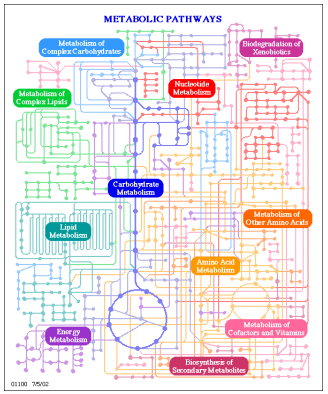
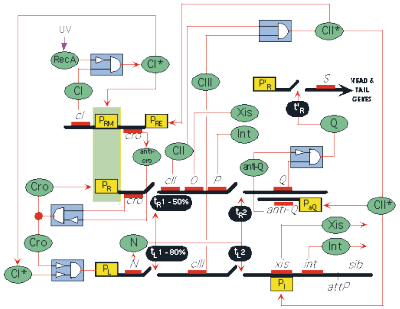
# General overview of networks (graphs) in biology & associated algorithms

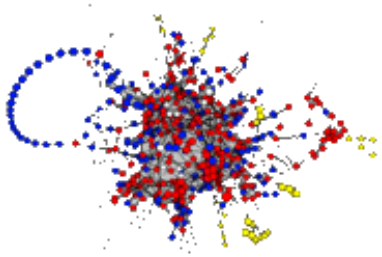
**The key abstract idea to retain is:**
**Interactions!**
**And so networks / graphs, as models or as tools**

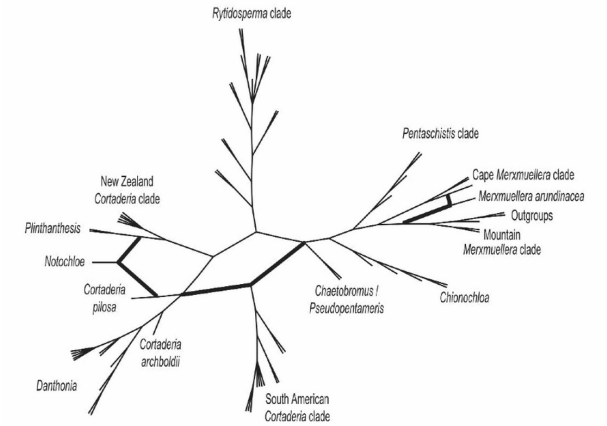
---

**Biochemical networks** … but also



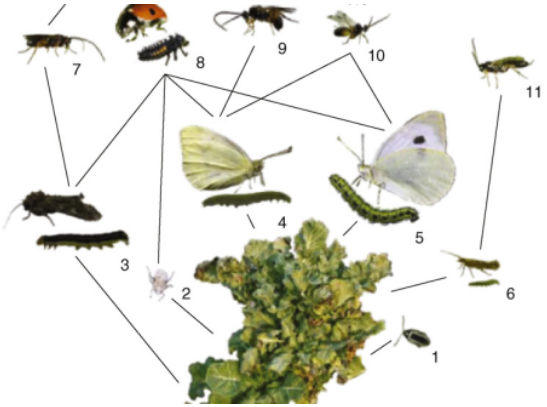



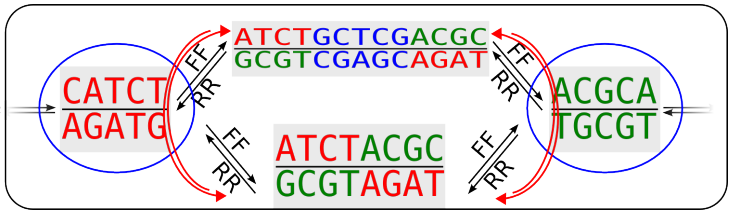



**Evolutionary network**

**"Symbiotic" network**




**Ecological network**

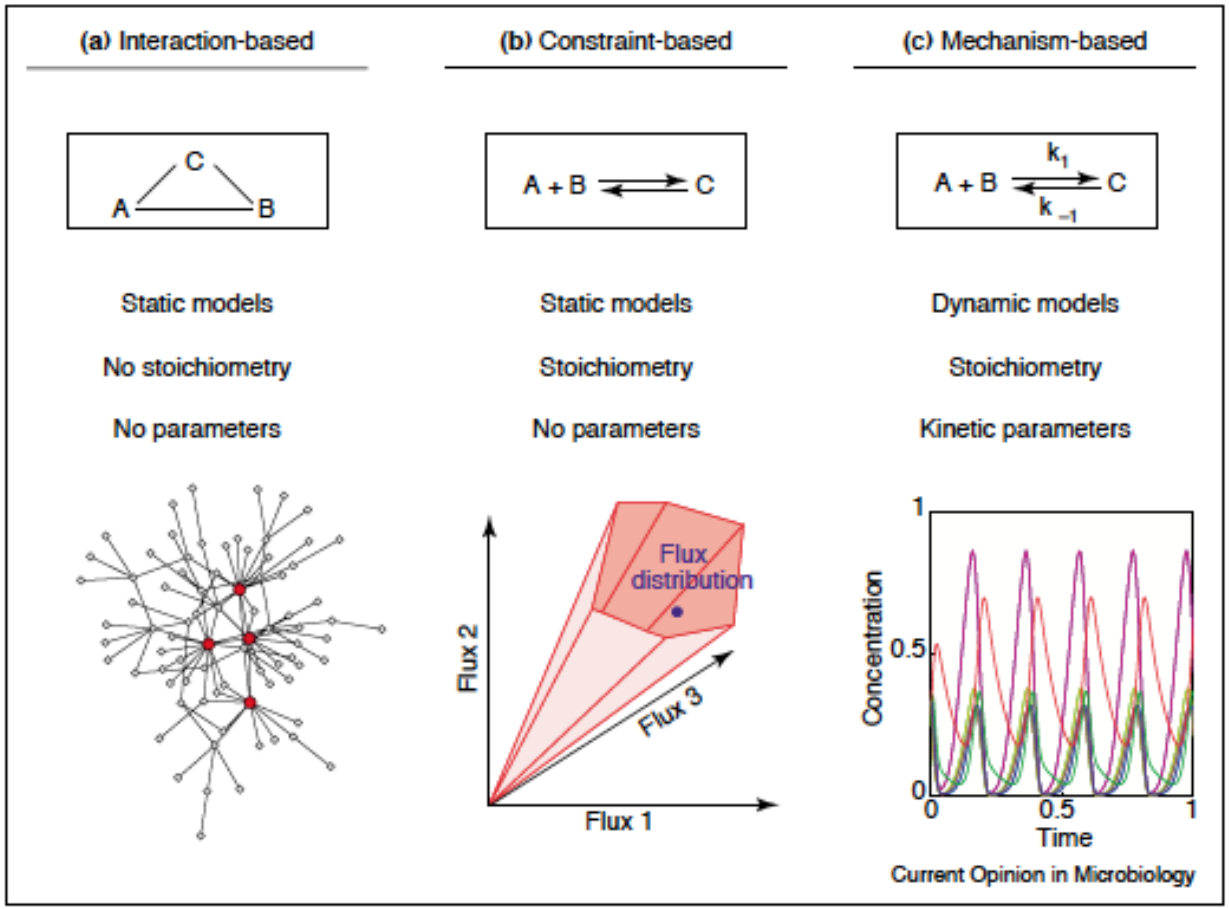



**Graphs as "tools"**
**De Bruijn graphs for NGS data**

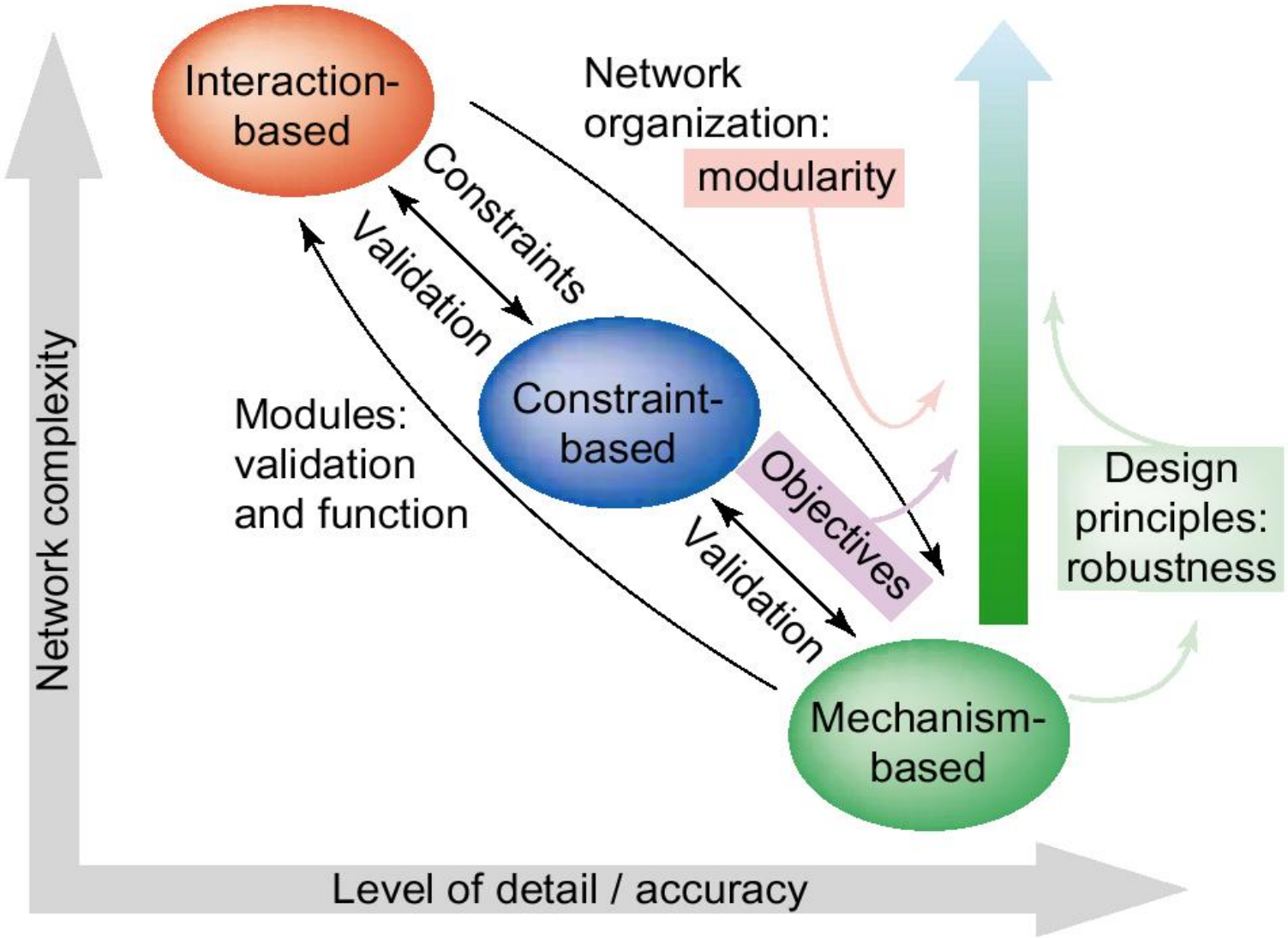
**This overview will be biased**
**towards a "graph-view" of networks**
**(… and even then, very far from being exhaustive!)**

