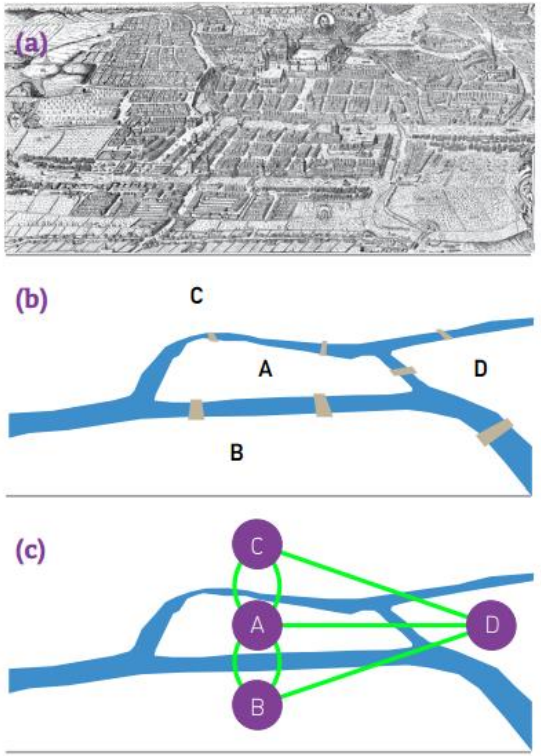
Introduction

Funded by
the European Union

Marie
Skłodowska-
Curie Actions

Network science

- Building on the field of graph theory
  - 1735 Koenigsberg (now Kaliningrad, Russia)
  - Leonard Euler (Swiss mathematician)
  - Walk across all seven bridges and never cross the same twice
  - Euler offered mathematical proof that such a path does not exist - using a graph representation

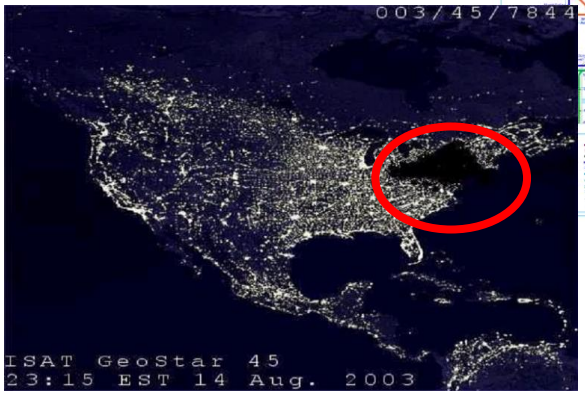


Networks are everywhere

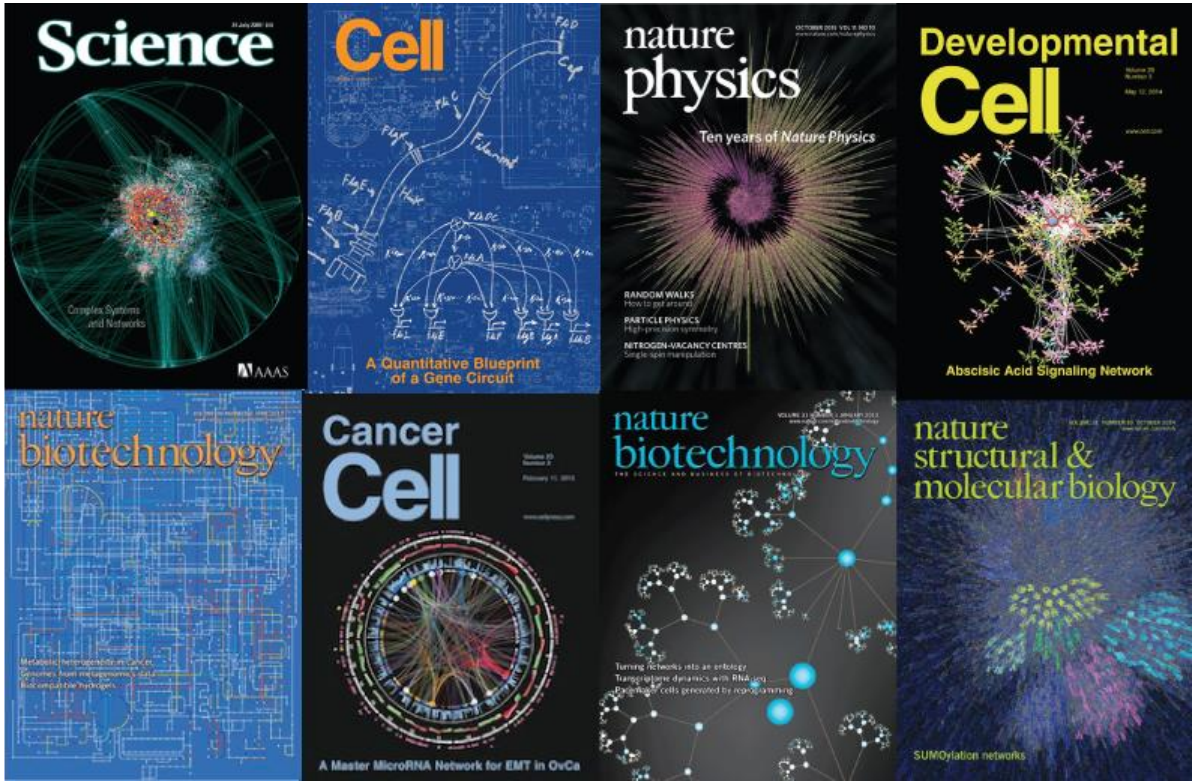
What networks do you use / know from your everyday life?

