Bioinformática y Tratamiento de Datos (BIF) 2023

Tema 5

Redes Biológicas y Teoría de Grafos

Florencio Pazos (CNB-CSIC)
Biological Networks and Graph Theory

- Networks
- Biological and molecular networks
- Short intro to Graph Theory
- Characteristics of the main molecular networks
- Network applications to human pathologies

- Cytoscape practical

http://csbg.cnb.csic.es/pazos/cursos/UMA_BIF/
Network representation of a phenomenon

Some phenomena/data can be represented as entities ("nodes"/"vertices") linked by relationships ("edges")

**Node** – Generic entity, physical or not. (gene, protein, metabolite, cellular state, disease, …) It can have associated features (quantitative or qualitative)

**Edge** – Generic relationship, in the broadest sense. (interaction (physical or functional), transcriptional control, chemical transformation (reaction), …) It can have associated features (quantitative or qualitative): weight, direction, ..
Large networks
Network theory / Graph theory

Once a given phenomenon is modelled as a large network, it can be studied using mathematical approaches (Graph Theory) in order to extract information hidden in its structure and topological patterns.

---

Networks have been used to model phenomena in all scientific disciplines.

Connections between computers (technological network)

Sexual contacts (social network)

Predator-prey relationships (food web) (biological/ecological network)

Relationships between banks (economic network)

Co-mentioning of characters in a novel (literature network)

In Biology, networks have been used to model diverse phenomena at all biological levels.

Molecular networks

Network approaches have been applied to the study of molecular phenomena as the data required to assemble these networks became available (-omics techniques).

Figure 1. Examples of the five major biological networks. (A) A yeast transcription factor-binding network, composed of known transcription factor-binding data collected with large-scale ChIP-chip and small-scale experiments. This figure was generated with the program Paclt (de Nooy et al. 2005). (B) A yeast protein-protein interaction network, containing protein-protein interactions identified by yeast two-hybrid and protein complexes identified by affinity purification and mass spectrometry (Barabasi and Odom 2003). (Reprinted by permission from Macmillan Publishers Ltd: Nature [Jeong et al. 2001], © 2001.) Nodes are colored according to the mutant phenotype. (C) A yeast phosphorylation network comprised primarily of in vitro phosphorylation events identified using protein microarrays (Penk et al. 2005). The figure was generated with Osprey 1.2.0. (Brenkste et al. 2003). (D) A yeast metabolic network with 574 reactions and 473 metabolites colored according to their modules (Reprinted by permission from Macmillan Publishers Ltd: Nature [Guimera and Amaral 2005], © 2005). (E) A yeast gene network constructed with synthetic lethal interactions using SGA analysis on eight yeast genes (From Tong et al. 2001; reprinted with permission from AAAS). Nodes are colored according to their YPD cellular roles.
-omics and molecular networks in the context of the Central Dogma

**DNA**

- Replication

**RNA**

- Transcription
- Translation

**Proteins**

- Proteomics

**Genomic**

- Genomics

**Transcriptomic**

- Transcriptomics

**Proteomic**

- Protein nets. (interaction, fosforilation, …)

**Metabolomic**

- Metabolomics
- Metabolic nets.

**Phenomic**

- Phenomics
- Genetic nets, phenotypic nets, disease-related nets, …
Systems biology is the study of biological systems whose behaviour cannot be reduced to the linear sum of their parts’. It is a biology-based interdisciplinary field of study that focuses on complex interactions within biological systems, using a holistic approach (holism instead of the more traditional reductionism) to biological research.

(nature.com, Wikipedia)
Is the reductionist approach of molecular biology enough?

Biological systems are a prototype of complex systems: systems characterized by a large number of components interacting non-linearly. They can not be (fully) modelled by first principles: a minimum difference in details leads to a totally different outcome.

“The whole is more than the sum of the parts”

The main idea behind complex systems is that the ensemble behaves in ways not predicted by its components. The interactions matter more than the nature of the units. Studying individual ants will almost never give us a clear indication of how the ant colony operates. For that, one needs to understand an ant colony as an ant colony, no less, no more, not a collection of ants. This is called an “emergent” property of the whole, by which parts and whole differ because what matters are the interactions between such parts. And interactions can obey very simple rules.

N. N. Taleb.
Skin in the game - Hidden asymmetries in daily life.
Reductionist vs. systemic approaches in other scientific areas
A graph is a mathematical object used to represent entities and relationships between them (understanding entity and relationship in the broadest possible way).