---

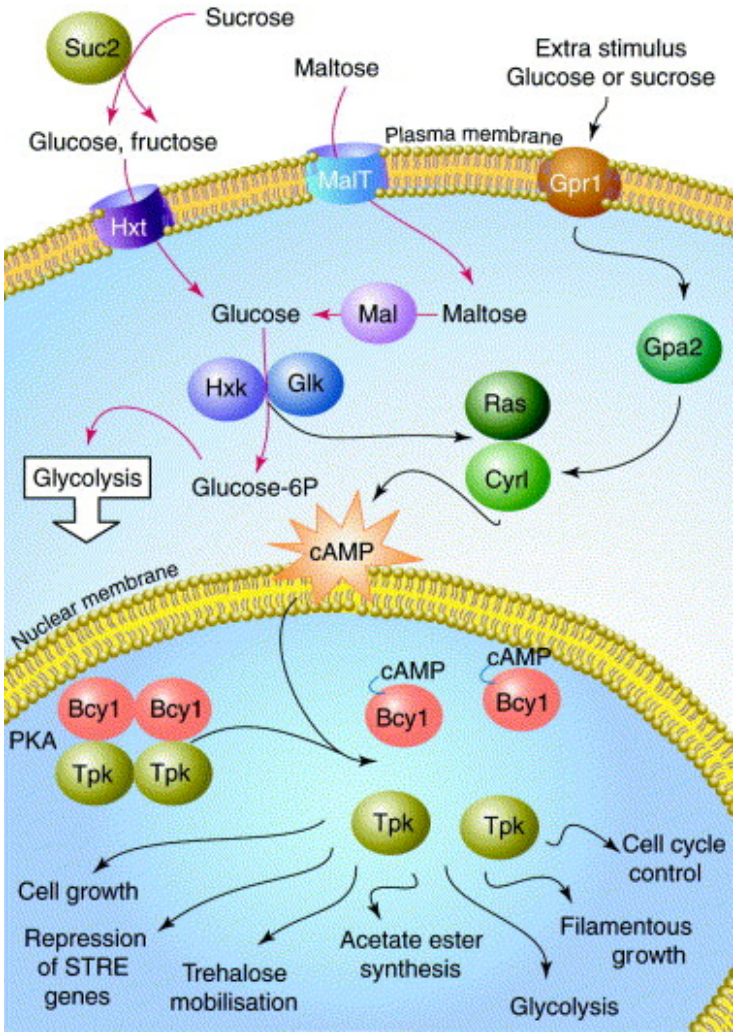
**Case of biochemical networks, more precisely of metabolism**
**Modelling of the biochemical reaction: A + B ⟷ C**



# Another view



# Modelling




$$
\begin{aligned}
\max \quad & v_{r^o} \\
\text{s.t.} \quad & Sv = 0 \\
& v_j = 0 \; \forall j \in F \\
& \textstyle\sum_j v_j \leq 1 \\
& v_j \geq 0 \; \forall j \notin F \cup r^o.
\end{aligned}
$$

**Algorithm RC (Reaction Cut)**

**input:**
a stoichiometric matrix $S$, a weight function $w$, a reaction $r^o$ to be cut;

**phase 1**
$F = \emptyset$;
**while** $F$ is not a reaction cut of $r^o$
**do begin**
let $C$ be the set of reactions defining an elementary mode in $S_F$ that includes $r^o$
let $\bar{w} = \min_{r \in C} w(r)$
**for** each reaction $r$ in $C$
**do begin**
$w(r) = w(r) - \bar{w}$
**if** $w(r) = 0$ **then** $F = F \bigcup \{r\}$
**end**
**end**

**phase 2**
let $r_1, r_2, \ldots, r_k$ be the reaction in $F$
**for** $j = 1$ **to** $k$ **do**
**if** $F - r_j$ is a reaction cut of $r^o$ **then** $F = F - r_j$

**output:** F

# Modelling

**Taken from "The role of Modeling in Systems Biology", Douglas Kell and Joshua Knowles**
**Chapter of "System Modeling in Cellular Biology: From Concepts to Nuts and Bolts", eds. Zoltan Szallasi, Jorg Stelling, Vipul Periwal, MIT Press 2006**

| ***Dimension or Feature*** | ***Possible choices*** | ***Comments*** |
|---|---|---|
| Stochastic or deterministic | Stochastic: Monte Carlo methods or statistical distributions<br>Deterministic: equations such as ODEs | Phenomena are not of themselves either stochastic or deterministic; large-scale, linear systems can be modeled deterministically, while a stochastic model is often more appropriate when nonlinearity is present. |
| Discrete versus continuous (in time) | Discrete: Discrete event simulation, for example, Markov chains, cellular automata, Boolean networks.<br>Continuous: Rate equations. | Discrete time is favored when variables only change when specific events occur (modeling queues). Continuous time is favored when variables are in constant flux. |

# Modelling

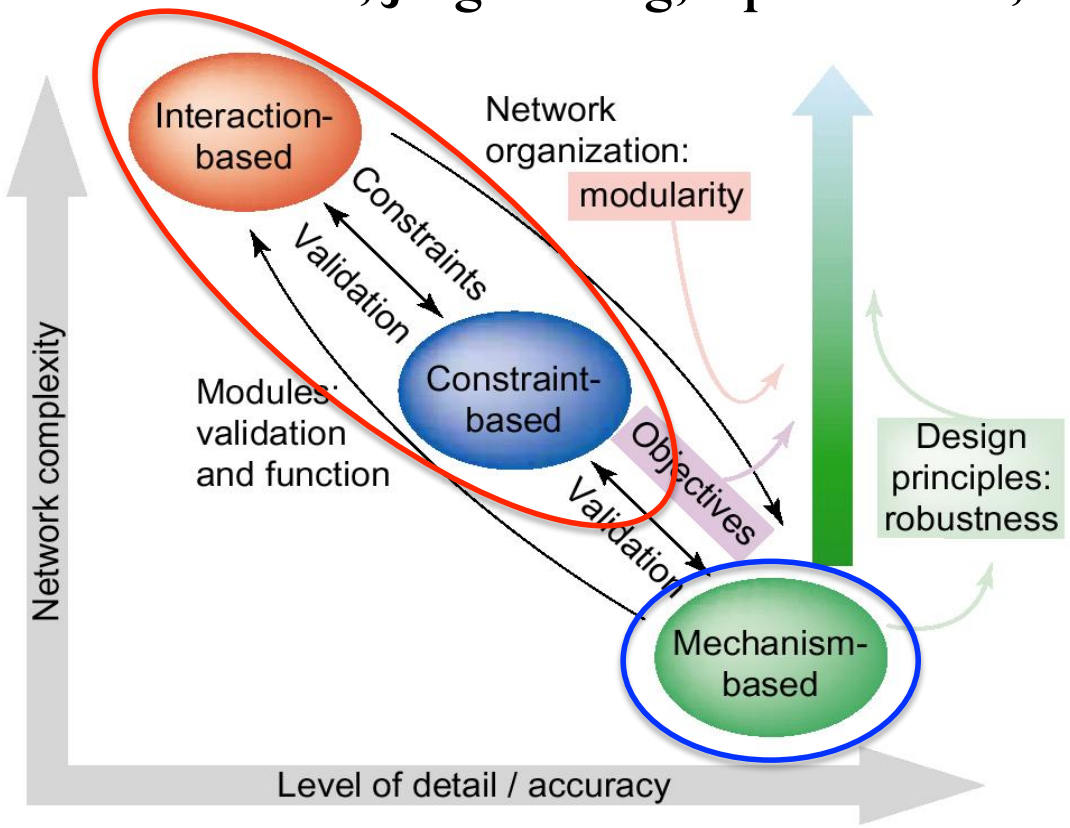
| *Dimension or Feature* | *Possible choices* | *Comments* |
|---|---|---|
| Macroscopic versus microscopic | Microscopic: Model individual particles in a system and compute averaged effects as necessary. Macroscopic: Model averaged effects themselves, for example, concentrations, temperatures, etc. | Are the individual particles or subsystems important to the evolution of the system, or is it enough to approximate them by statistical moments or ensemble averages? |
| Hierarchical versus multi-level | Hierarchical: Fully modular networks. Multi-level: Loosely connected components. | Can some processes/variables in the system be hidden inside modules or objects that interact with other modules, or do all the variables interact, potentially? This relates to reductionism versus holism. |
| Fully quantitative versus partially quantitative versus qualitative | Qualitative: Direction of change modeled only, or on/off states (Boolean network). Partially quantitative: Fuzzy models. Fully quantitative: ODEs, PDEs, microscopic particle models. | Reducing the quantitative accuracy of the model can reduce complexity greatly and many phenomena may still be modeled adequately. |