Network science

- Building on the field of graph theory

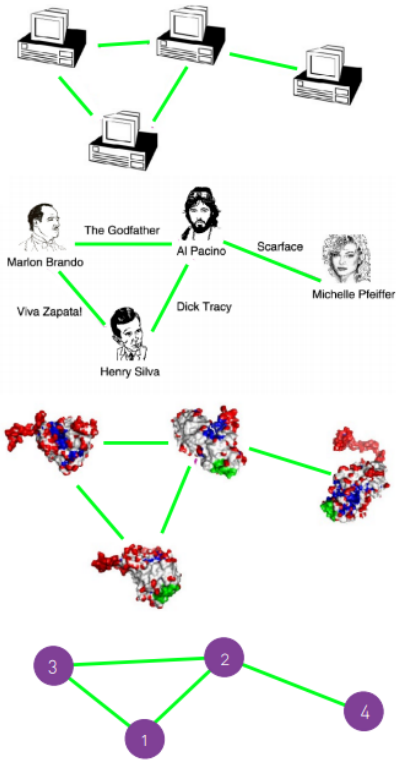


Networks in science



Network science

- Building on the field of graph theory



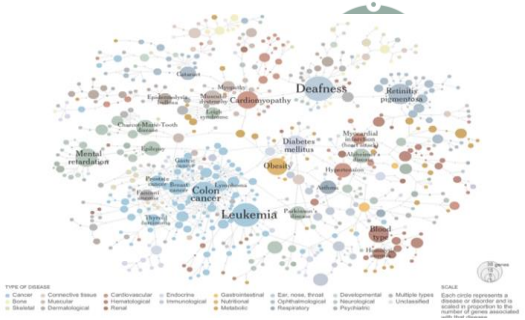
Network as universal concept

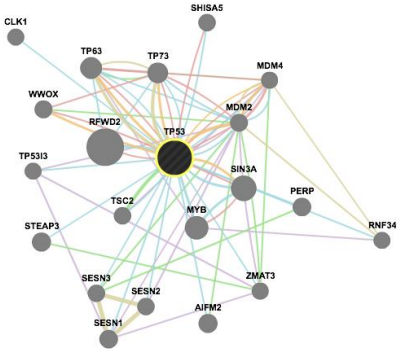
Why networks in biology?

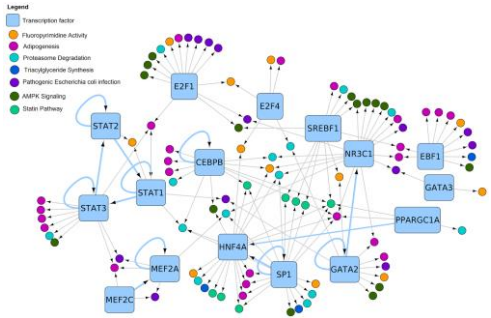
- Study biological complexity
- More efficient than tables
- Great for data integration
- Intuitive visualization

Why networks in biology?

- **Types of networks**
  - Molecular networks
    - Protein-protein interaction networks
    - Metabolic networks
    - Regulatory networks
  - Cell-cell communication
  - Nervous systems
  - Human disease network
  - Social networks

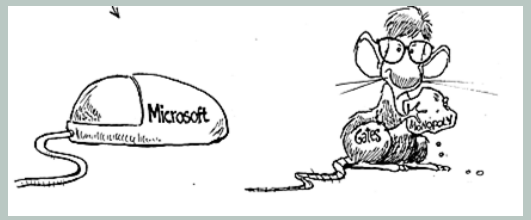






NUTRIOME

Terminology



Funded by
the European Union

Marie
Skłodowska-
Curie Actions

Network

- A network is a graphical representation of a set of **objects** where some pairs of objects are connected by **links**.

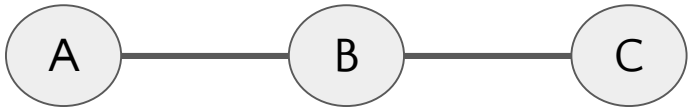
Network

- A network is a graphical representation of a set of **objects** where some pairs of objects are connected by **links**.

Objects in the network are called **nodes** (A and B).
Links in the network are called **edges** or interactions.

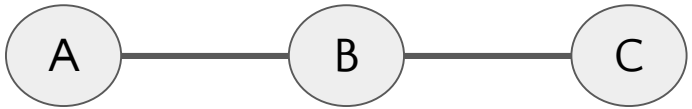
Neighbour

A neighbour is a node that is **linked** with another **node** through a **direct edge**.



Neighbour

A neighbour is a node that is **linked** with another **node** through a **direct edge**.



A is a neighbour of B but not of C.
B has two neighbours - A and C.
C is a neighbour of B but not A.

Path

- A path is a **sequence of edges** which connect a sequence of nodes.
- A path can intersect itself and pass through the same node/edge repeatedly.

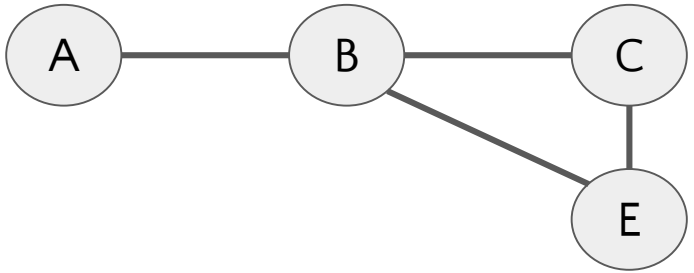
Path

- A path is a **sequence of edges** which connect a sequence of nodes.
- A path can intersect itself and pass through the same node/edge repeatedly.