Formally, graph $G$ is a pair of sets $(V,E)$

- $V = \{v_1,v_2,\ldots,v_n\}$ is the set of vertices/nodes
- $E = \{(v_i,v_j),(v_{i'},v_{j'})\ldots\}$ is a set of edges (pairs formed by the elements in $V$).

- The number of nodes is called **order** of the graph.
- The number of edges is called **size** of the graph.
Example of a graph (order: 8 , size: 11).

\[ V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8\} \]

\[ E = \{(v_1, v_2), (v_1, v_3), (v_2, v_4), (v_3, v_5), (v_4, v_6), \\
(v_5, v_7), (v_5, v_8), (v_5, v_8), (v_8, v_2), (v_8, v_7), (v_4, v_5), (v_6, v_7)\} \]
An hyperedge is an edge connecting more (or less) than two nodes. Graphs with hyperedges are called hypergraphs.

\[ V = \{ v_1, v_2, v_3, v_4, v_5, v_6, v_7 \} \]

\[ E = \{ e_1, e_2, e_3, e_4 \} = \{ \{ v_1, v_2, v_3 \}, \{ v_2, v_3 \}, \{ v_5 \}, \{ v_3, v_6, v_7 \} \} \]
A directed graph (or *digraph*) is a graph in which the edges have directionality: \((v_i, v_j) \neq (v_j, v_i)\)
Weighted graph

A **weighted graph** is a graph in which the edges have associated numerical values (to quantify some characteristic of the relationship (importance, distance, capacity …))

A **weighted directed graph**…
Path

A **path** is a alternating sequence of nodes and edges connected in the network in which all nodes (and consequently all edges) are distinct.

A **closed path** (aka **cycle**) is a closed sequence of nodes and edges in which all nodes (and consequently all edges) are distinct.

![Diagram of a network with nodes and edges]

The **shortest path** between two nodes...

**C**={v₁,v₂,v₅,v₄,v₆,v₇,v₈}  
k=6

**C**={v₁,v₂,v₄,v₆,v₈}  
k=4

**Distance** between two nodes  
= length of shortest path connecting them

Distance(v₁,v₈)= 4
Clusters/modules/communities

Densely connected subgraphs = Subgraphs with many internal connections and few connections to the rest of the graph.
Signal diffusion in networks – aka “network propagation”

Detect the network regions “affected” by a set of nodes by propagating a signal from them.

Examples:

**Random walks**
Simulates a set of \( n \) walkers falling in the node(s) of interest and moving \( m \) steps randomly following the network edges.

**Heat diffusion**
Simulates that the node(s) of interest are heat sources and the network edges wires able to transmit it.

**Graph Kernels**
Mathematical operations with matrix representations of graphs that end up in a vector/matrix with the signal associated to each node and/or measures of “network distance” between nodes.
Main graph metrics

Local (node/edge) metrics

- Degree: $k_i$
- Clustering coefficient: $C_i = 2 \cdot n / (k^2 - k)$
- Betweenness (node or edge): $B_i = \text{number of shortest paths using node/edge } "i"$
- Closeness: $C_i = 1 / \sum d_{ik}$

Global (graph) metrics

These metrics summarize the main topological features of graphs and can be used to classify them.

- Size $|E|$ and order $|V|$
- Dispersion ($|E|/|V|$)
- Degree distribution: $k$ vs. $p(k)$
- Average degree ($<k>$)
- Clustering coefficient ($C$)
- Average/characteristic path length ($L$)
  The characteristic path length is the average of the lengths of the shortest paths (distances) between all nodes $L = <d_{ij}>$
- Diameter ($D$)
  The diameter of a graph is the maximum path length. $D = \text{max}(d_{ij})$
Degree (v2) = 3

Out-degree (v2) = 1
in-degree (v2) = 2

regular graph: $\text{degree}(i) = \text{degree}(j) \ \forall \ i, j$
**Node/Edge betweenness-centrality**
And closeness-centrality

**Betweenness-centrality**: Number (or fraction) of shortest-paths passing through that node/edge

\[ B(V) = 9 \]

absolute or...

\[ = 9/N \]

relative to the number of shortest paths

N: #shortest paths

= #node pairs

= \((7^2-7)/2\)

\[ C(V) = 1/(1+2+2+1+2+2) = 1/10 \]

**Closeness-centrality**: Inverse of the sum of distances to all other nodes/edges

(graph above)

A is the most central according to the degree

B is the most central according to closeness and betweenness

Not the same as “degree”
Clustering coefficient

The clustering coefficient of a node is the number of connections between its neighbors \((n)\) over the maximum number of possible connections between them.