# Modelling

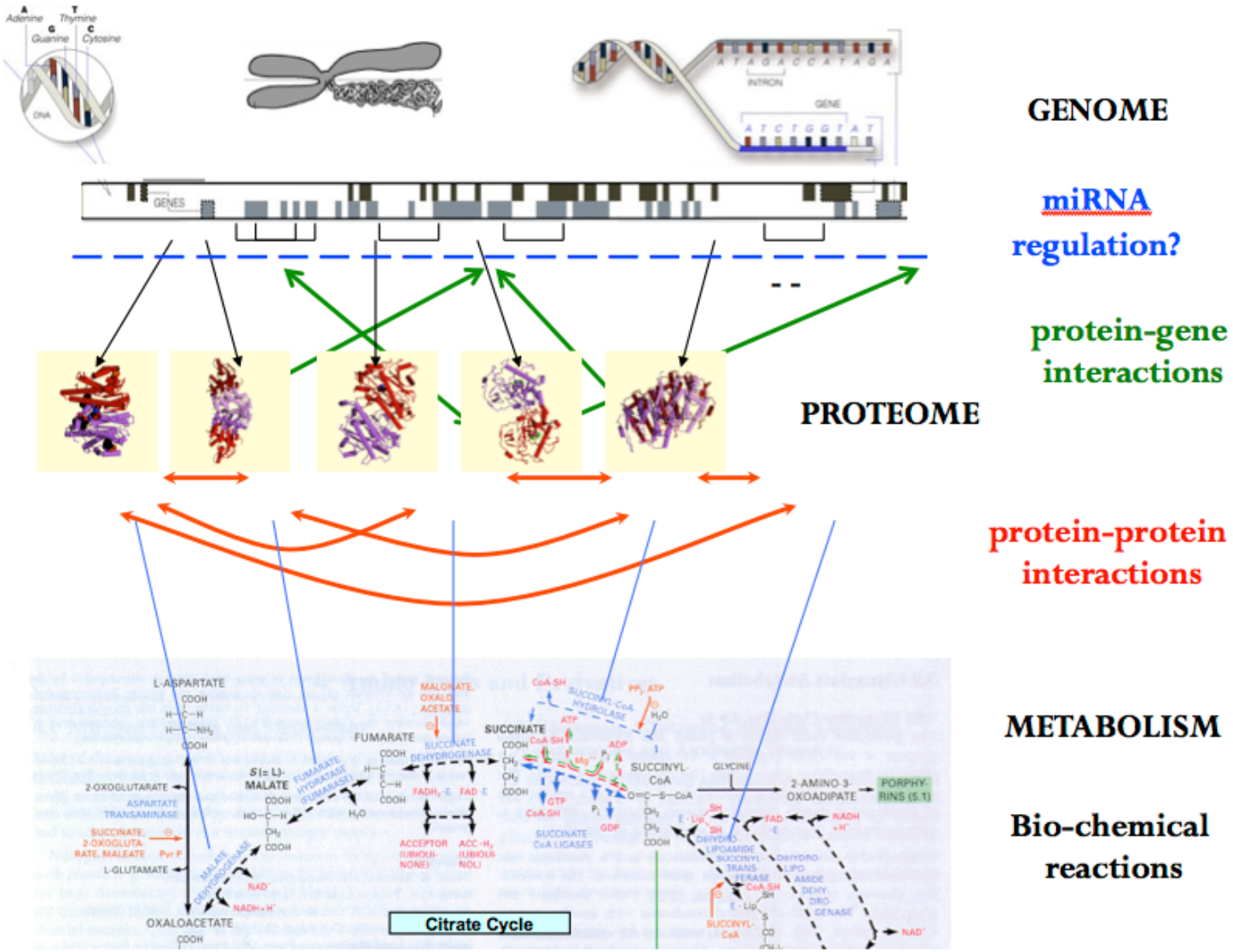
| *Dimension or Feature* | *Possible choices* | *Comments* |
|---|---|---|
| Predictive versus exploratory/explanatory | Predictive: Specify every variable that could affect outcome.<br>Exploratory: Only consider some variables of interest. | If a model is being used for precise prediction or forecasting of a future event, all variables need to be considered. The exploratory approach can be less precise but should be more flexible, for example, allowing different control policies to be tested. |
| Estimating rare events versus typical behavior | Rare events: Use importance sampling.<br>Typical behavior: Importance sampling not needed. | Estimation of rare events, such as apoptosis times in cells is time-consuming if standard Monte Carlo simulation is used. Importance sampling can be used to speed up the simulation. |
| Lumped or spatially segregated | Lumped: Treat cells or other components/compartments as spatially homogeneous.<br>Spatially segregated: Treat the components as differentiated or spatially heterogeneous. | If heterogeneous it may be necessary to use the computationally intensive partial differential equation, though other solutions are possible (Mendes and Kell, 2001) |

# Modelling

**"The role of Modeling in Systems Biology", Douglas Kell and Joshua Knowles**
**Chapter of "System Modeling in Cellular Biology: From Concepts to Nuts and Bolts", eds. Zoltan Szallasi, Jorg Stelling, Vipul Periwal, MIT Press 2006**

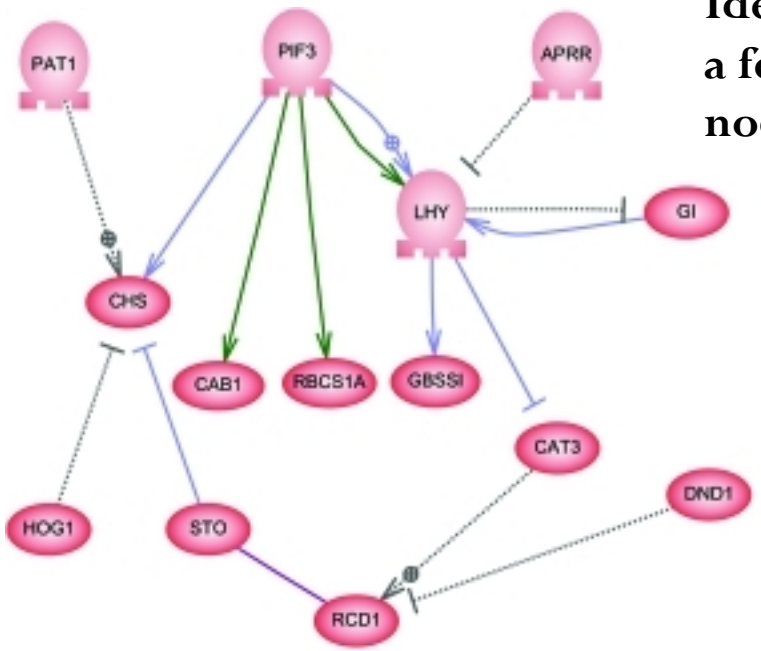


# Biochemical networks
## An overview

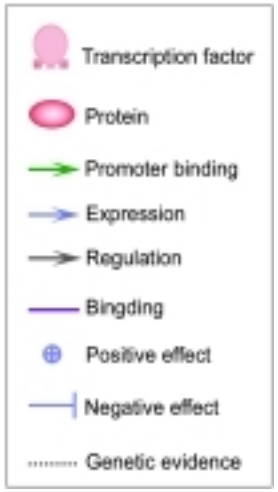


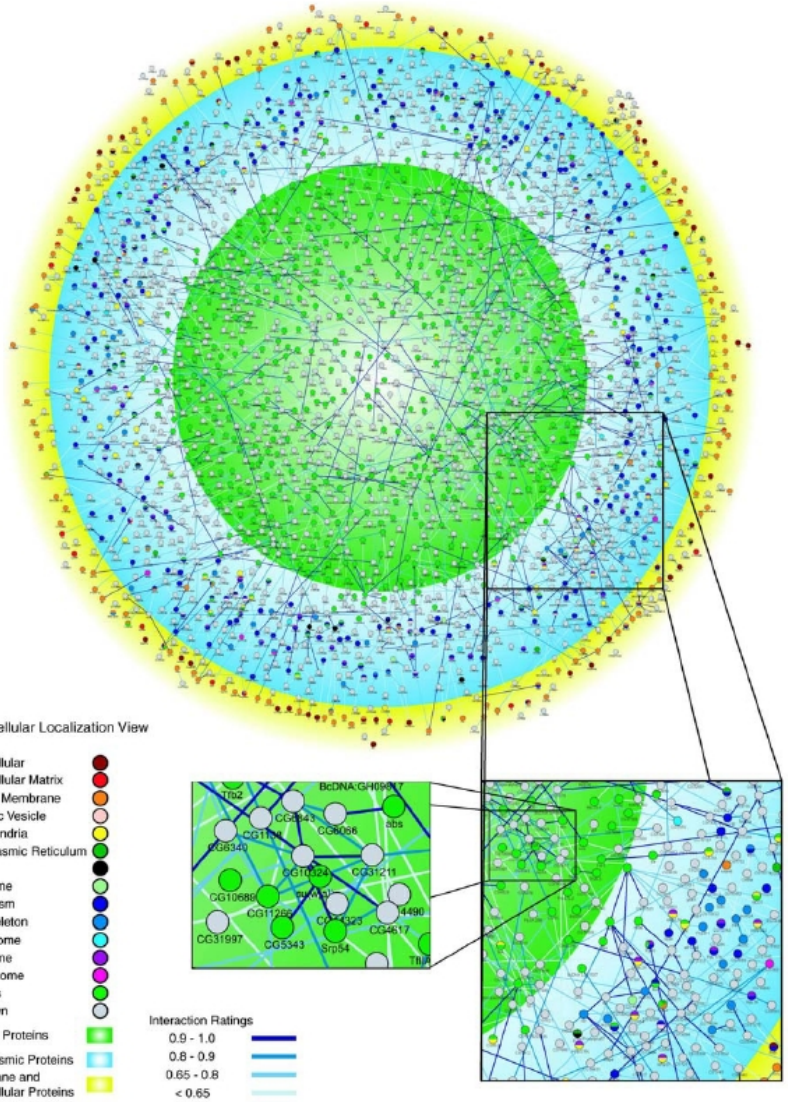
# Gene-protein and protein-protein networks