Path(s) from A to E?

A - B - C - E

A - B - E



Distance

- Distance between two nodes is the number of edges along the path connecting them.
- If two nodes are disconnected, the distance is infinity.
- The **shortest path** is the path with the minimal number of edges necessary to get from one node to another.

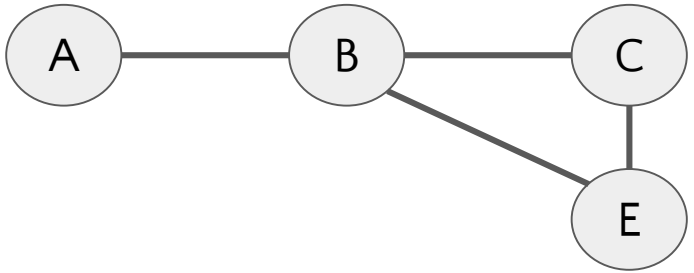
Distance

- Distance between two nodes is the number of edges along the path connecting them.
- If two nodes are disconnected, the distance is infinity.
- The **shortest path** is the path with the minimal number of edges necessary to get from one node to another.

Distance from A to E?

A - B - C - E 3

A - B - E 2



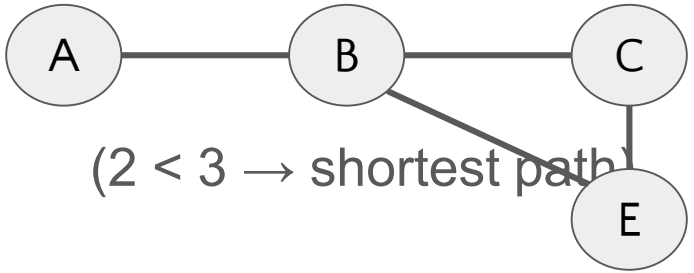
Distance

- Distance between two nodes is the number of edges along the path connecting them.
- If two nodes are disconnected, the distance is infinity.
- The **shortest path** is the path with the minimal number of edges necessary to get from one node to another.

Distance from A to E?

A - B - C - E 3

A - B - E **2**



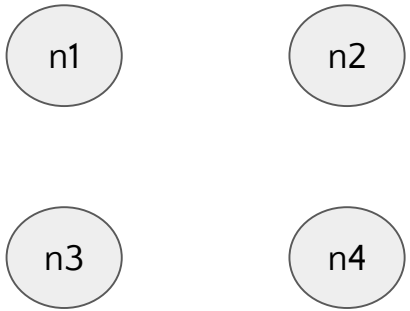
Adjacency matrix

Mathematical representation

A_{ij} = 1 there is an edge between node i and j

A_{ij} = 0 there is no edge between node i and j

$$A_{ij} = \begin{array}{c} \\ n1 \\ n2 \\ n3 \\ n4 \end{array} \begin{array}{c} \begin{array}{cccc} n1 & n2 & n3 & n4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$



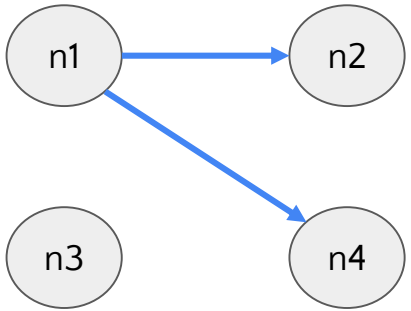
Adjacency matrix

Mathematical representation

A_{ij} = 1 there is an edge between node i and j

A_{ij} = 0 there is no edge between node i and j

$$A_{ij} = \begin{array}{c} \\ n1 \\ n2 \\ n3 \\ n4 \end{array} \begin{array}{c} \begin{array}{cccc} n1 & n2 & n3 & n4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$



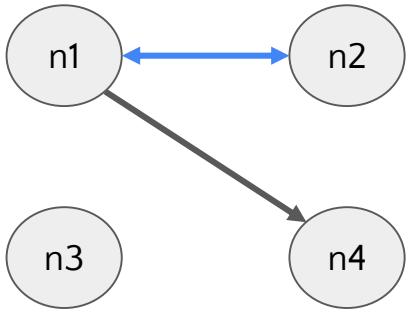
Adjacency matrix

Mathematical representation

A_{ij} = 1 there is an edge between node i and j

A_{ij} = 0 there is no edge between node i and j

$$A_{ij} = \begin{array}{c} \\ n1 \\ n2 \\ n3 \\ n4 \end{array} \begin{array}{c} \begin{array}{cccc} n1 & n2 & n3 & n4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$



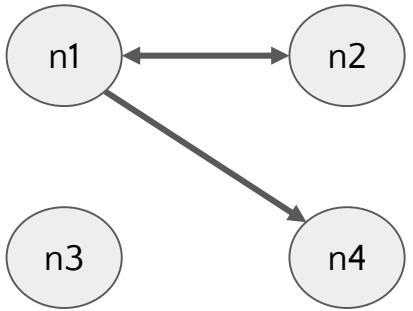
Adjacency matrix

Mathematical representation

A_{ij} = 1 there is an edge between node i and j

A_{ij} = 0 there is no edge between node i and j

$$A_{ij} = \begin{array}{c} \\ n1 \\ n2 \\ n3 \\ n4 \end{array} \begin{array}{c} \begin{array}{cccc} n1 & n2 & n3 & n4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$