\[ Ci = \frac{n}{k_i(k_i-1)/2} \]  
(being \(k_i\) the degree of the node = number of neighbors)

It gives an idea how clustered or sparse is the neighborhood of a node.

\[ Ci = 3/3 = 1.0 \quad Ci = 2/3 = 0.67 \quad Ci = 1/3 = 0.33 \quad Ci = 0/3 = 0.0 \]

The clustering coefficient of a graph is the average of the clustering coefficients of all its nodes. It gives an idea of how clustered/interconnected a graph is (vs. “star-like” graphs and trees)

\[ C = \langle Ci \rangle \]
Global metrics - Degree distribution

Main graph types

**Random graphs**
Generation: assign $E$ edges randomly, from the $(N^2-N)/2$ possible. Equivalent to create each edge with probability $p = E/(N^2-N)/2$
Degree distribution: poison.

**Regular graphs**
Degree distribution: single value $k$.
Mathematical “curiosities”. There are analytic expressions for all the metrics.
For small $k$ (disperse graphs), $C \sim 0.75$ (high) and $L \sim N/k$

**Scale-free graphs**
Arise from various possible generation mechanisms:
- Preferential attachment: add nodes to an existing network connecting them PREFERENTIALLY to already highly connected nodes (“rich gets richer”, …)
- Node duplication preserving links (biological networks)

Characterized by “hubs” (highly connected nodes, in low number)
“Small world” in some cases (hubs act as shortcuts).
Degree distribution: power-law: $p(k) = C \cdot k^{-\gamma}$
$\gamma$ characterizes the connectivity pattern of the network (proportion of hubs) and in real networks it ranges $\sim 2.0 - 3.0$
Adequately represent a number of real-word networks such as many biological networks and the Internet
Graph layout

A “layout” is an arrangement of the graph nodes (and edges) in 2D or 3D which facilitates its visualization and/or makes more evident some properties of the graph.

The same graph can be visualized with many different layouts

\[ G = \{ a, b, c, d, e, f, (a,b), (a,c), (a,d), (a,e), (a,f), (b,d) \} \]
Main Biological Networks

**Interactome:** Undirected network. Nodes=proteins, edges=protein interactions. Scale-free structure (due to gene duplication) => resistance to random perturbation while sensitivity to direct “attacks”. Hubs=conserved and related to essentiality. Small world. Modular (topological clusters= functional clusters (biological processes, macromolecular complexes, …).

**Metabolome:** Directed network. Nodes=metabolites, edges=chemical transformations. Scale-free structure => resistance to random perturbation while sensitivity to direct “attacks”. Small world (=> short and efficient chemical transformations). Modular, topological clusters~ classical metabolic pathways.

**Regulome:** Directed network. Nodes=genes/proteins, edges=gene regulatory relationships (TF->gene). Scale free for out-degree, exponential for in-degree. Modular, topological clusters= biological processes. Bottlenecks more related to essentiality than hubs. Presence of overrepresented small functional motifs such as the “feed-forward loop”.

**Other:** genetic nets, phosphorilome, co-expression networks, miRNA-target networks, …
Monogenic vs. complex diseases

Monogenic disease

CFTR gene → Mutation → Cystic fibrosis

Complex disease

Gene
Environment parameter
Gender → Age → Anatomy

- Anemia falciforme
- Fibrosis quística
- Fenilcetonuria
- Enfermedad de Batten
- Hemocromatosis
- Deficiencia de alfa-1 antitripsina
- Enfermedad de Huntington
- Enfermedad de Marfan
- Distrofia muscular de Duchenne
- Síndrome de cromosoma X frágil
- Hemofilia A
- ...

- Cáncer
- Autismo
- Diabetes
- Obesidad
- Alzheimer
- Asma
- Parkinson
- Esclerosis múltiple
- Osteoporosis
- ...

The traditional reductionist approach to diseases is based on locating THE gene associated to the disease, so that it can serve as marker for diagnosing the disease and, eventually, as target for curing it.
The R&D investment required to bring a new drug to market has tripled, from $770 million per new molecular entity in 1999 to $2.3 billion in 2010. http://www.genengnews.com/

A limit in the traditional drug-development workflow?

The limitations of the reductionist approach are reflected, among other things, in the reduction of new drugs brought to the marked in spite of the increased inversion.
Networked systems require “networked” drugs

The intrinsic complexity of biological systems, reflected in the molecular networks is the main cause for the limitations of the reductionist approach.