**Protein ~ Macromolecule**
**Edges/Arcs may have a sign**
**(indicating positive or negative effect)**




**Idea of sizes:**
**a few thousands**
**nodes and edges/arcs**



## Metabolic networks

**You'll have more details later in the course, but for now already, the basic information on such networks**

**As I mentioned, three main types of representations:**

**Graph representation: Connectivity of reactions/metabolites, structure of the metabolic network**

**Stoichiometric (reaction equation) representation: capabilities of the network, flow analysis, steady-state analyses**

**Kinetic models: dynamic behaviour under changing conditions**

## The representations that will be used

**You'll have more details later in the course, but for now already, the basic information on such networks**

**As I mentioned, three main types of representations:**

**Graph representation: Connectivity of reactions/metabolites, structure of the metabolic network**

**Stoichiometric (reaction equation) representation: capabilities of the network, flow analysis, steady-state analyses**

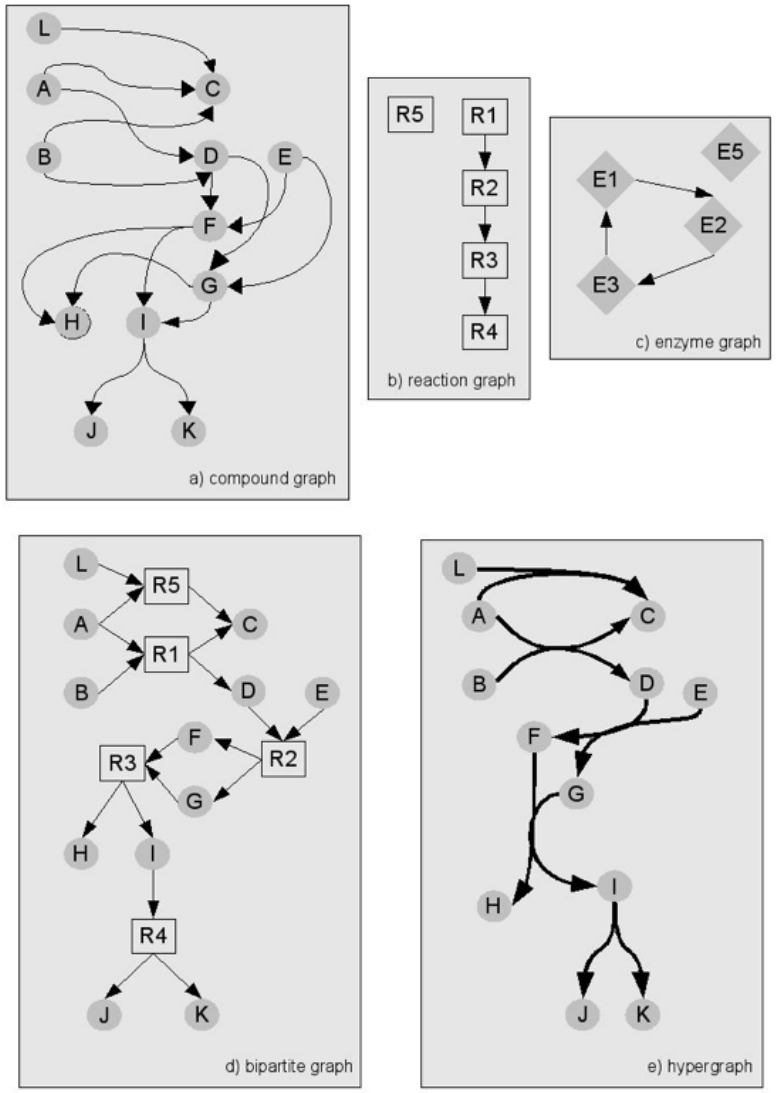
**Kinetic models: dynamic behaviour under changing conditions**

# Graph representation, or directed hypergraph

reaction

compound / metabolite

R1: A + B -> C + D
R2: D + E -> F + G
R3: F + G -> H + I
R4: I -> J + K
R5: A + L -> C



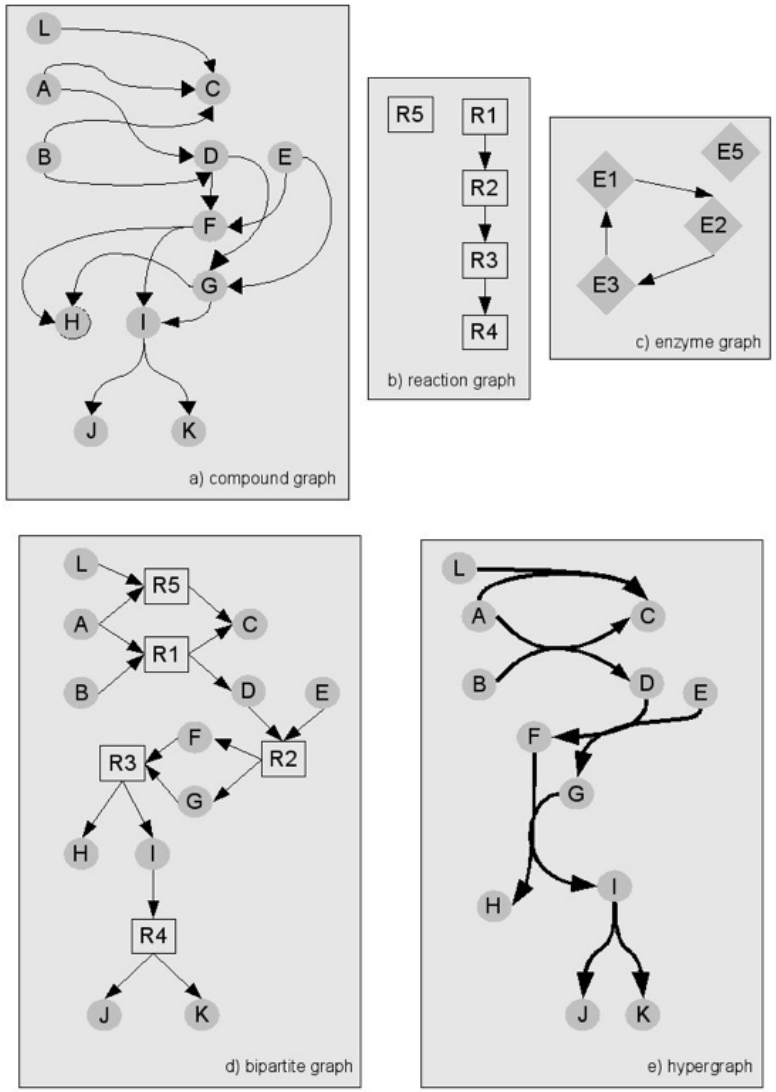
# Valued directed (hyper)graphs

compound / metabolite

reaction

R1: 1A + 2B -> 2C + 3D
R2: 3D + 1E -> 2F + 2G
R3: 2F + 1G -> 2H + 1I
R4: 3I -> 1J + 2K
R5: 1A + 3L -> 2C

Valued directed (hyper)graphs



# Stoichiometric matrix

reaction

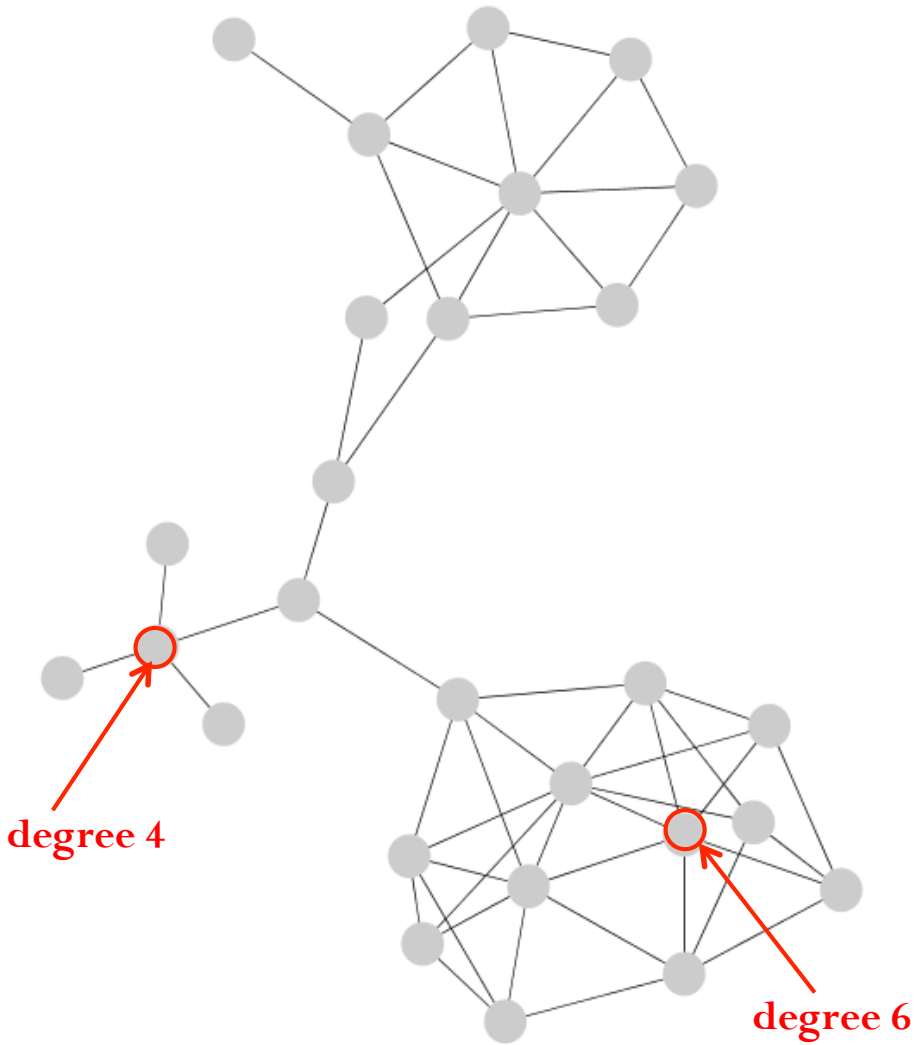
compound / metabolite

R1: 1A + 2B -> 2C + 3D
R2: 3D + 1E -> 2F + 2G
R3: 2F + 1G -> 2H + 1I
R4: 3I -> 1J + 2K
R5: 1A + 3L -> 2C

| | R1 | R2 | R3 | R4 | R5 |
|---|---|---|---|---|---|
| A | -1 | 0 | 0 | 0 | -1 |
| B | -2 | 0 | 0 | 0 | 0 |
| C | +2 | 0 | 0 | 0 | +2 |
| D | +3 | -3 | 0 | 0 | 0 |
| E | 0 | -1 | 0 | 0 | 0 |
| F | 0 | +2 | -2 | 0 | 0 |
| G | 0 | +2 | -1 | 0 | 0 |
| H | 0 | 0 | +2 | 0 | 0 |
| I | 0 | 0 | +1 | -3 | 0 |
| J | 0 | 0 | 0 | +1 | 0 |
| K | 0 | 0 | 0 | +2 | 0 |
| L | 0 | 0 | 0 | 0 | -3 |