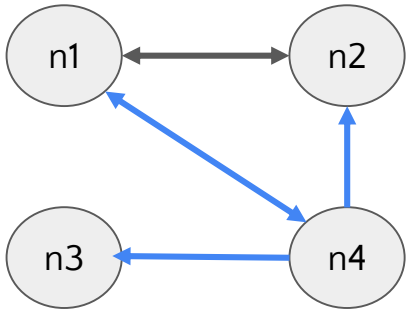
Adjacency matrix

Mathematical representation

A_{ij} = 1 there is an edge between node i and j

A_{ij} = 0 there is no edge between node i and j

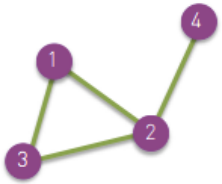
$$A_{ij} = \begin{array}{c} \\ n1 \\ n2 \\ n3 \\ n4 \end{array} \begin{array}{c} \begin{array}{cccc} n1 & n2 & n3 & n4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$



Undirected vs. directed networks

Undirected

Links: undirected (symmetrical)



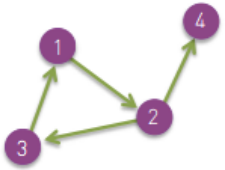
$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Examples:

Coauthorship, actor network, protein interactions

Directed

Links: directed (arcs)



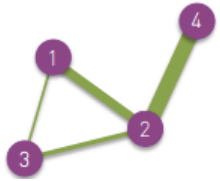
$$A_{ij} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Examples:

URLs (internet), phone calls, metabolic reactions

Weighted networks

Edges have a defined weight, strength or flow parameter



$$A_{ij} = \begin{pmatrix} 0 & 2 & 0.5 & 0 \\ 2 & 0 & 1 & 4 \\ 0.5 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \end{pmatrix}$$

Examples:

Correlation networks, route planning, mobile phone calls

Centrality measures

NUTRIOME

- Indicators which identify the **most important nodes** and/or **edges** in the network
  - Degree centrality
  - Betweenness centrality
  - Clustering coefficient
  - ...

Help to answer the following questions:

How influential is a person?
How important is a room in a building?
How much influence has a mutation in a protein?

Degree centrality

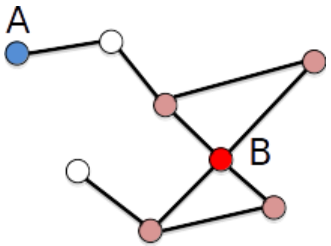
- Undirected:
  - node degree = number of edges connected to the node

- Directed:
  - in-degree = number of edges pointing towards a node (regulators)
  - out-degree = number of edges going out of a node (targets)

Degree centrality

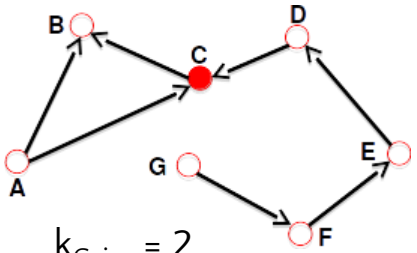
- Undirected:
  - node degree = number of edges connected to the node

- Directed:
  - in-degree = number of edges pointing towards a node (regulators)
  - out-degree = number of edges going out of a node (targets)



$k_A = 1$

$k_B = 4$



$k_{C\_in} = 2$

$k_{C\_out} = 1$

Degree centrality

- Biological interpretation
  - Nodes with a high degree tend to be essential
  - Nodes with a high degree are also called **hub nodes**

Betweenness centrality

- Betweenness = number of shortest paths going through a node

$$C_b(n) = \sum \frac{\delta_{st}(n)}{\delta_{st}}$$

δ_{st} = number of shortest path from s to t

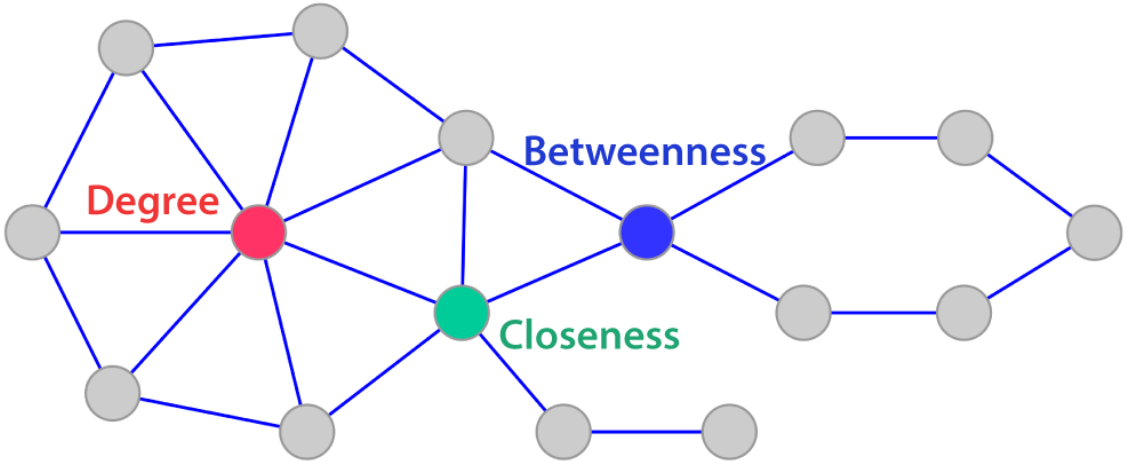
$\delta_{st}(n)$ = number of shortest path from s to t that go through n