Networked systems might require multicomponent interventions to modulate signalling outputs. a | Targets at divergent pathway nodes might cause undesired side effects when acted on in isolation. For example, AKT regulates several downstream pathways, for example, without affecting glycogen metabolism, we would need multicomponent drugs to specifically inhibit these two downstream pathways rather than using a single AKT inhibitor. b | Redundant pathways can compensate for inhibition of another pathway. For example, SMAD2 and SMAD3 perform largely similar functions in tissue culture experiments. A small-molecule inhibitor of either SMAD2 or SMAD3 alone would therefore not be effective at blocking transforming growth factor-β (TGF-β) signalling if cells responded by upregulating a redundant SMAD. TGF-β regulates several downstream outputs, so inhibiting this protein on its own could cause undesired effects by inhibiting SMAD-independent TGF-β effects, such as activation of mitogen-activated protein kinase (MAPK) signalling. Using multicomponent interventions to simultaneously inhibit SMAD2 and SMAD3 would overcome both these problems by blocking SMAD-dependent TGF-β effects without inhibiting SMAD-independent TGF-β effects. TGFR, TGF-β receptor.

Even in “classic” monogenic diseases the causative gene(s) are immersed in molecular networks and hence, at least the severity of the disease depends on many other genes/mutations. Eg. Cistic fibrosis (CFTR gene).

Reductionist vs. systemic approaches to diseases

Diagnosis (markers)

Treatment
Diseases as re-wiring of molecular networks

For the network/systemic approaches to diseases, these are caused by perturbations (e.g. re-wiring) of large networks, instead of single genes.

Human molecular networks

- Protein interactions
- Metabolic network
- Gene regulatory network
- Others… co-expression nets, phosphorilome, RNA nets, combinations of the above …

Associated dynamical data:
- Gene expression
- Metabolomics

Main characteristics of the human networks:
- Modularity
- Scale free structure => hubs
- Small world
- Presence of small motifs
Hubs

- Upregulated genes associated to lung carcinoma tend to have higher degree than unchanging ones.

- 346 proteins involved in Cancer have twice as many interactors as “non-cancer” proteins

- Disease proteins in the OMIM Morbid Map have more protein–protein interactions than do non-disease proteins

- Nevertheless, disease genes ≠ essential genes (not viable => no disease)

Disease-related and essential genes in the interactome

Disease-related proteins tend to have more interactors than the average but they are not hubs. Hubs tend to be essential proteins and hence do not lead to disease but to (embryonic) death.

“Hubs” are more related to essentiality than to disease. “Bottlenecks” are more related to diseases, at least in Cancer.

Network modules

**Module**: group of nodes highly connected among themselves and poorly connected to the rest of the network

In general... function/role separable from the rest of the network

In biological nets... correspondence with “functional” modules/pathways
Diseases and symptoms are related to Network modules

Example of a modular clinical sign: café-au-lait spots.

Genes known to be associated to diseases and clinical signs tend to cluster in molecular networks.


Diseases and symptoms are related to network modules (=functional pathways)

Even in very complex diseases involving hundreds/thousands of genes, these tend to concentrate in a reduced number of modules/pathways

Relationship between network topological modules, functional pathways and diseases

Working model of network-based approaches to diseases:

“Disease Z is due to a malfunctioning in biological system Y, whose function is carried out by a group of proteins working together (reflected in topological module X)”

Identifying disease-related modules
General strategy

Altered genes
a
b
c
e
i
m
...

Map genes to net

"Extend/restrict" based on network criteria (e.g. topological clusters, minimum paths, network propagation, ...)

Kim, Y.-A. & Przytycka, T. M. Bridging the Gap between Genotype and Phenotype via Network Approaches. Front. Genet. 3,
Identifying disease-related modules
General advantages

- Identify genes not altered (or not passing threshold for being selected as “altered”) but eventually important. E.g. “f” (connecting a and e)

- Discard “unrelated” genes (experimental errors, among other things). E.g. “I”.

- Provide additional molecular information on the disease/alteration. E.g. Disease related to pathway “1”

- Identify other potential targets eventually more “drugeable” (e.g. f, d)

- Design re-wiring strategies for recovering, for example, a malfunctioning module.