# What has been done in the literature
# Computing indices

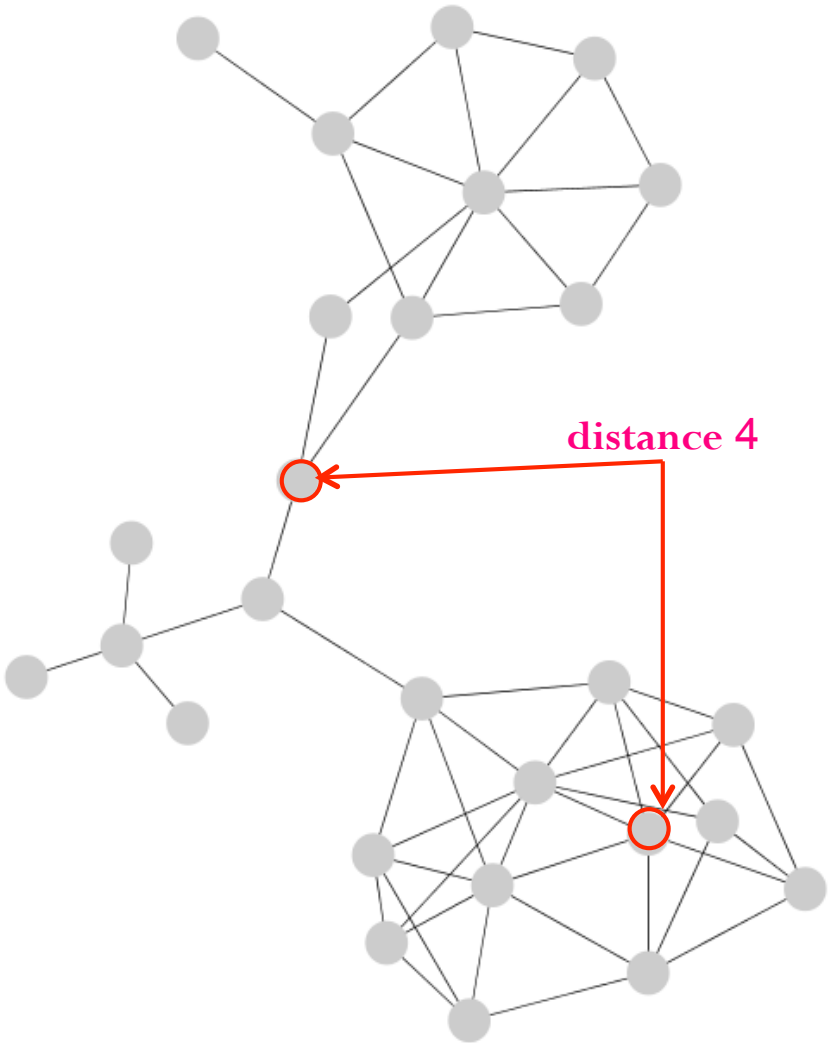
## Degree distribution



# Computing indices

**Degree distribution**

**Distance distribution & diameter**

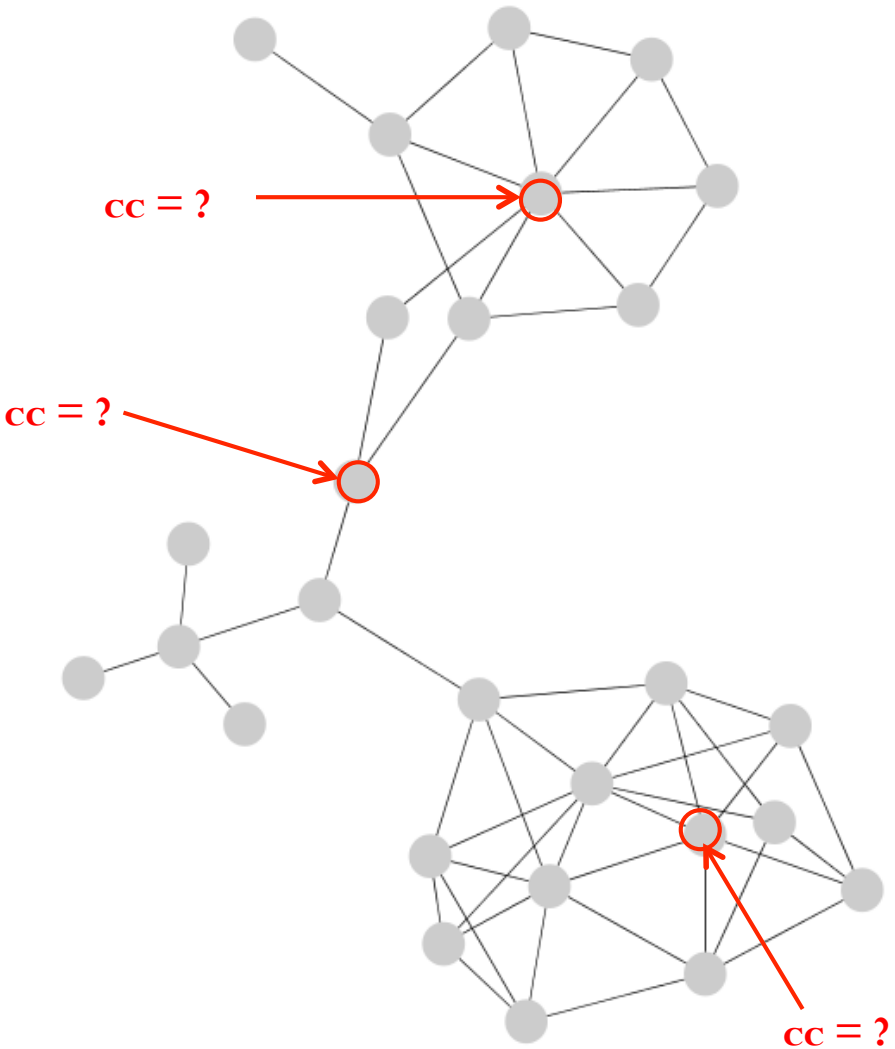


# Computing indices

**Degree distribution**

**Distance distribution & diameter**

**Clustering coefficient**



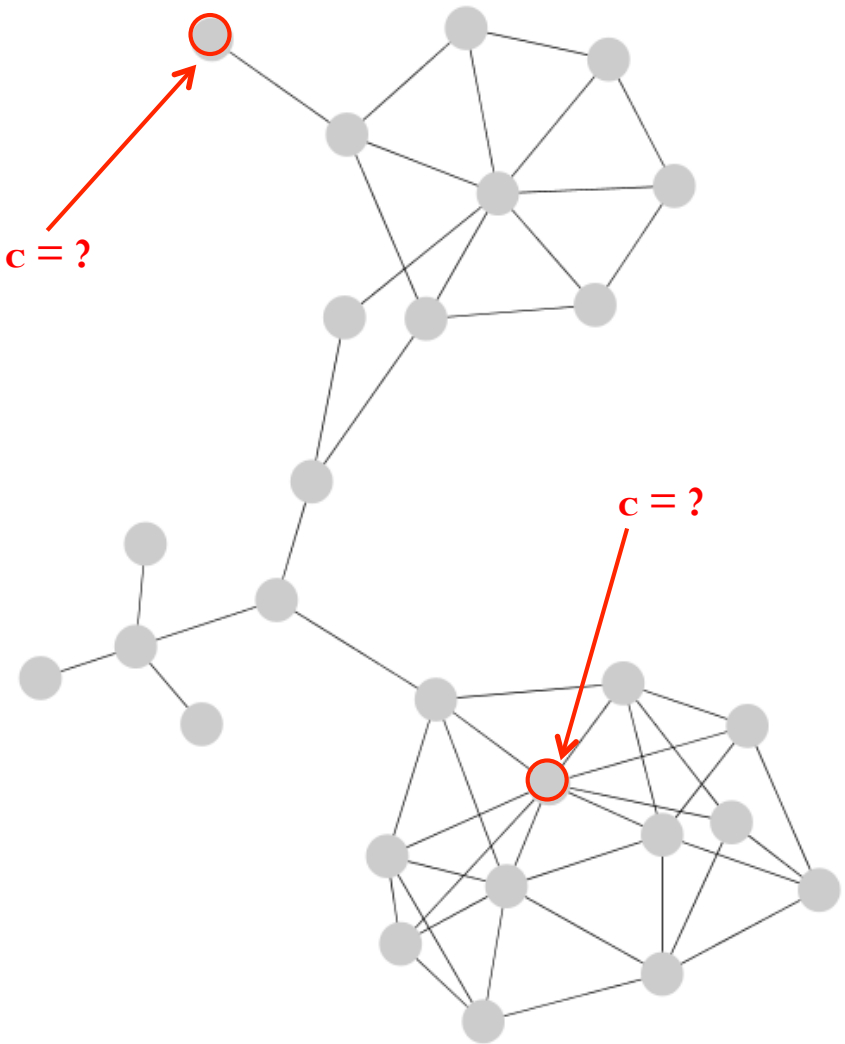
# Computing indices

**Degree distribution**

**Distance distribution & diameter**

**Clustering coefficient**

**Closeness centrality**



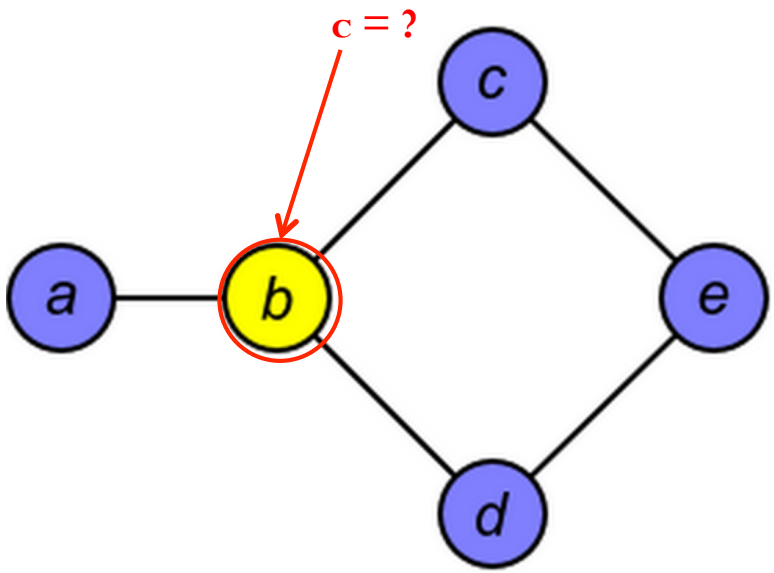
# Computing indices

**Degree distribution**

**Distance distribution & diameter**

**Clustering coefficient**

**Closeness centrality**



# Computing indices

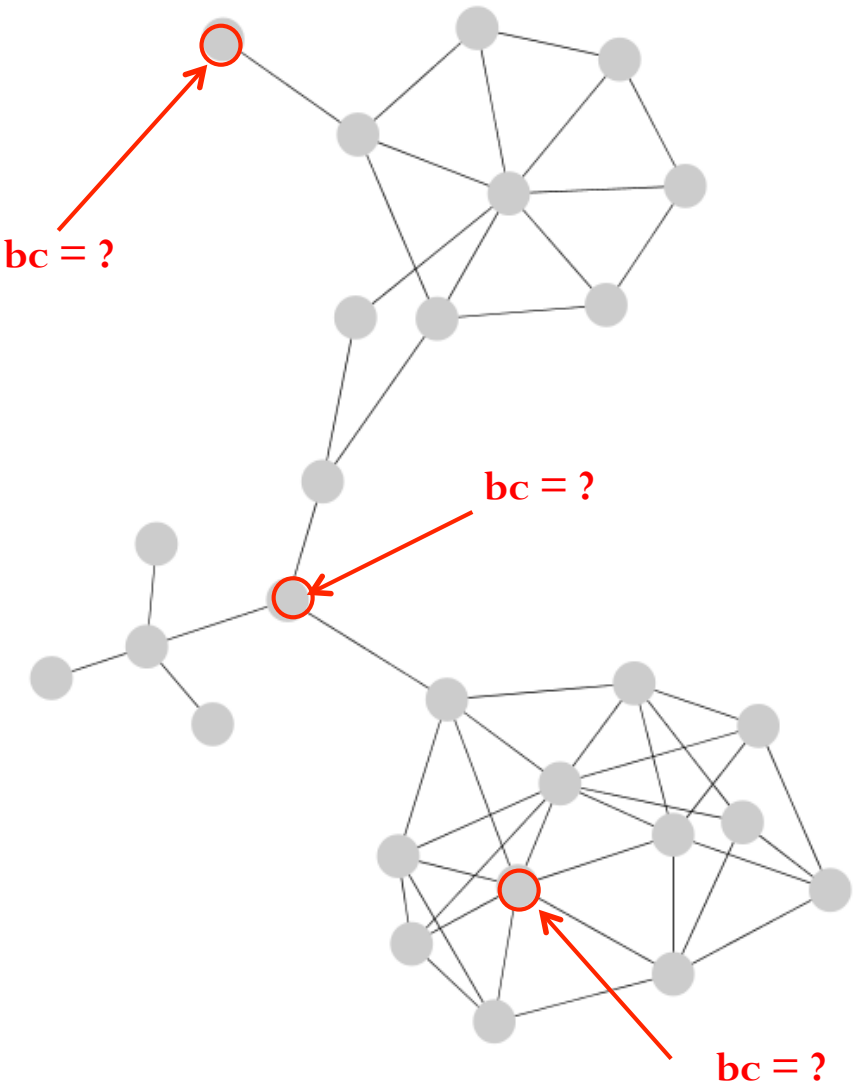
**Degree distribution**

**Distance distribution & diameter**

**Clustering coefficient**

**Closeness centrality**

**Betweenness centrality**



# Computing indices

**Degree distribution**

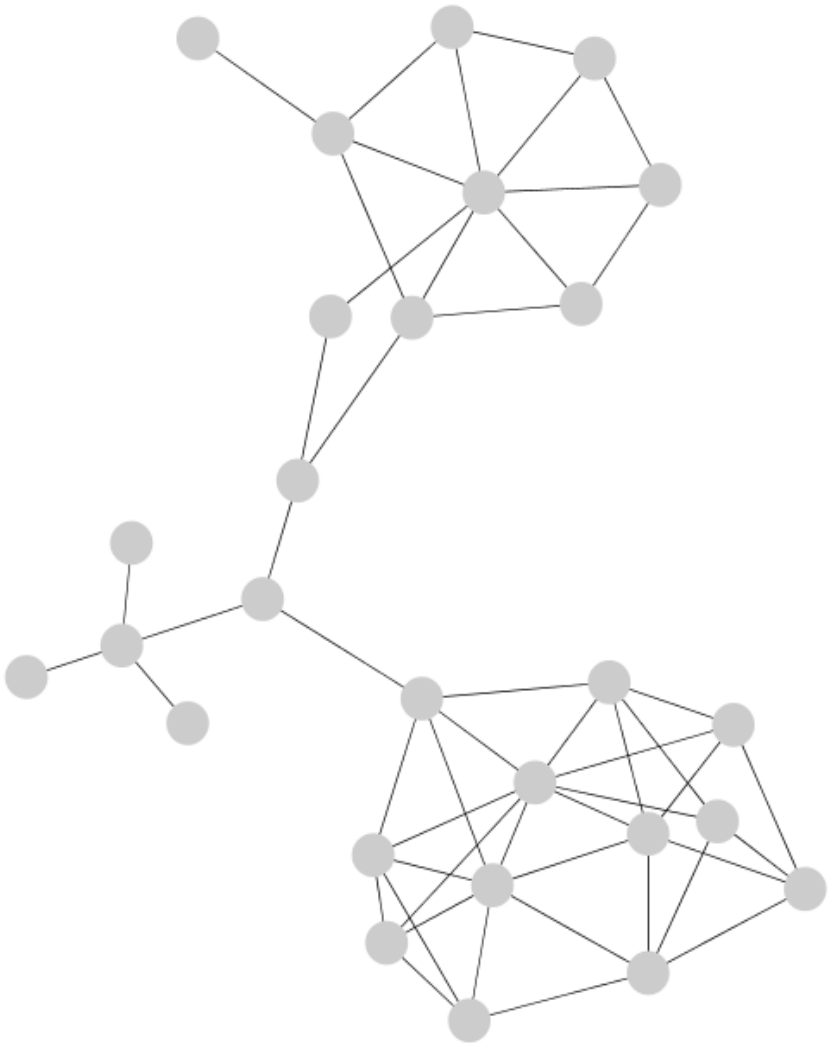
**Distance distribution & diameter**

**Clustering coefficient**

**Closeness centrality**

**Betweenness centrality**

**And there are many others…**



# Complexity of computing indices?

**Exercise**

## Some controversies…
## But first a definition:
## Scale-free property

**Scale-free network (graph): invariant to changes in scale**

**Any part of a scale-free network is stochastically similar to the whole network, and parameters are assumed to be independent of the system size (sometimes called the "self-similarity property")**

**Consider for instance the connectivity of a network: A network is defined as being scale-free in terms of its connectivity if a randomly picked node has $k$ connections with other nodes with a probability that follows a power-law $P(k) \sim k^{-\gamma}$, where $\gamma$ power-law exponent**

**Actually, literature a bit fuzzy on definition of "scale-free"**

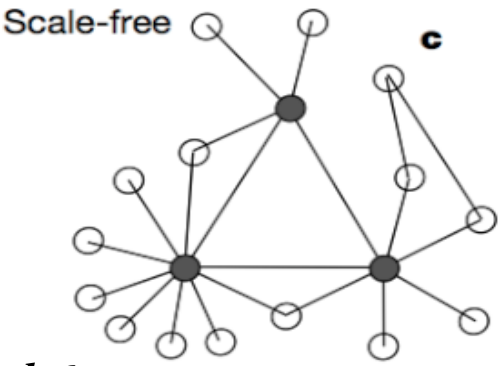
**Khanin *et al.*, *J. Comp. Biol.*, 2006**
**Li *et al.*, *Internet Math.*, 2(4):431-523, 2006**
**Bollobás & Riordan, *Internet Math.*, 1(1):1-35, 2003 / *Combinatorica*, 24(1):5-34, 2004**

## Some controversies

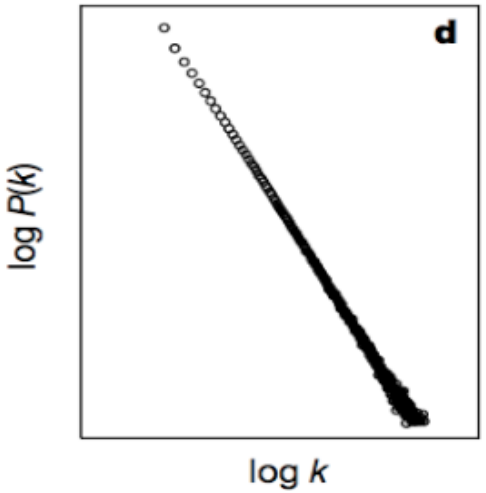
**Scale-freeness of biological networks, at least asymptotically**

**…according to Barabási and colleagues**

**Albert *et al.*, *Nature*, 1999**
**Barabási *et al.*, *Science*, 1999**
**Jeong *et al.*, *Nature*, 2000**
**Jeong *et al.*, *Nature*, 2001 etc.**

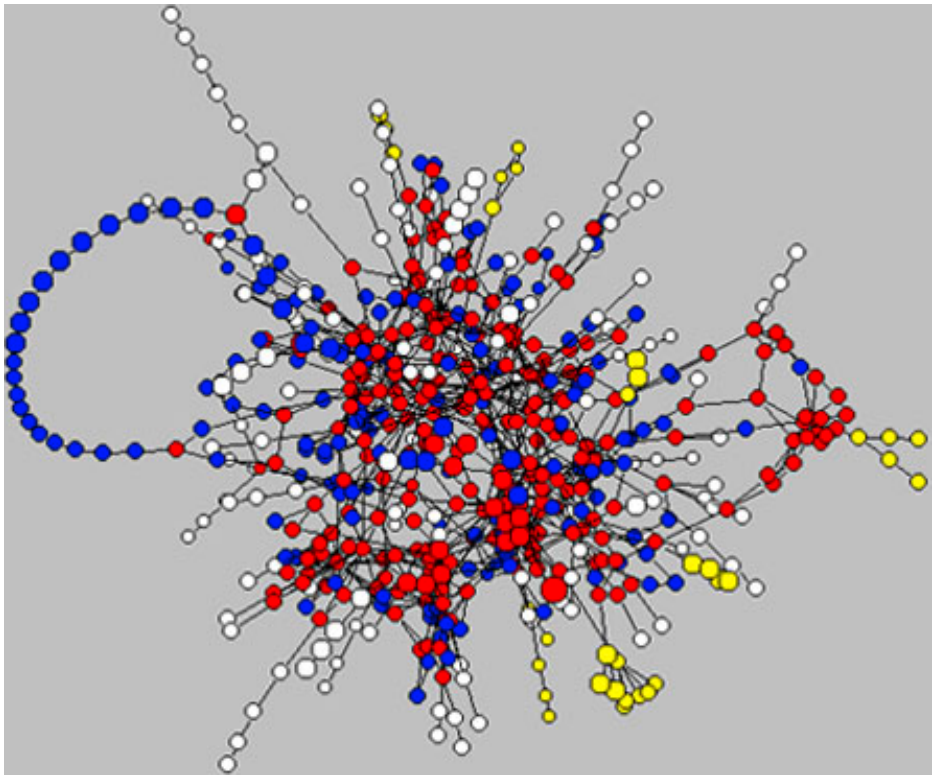