- Betweenness = 0 no shortest paths go through this node
- Betweenness = 1 all shortest paths go through this node

Betweenness centrality

- Biological interpretation
  - Information load on a node
  - Control of the node over the connectivity of the network
  - Connection of two subnetworks
  - Weak links
  - Can be calculated for edges too

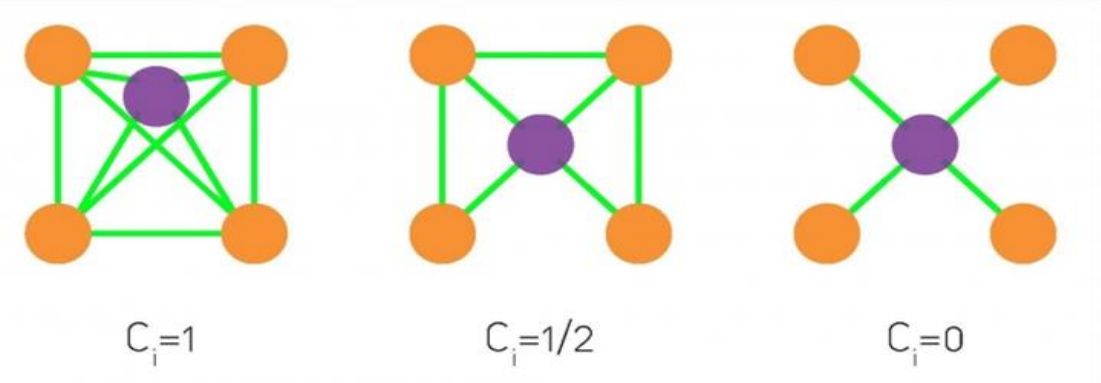
Betweenness centrality



Degree	ClosenessCentrality	BetweennessCentrality
7	0.45454545	0.29047619
5	0.51724138	0.42380952
4	0.48387097	0.4952381

Clustering coefficient

- Connectivity of the neighborhood - measure for the network's local edge density
  - How many of a nodes neighbors are connected to each other?

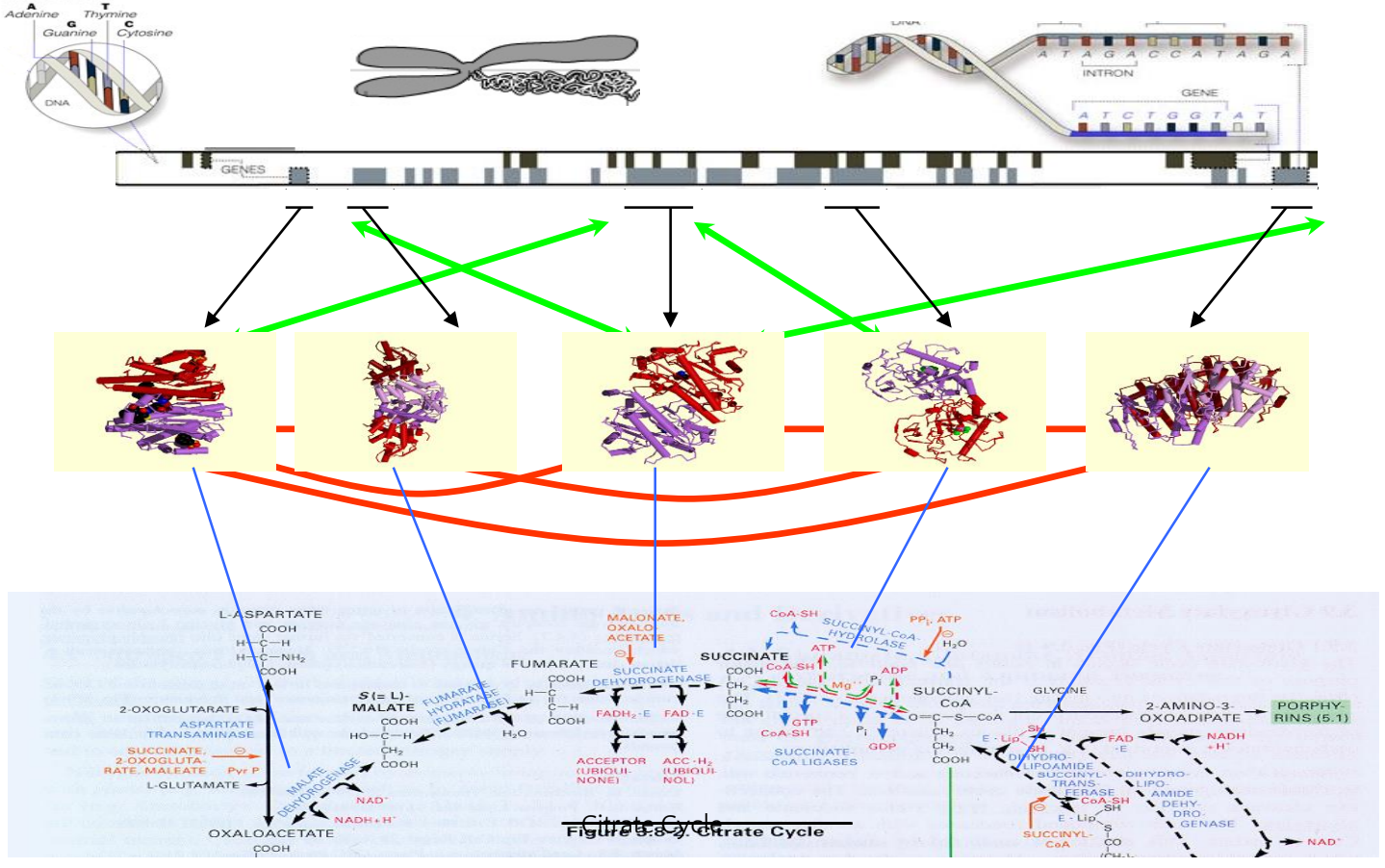


Where do I find *the* network?