Identifying disease-related modules
Network propagation – General strategy

Identifying disease-related modules
Network propagation with Cytoscape

Identifying disease-related modules
Network propagation

“HOTNET”

Identifying disease-related modules

Identifying disease-related modules


Identifying disease-related modules
Combining genotypic and phenotypic alterations

Altered genes

Map genes to net

“Extend/restrict” based on network criteria (e.g. topological clusters, minimum paths, network propagation, …)

Kim, Y.-A. & Przytycka, T. M. Bridging the Gap between Genotype and Phenotype via Network Approaches. Front. Genet. 3,
Identifying disease-related modules
Combining genotypic and phenotypic alterations

Genotypic variation

Phenotypic variation

Distance-based Information/"electric current"/flow -based

Identifying disease-related modules
Combining genotypic and phenotypic alterations
Distance-based approaches

Look for (minimum) pathways of protein/genes which best explain the long-range relationships observed

Identifying disease-related modules
Combining genotypic and phenotypic alterations
Flow-based approaches


Design therapeutic strategies with networks in mind

Sometimes it is better to break than trying to repair

Design therapeutic strategies with networks in mind

Select which nodes to touch (inhibit with drug) in the network to get the desired rewiring

<table>
<thead>
<tr>
<th>GENOTOXIN</th>
<th>INHIBITOR</th>
<th>DMSO</th>
<th>INHIB</th>
<th>GENOTOX COMBO</th>
<th>PRE</th>
<th>POST</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR (10 GY)</td>
<td>ERLOTINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P0-90059</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAPAMYCIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NVP-BEX-235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WORTMANNIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAMPTOTHECIN</td>
<td>ERLOTINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CISPLATIN</td>
<td>ERLOTINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETOPOSIDE</td>
<td>ERLOTINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOXORUBICIN</td>
<td>ERLOTINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P0-90059</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BMS-345541</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAPAMYCIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NVP-BEZ-235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WORTMANNIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEMOZOLOMIDE</td>
<td>ERLOTINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAXOL</td>
<td>ERLOTINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IRETTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LAPATINIB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAPAMYCIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NVP-BEZ-235</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WORTMANNIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Create a synthetic net that restores the function of the malfunctioning one


Biomarkers, for diagnosis/prognosis

Networks are better markers of an alteration/disease than single genes or combinations of them.

Biomarkers, for diagnosis/prognosis

Network-based patient stratification (personalized medicine)

Patient networks can be used to identify patient subgroups which, for example, present different responses to drugs.


Drug repositioning

It often takes 10–15 years of research and 0.8–1.5 billion dollars to bring a drug from abstract concept to market-ready product [1]. Every year, ~90% of drugs fail during FDA evaluations, preventing their use in actual therapy.

Drug repositioning (DR, aka d. repurposing, redirecting, retargeting, …) seeks to find new uses for existing drugs, with established and demonstrated human safety. Usually this involves finding new targets for approved drugs.

Computational methods for DR include molecular docking, machine learning, literature mining, … and network-based approaches.
“We deployed algorithms relying on artificial intelligence, network diffusion, and network proximity […] to rank 6,340 drugs for their expected efficacy against SARS-CoV-2. We screened in human cells the top-ranked drugs, obtaining a 62% success rate. Of the six drugs that reduced viral infection, four could be directly repurposed to treat COVID-19. We also found that 76 of the 77 drugs that successfully reduced viral infection do not bind the proteins targeted by SARS-CoV-2, indicating that these network drugs rely on network-based mechanisms that cannot be identified using docking-based strategies.

Other Networks used for studying human pathologies

Networks specifically generated for studying human pathologies

- Drug-target interactions (DTI)
- Drug-drug interactions (based on chemical similarity, biological effect similarity, target(s) similarity, …)
- Drug-disease associations
- Drug-side effect associations
- Disease-disease associations
- Integrated (combined) networks
- …

Apart from the “generic” molecular networks (interactome, metabolome, gene regulation, …) some networks were generated to specifically represent pathology-related information

Other networks used for studying human pathologies

- Social Networks
- Technological networks
- …

"network pharmacology"
Drug-target network

Summary

• Any phenomenon that can be modelled as entities linked by relationships can be represented as a network and studied with the tools of Graph Theory.

• Many phenomena in different disciplines have been studied from this point of view, including molecular phenomena as soon as the data required to assemble these networks became available.

• Biological (molecular) networks are the prototypic subject of study of Systems Biology, that focus on the complex network of relationships between components instead of the components themselves, hence complementing the reductionist approach of Molecular Biology.

• From the topological features of molecular networks, biological information related to their functioning, resistance to perturbations, ability to transmit information, etc. has been extracted.

• These network approaches are now being used to study and treat human pathologies.
Bibliography


*PDFs available at* [http://csbg.cnb.csic.es/pazos/cursos/UMA_BIF/](http://csbg.cnb.csic.es/pazos/cursos/UMA_BIF/)

Florencio Pazos  
@fpazos_bioinf  
pazos@cnb.csic.es  
http://csbg.cnb.csic.es