$P(k) \sim k^{-g}$



## Some controversies

**According to them also**
**Hubs are correlated with essential (critical for survival) genes ("centrality principle")**



**Jeong *et al.*, *Nature*, 2001**

## Centrality principle contradicted

**For instance, in the case of protein-protein interaction (PPI) networks, correcting for bias in data shows no correlation between essentiality of a gene and:**

- **its degree in PPI network**
- **the average degree of its neighbours**
- **its clustering coefficient**

**Main type of bias: data collected from literature, but essential genes are the objects of more papers than non-essential ones**

**Coulomb *et al.*, *Proc. Royal Society*, 2006**
**Ito and Xenarios data**

## Scale-freeness contradicted also in terms of statistical analysis

**Fitting of power-law to the data using maximum-likelihood method and goodness-of-fit test on various biological datasets:**

**6 PPI (Uetz, Schwikowski, Ito, Li, Rain, Giot); 1 gene interaction (Lee); 1 metabolic network (Ma); 2 synthetic lethal interaction data (Tong and Guelzim)**

**showed that ALL those networks DIFFER SIGNIFICANTLY from the power-law distribution, and from truncated power-law except sometimes for very small ranges, this based on a chi-squared goodness-of-fit test**

**Khanin *et al.*, *J. Comp. Biol.*, 2006**

## Other reported characteristics appear (more) robust

**There is a short path from any node to another… BUT… see later**

**There are many nodes with few connections and a few nodes with very many connections, which is what is actually observed in biological networks**

**However, many other distributions apart from power-law have similar properties (generalised Pareto law, stretched exponential, geometric distribution, geometric random graph…)**

**Khanin *et al.*, *J. Comp. Biol.*, 2006**

## Self-similarity

**What those other distributions have NOT is the self-similarity property**

**Self-similarity property:**

**any part of a scale-free network is stochastically similar to the whole network**

**Khanin *et al.*, *J. Comp. Biol.*, 2006**

## A story not without some deep consequences

**"Often, the underlying principles and assumptions of evolutionary models are adjusted so that they yield the scale-free topology of the network"**

**Khanin *et al.*, *J. Comp. Biol.*, 2006**

## A story not without some deep consequences

### Preferential attachment

Preferential attachment (PA) is a mechanism that is proposed to generate many networks occurring in nature.