- There is no such thing!
- >700 different interaction databases

www.pathguide.org

Molecular networks



GENOME

protein-gene interactions

PROTEOME

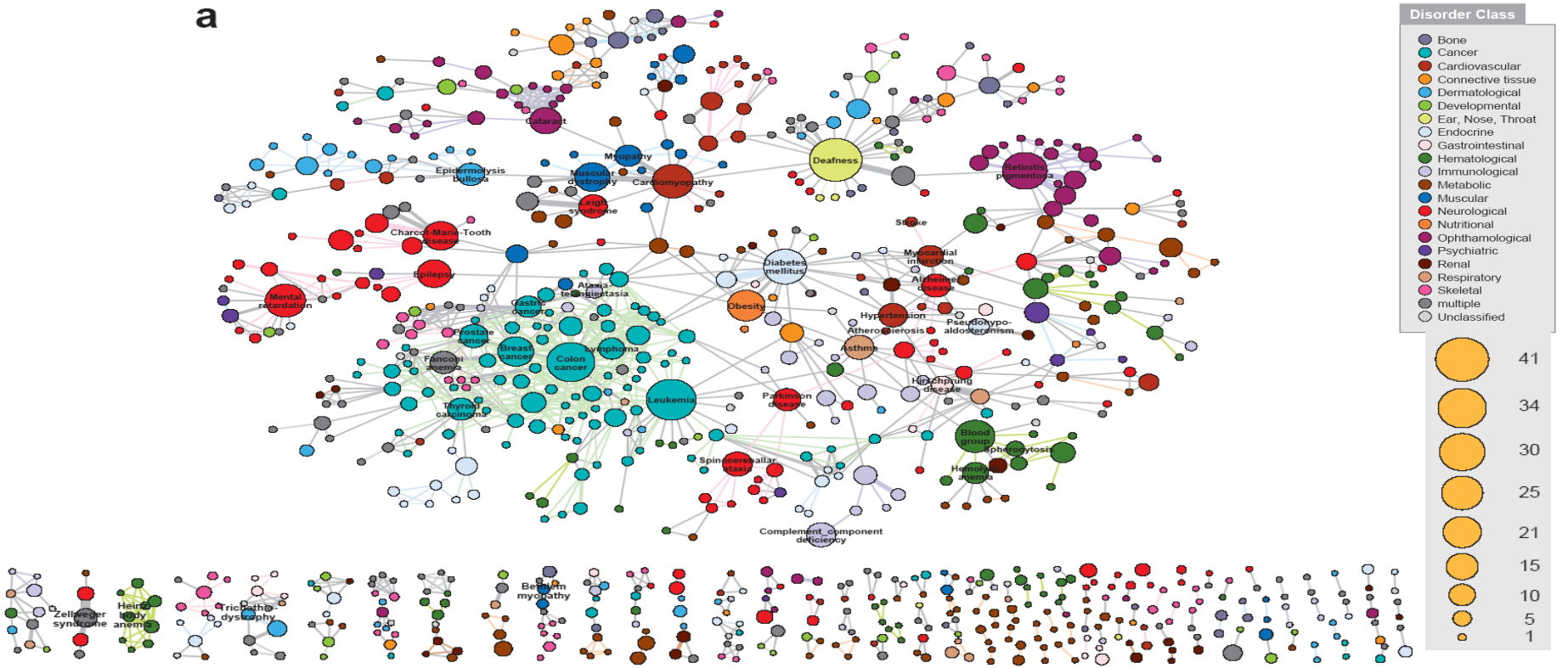
protein-protein interactions

METABOLISM

bio-chemical reactions

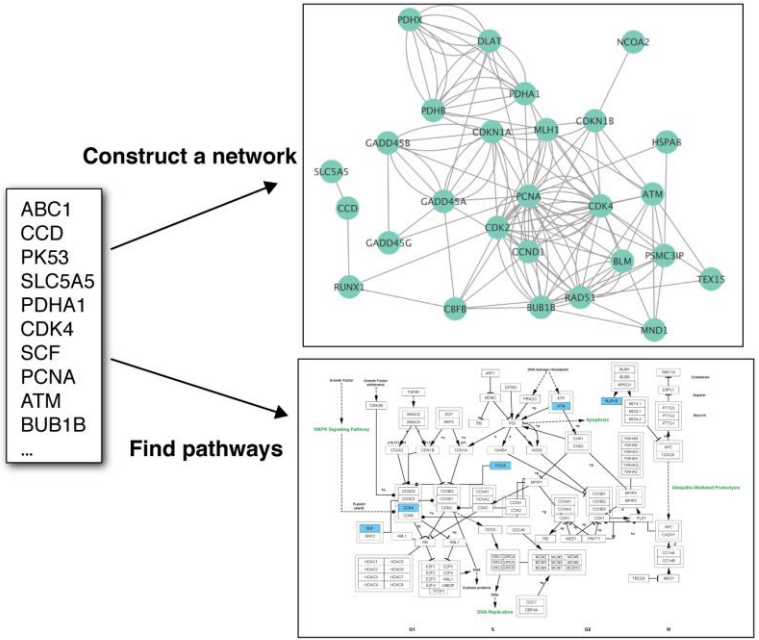
Human disease network

NUTRIOME



Finding network sources

- Depends on biological question and analysis plan
- Typical start: **gene list**



Finding network sources

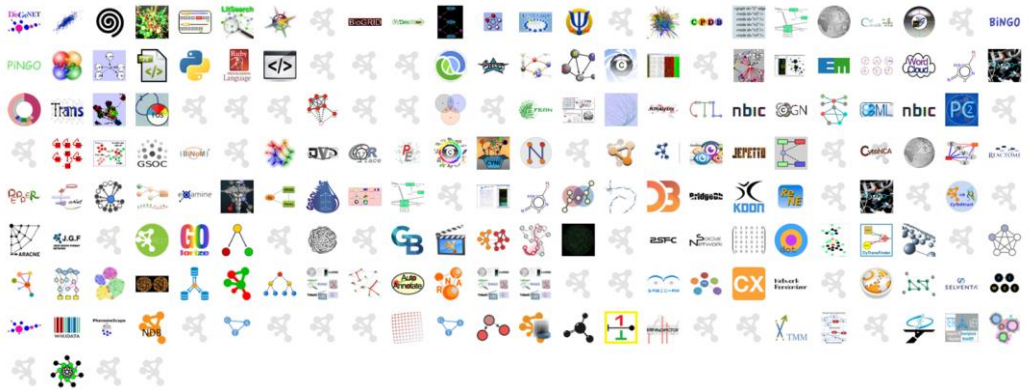
- **Networks**
  - Broad coverage / low resolution
  - **Databases:**
    - PSICQUIC, STRING, IntAct, GeneMANIA, NDEx, etc
  - **Interaction types**
    - Protein-protein interactions
    - Gene-regulatory interactions
    - Genetic interactions
    - Protein-compound interactions

Finding network sources

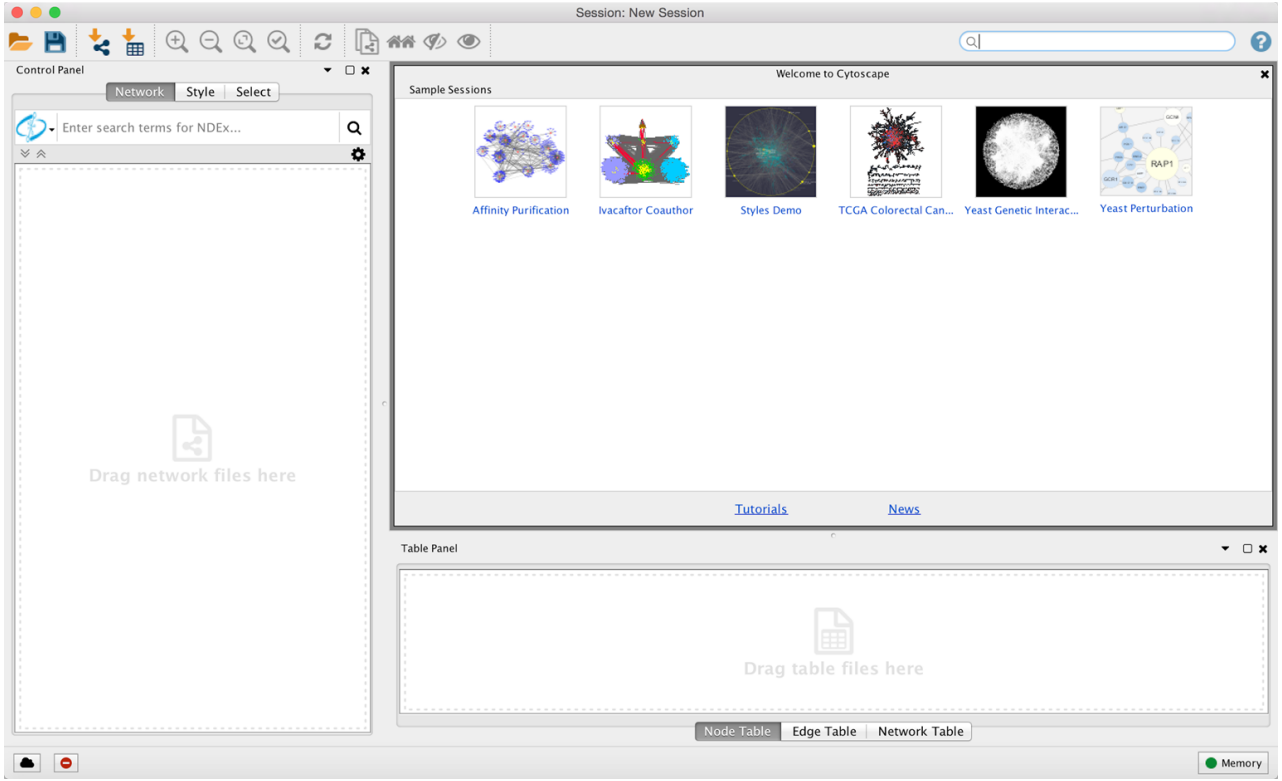
- **Pathways**
  - High resolution / limited coverage (~60% of genes)
  - **Databases:**
    - WikiPathways, Reactome, Pathway Commons, KEGG, etc
  - **Interaction types**
    - Signaling pathways
    - Metabolic pathways
    - Gene regulation pathways