- Start with a small number $n_0$ of nodes and no edges.
- Iterate the following:
  - insert a new node $v_j$,
  - draw $m \leq n_0$ edges from $v_i$ existing nodes $v_i$ with probability $p \sim \frac{k_i+1}{\sum_j (k_j+1)}$

When drawing new edges, nodes with many edges already are preferred over nodes with few or no edges.

**Barabási *et al.*, *Science*, 1999**

## A story not without some deep consequences

**"Often, the underlying principles and assumptions of evolutionary models are adjusted so that they yield the scale-free topology of the network"**

**Khanin *et al.*, *J. Comp. Biol.*, 2006**

**"Many attributed a deep significance to this fact (scale-freeness) inferring a universal architecture of complex systems. Closer examination, however challenges the assumptions"**

**Keller, *BioEssays*, 27(10):1060-1068, 2005**

# Another controversy

## Small-world graphs

Graphs fulfilling the following two criteria are called small-world graphs

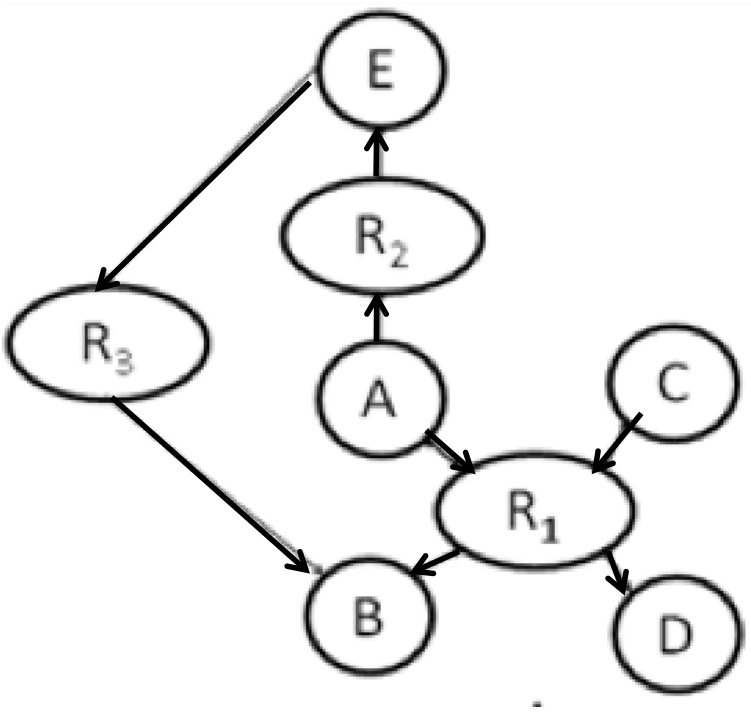
- Small average shortest path length between two nodes, the same level as ER graphs, lower than many regular graphs: shortcuts accross the graphs go via hubs
- High clustering coefficient compared to ER graph: the neighbors of nodes are more often linked than in ER graphs.

Graphs generated with preferential attachment are small-world graphs. However, small-world graphs can be generated with other mechanisms as well.

# Shortest paths in reaction or compound graphs
# May not be biologically relevant

**Example in metabolic network represented as bipartite graph**

**What is the shortest distance between A and B?**



## The main messages to retain

**The first is that even without considering the problem of noise in the data (see later), it's important to remember to do "good" mathematics/statistics/ computation (algorithmics)**

## The main messages to retain

**The first is that even without considering the problem of noise in the data (see later), it's important to remember to do "good" mathematics/statistics/ computation (algorithmics)**

**Even when this has been done, and even if the data was not noisy, one must be careful with biological interpretation:**

- **It may be wrong for various reasons**
- **It may be not informative even when it is correct**

## The main messages to retain

**The first is that even without considering the problem of noise in the data (see later), it's important to remember to do "good" mathematics/statistics/ computation (algorithmics)**

**Even when this has been done, and even if the data was not noisy, one must be careful with biological interpretation:**

- **It may be wrong for various reasons**
- **It may be not informative even when it is correct**

**These indices intervene also in another, difficult context: the one of obtaining "good" random models against which to compare biological networks in order to then draw some reliable biological conclusions**

## The main messages to retain

**The first is that even without considering the problem of noise in the data (see later), it's important to remember to do "good" mathematics/statistics/ computation (algorithmics)**

**Even when this has been done, and even if the data was not noisy, one must be careful with biological interpretation:**

**It may be wrong for various reasons**

**It may be not informative even when it is correct**

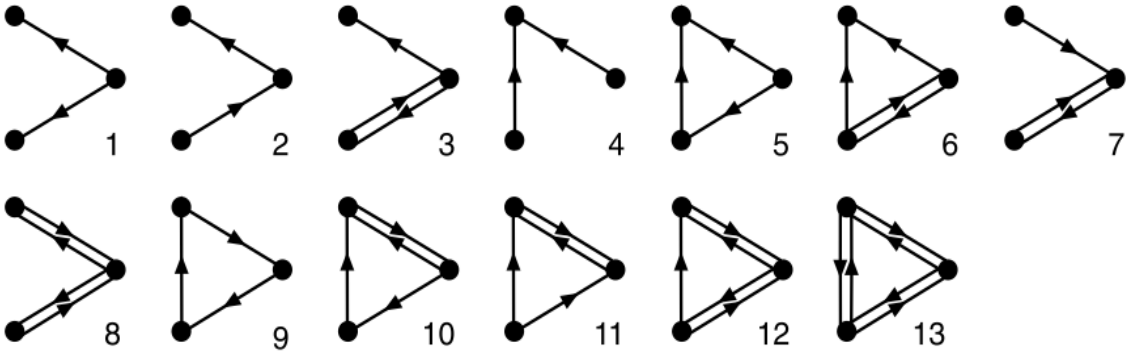
**These indices intervene also in another, difficult context: the one of obtaining "good" random models against which to compare biological networks in order to then draw some reliable biological conclusions**

**Besides the literature, you may be interested in reading some more informal comments such as those, possibly controversial, presented here:**

**https://liorpachter.wordpress.com/**

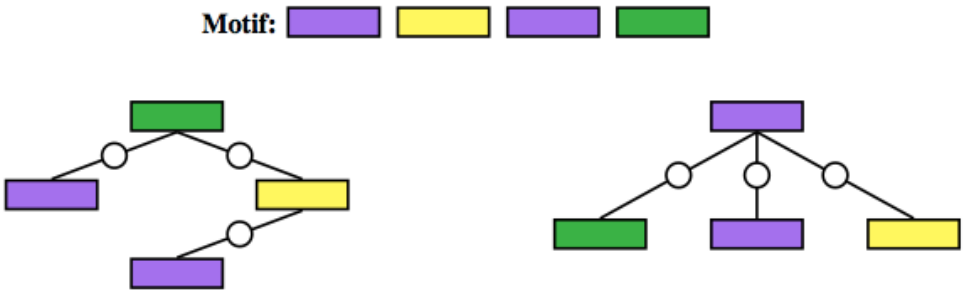
# Another topic that has been covered in the literature
# Enumerating motifs

**Different definitions have been used in the literature, depending also in whether gene-protein interaction, protein-protein interaction or metabolic networks where considered**




**Motifs as induced or non induced subgraphs**

**So-called coloured motifs**



# Enumerating motifs

**More on enumeration**

**And on motifs with Arnaud Mary**

## But recalling some main messages to retain

**The first is that even without considering the problem of noise in the data (see later), it's important to remember to do "good" mathematics/statistics/ computation (algorithmics)**

## But recalling some main messages to retain

**The first is that even without considering the problem of noise in the data (see later), it's important to remember to do "good" mathematics/statistics/ computation (algorithmics)**

**Even when this has been done, and even if the data was not noisy, one must be careful with biological interpretation:**

- **It may be wrong for various reasons**
- **It may be not informative even when it is correct**

## But recalling some main messages to retain

**The first is that even without considering the problem of noise in the data (see later), it's important to remember to do "good" mathematics/statistics/ computation (algorithmics)**

**Even when this has been done, and even if the data was not noisy, one must be careful with biological interpretation:**

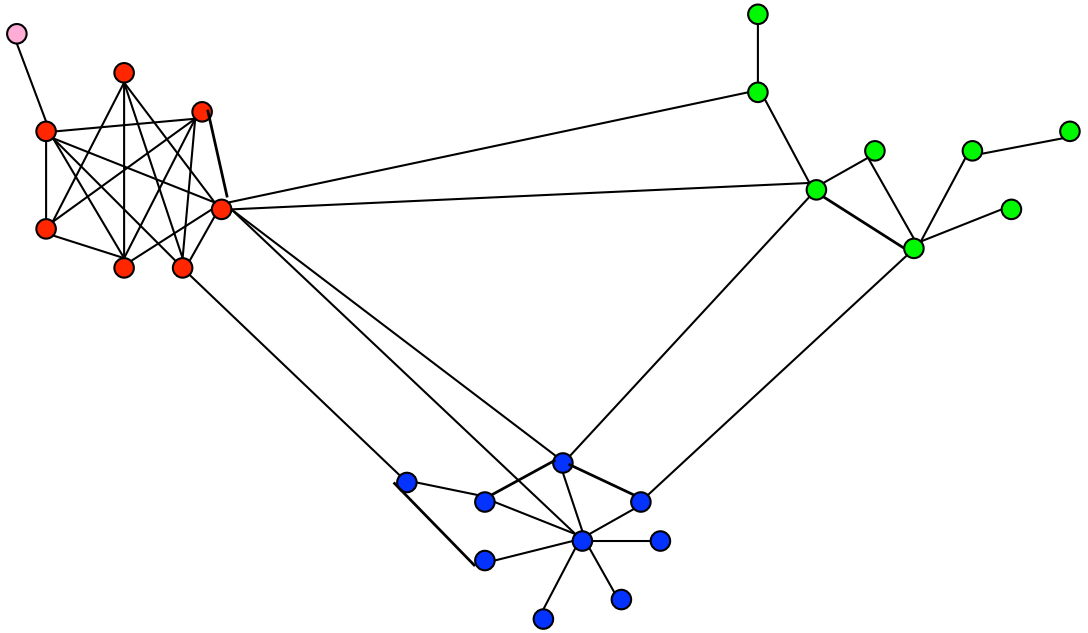
- **It may be wrong for various reasons**
- **It may be not informative even when it is correct**