Cytoscape

- Cytoscape (www.cytoscape.org) is a widely adopted network analysis and visualization toolbox
- Extendable with apps
  - 373 apps available (apps.cytoscape.org)

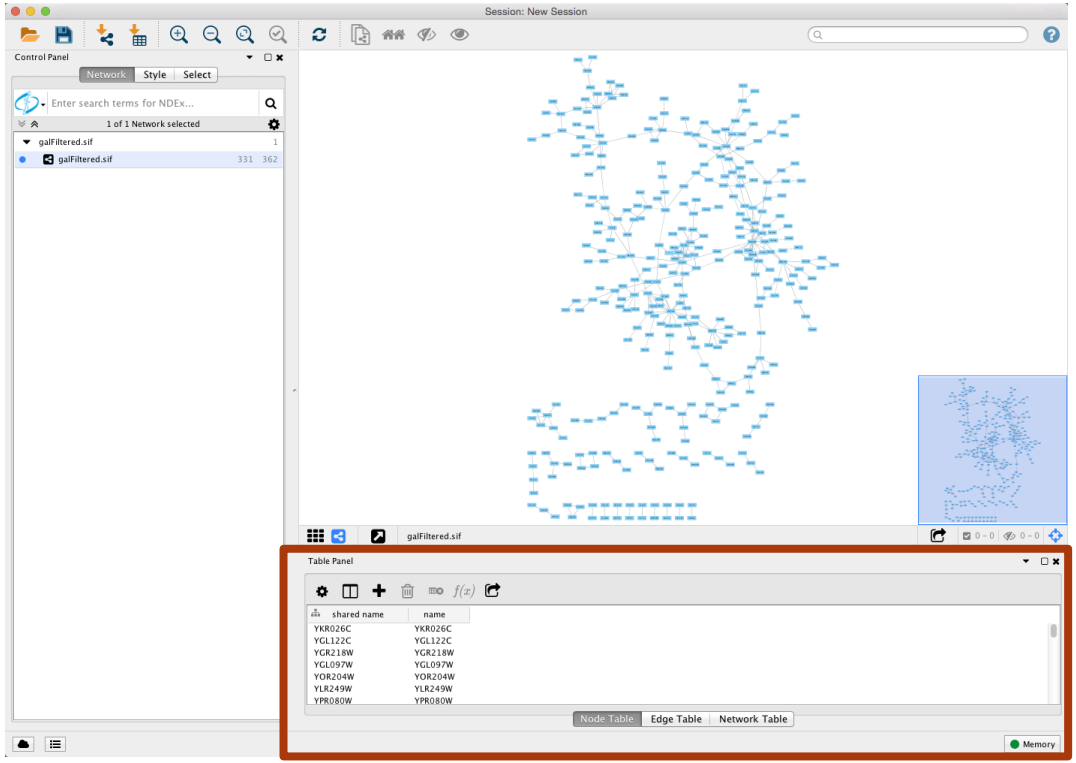


Cytoscape


Session: New Session
Control Panel
Network
Style
Select
Enter search terms for NDEx...
Drag network files here
Welcome to Cytoscape
Sample Sessions
Affinity Purification
Ivacaftor Coauthor
Styles Demo
TCGA Colorectal Can...
Yeast Genetic Interac...
Yeast Perturbation
RAP1
Tutorials
News
Table Panel
Drag table files here
Node Table
Edge Table
Network Table
Memory

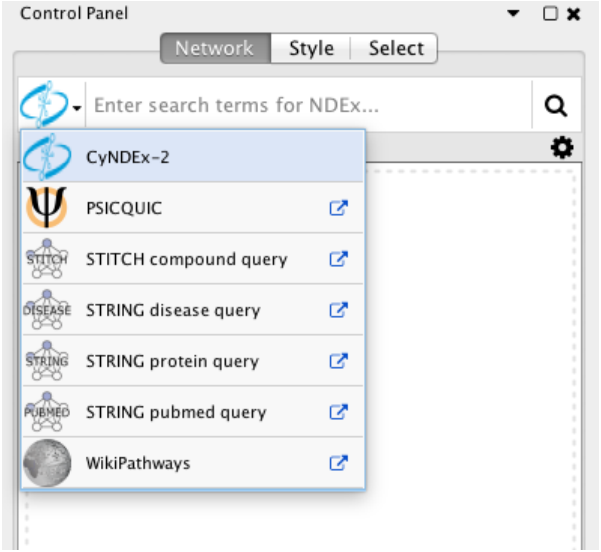
Cytoscape

NUTRIOME



Cytoscape

- Import networks from public databases



Cytoscape

- There is a lot more functionality than what we will show you in the practical:
  - http://manual.cytoscape.org/en/stable/
  - Detailed documentation and examples

NUTRIOME

Questions?

Martina Summer-Kutmon
martina.kutmon@maastrichtuniversity.nl
Maastricht Centre for Systems Biology (MaCSBio)

Maastricht University

Funded by the European Union

Marie Skłodowska-Curie Actions

macsbio
MAASTRICHT CENTRE FOR SYSTEMS BIOLOGY