**These indices intervene also in another, difficult context: the one of obtaining "good" random models against which to compare biological networks in order to then draw some reliable biological conclusions**

## Somewhat related but different from motifs: Enumerating “modules” (notice the “inverted commas”)

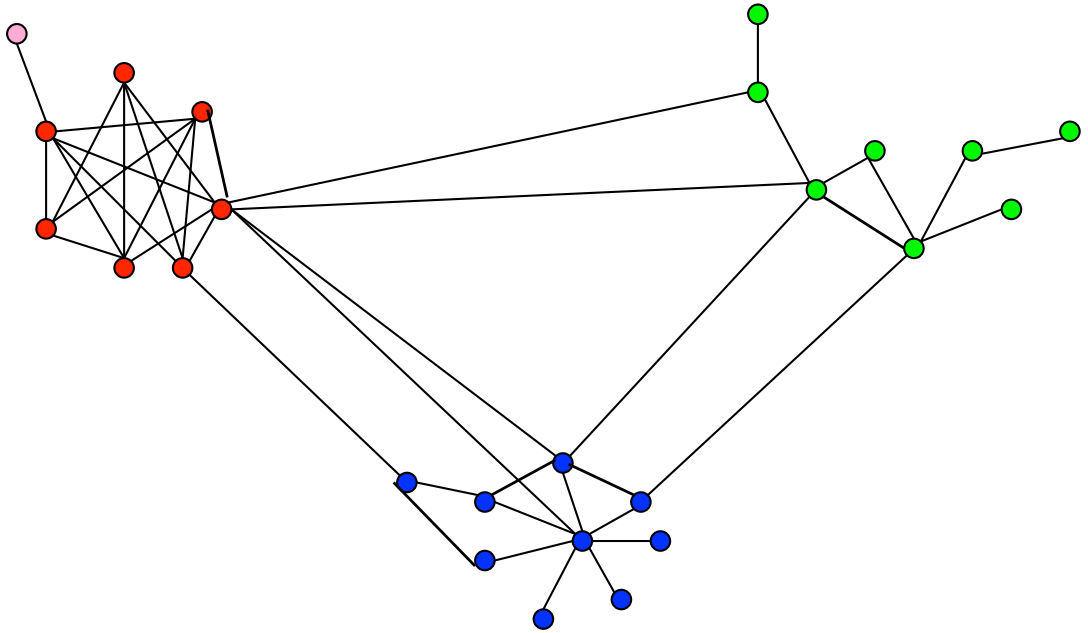
**One example of definition:**

**Subgraph *S* is a module if $M(S)=ind(S)/outd(S) > 1$**



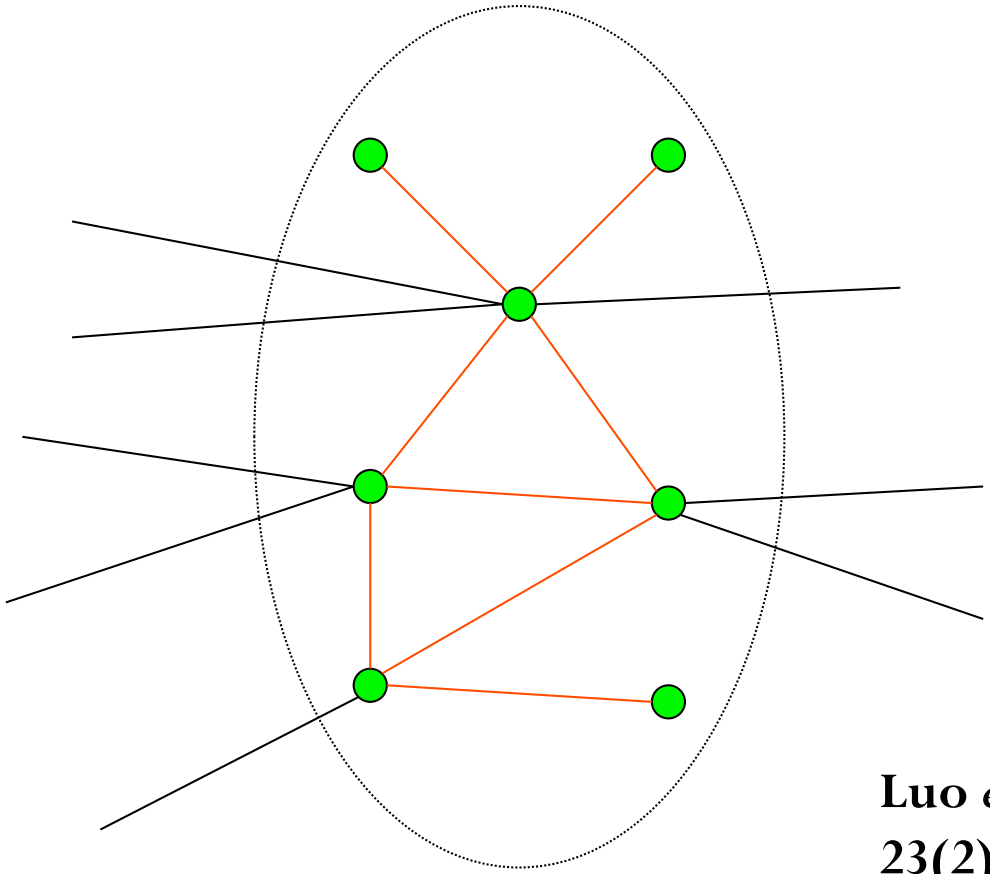
## Somewhat related but different from motifs: Enumerating “modules” (notice the “inverted commas”)

**One example of definition:**
**Subgraph *S* is a strong module if $M(S)=ind(S)/outd(S) > 1$ and the same is true for every node in *S***



## Somewhat related but different from motifs: Enumerating "modules"

**Is this module strong?**



Luo *et al*. *Bioinformatics*, 23(2): 207-214, 2007

**There are (many) other definitions of modules that have been used**
**Here is an interesting one based on fluxes**

---

**This is for metabolic networks**
**Informally: set of reactions that behave together like one reaction with a fixed flux**

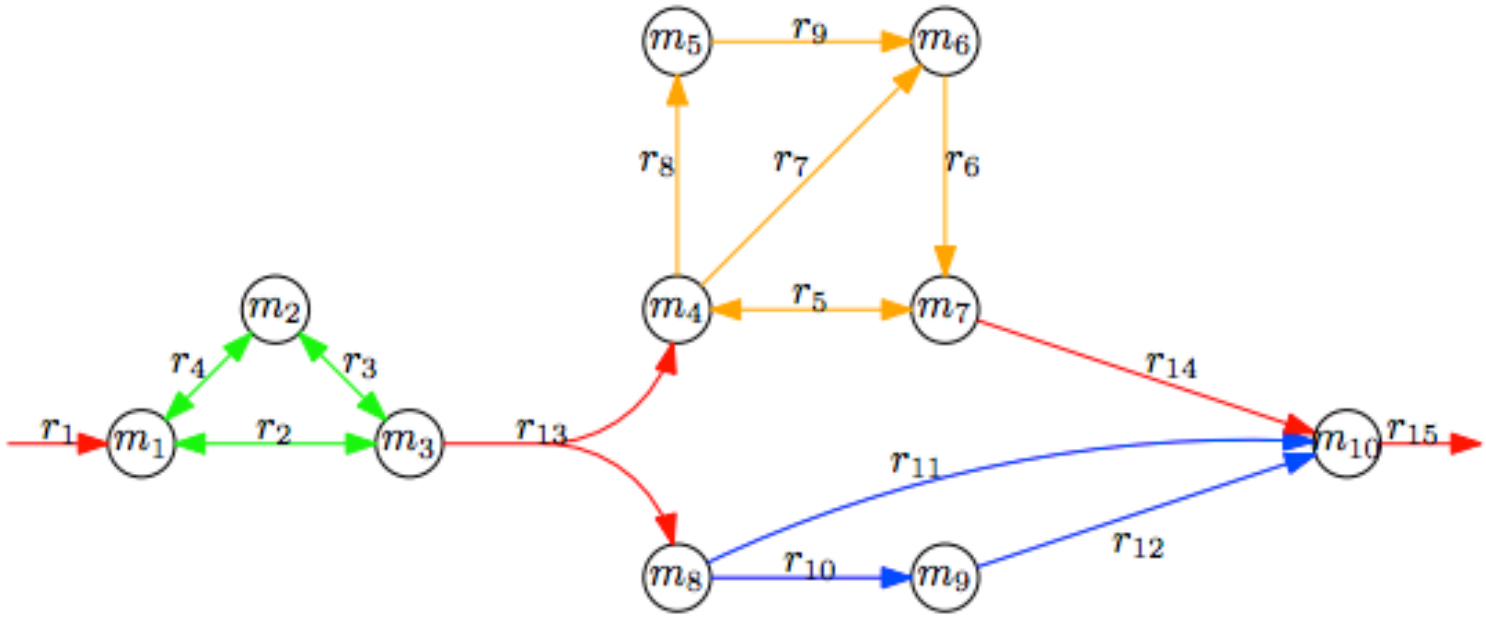



Figure 1: All stoichiometric coefficients in this example are 1. Assume flux through reaction $r_1$ is fixed to 1. Then flux through reactions $(r_1, r_{13}, r_{14}, r_{15})$ is fixed and we get the three modules $(r_2, r_3, r_4)$, $(r_5, r_6, r_7, r_8, r_9)$, and $(r_{10}, r_{11}, r_{12})$.

## Somewhat related again to subgraph identification

**This is for metabolic networks**

**Informally: set of reactions that behave together like one reaction with a fixed flux**

**Reimers *et al. J Comput Biol.* 22(5):414-424, 2015**

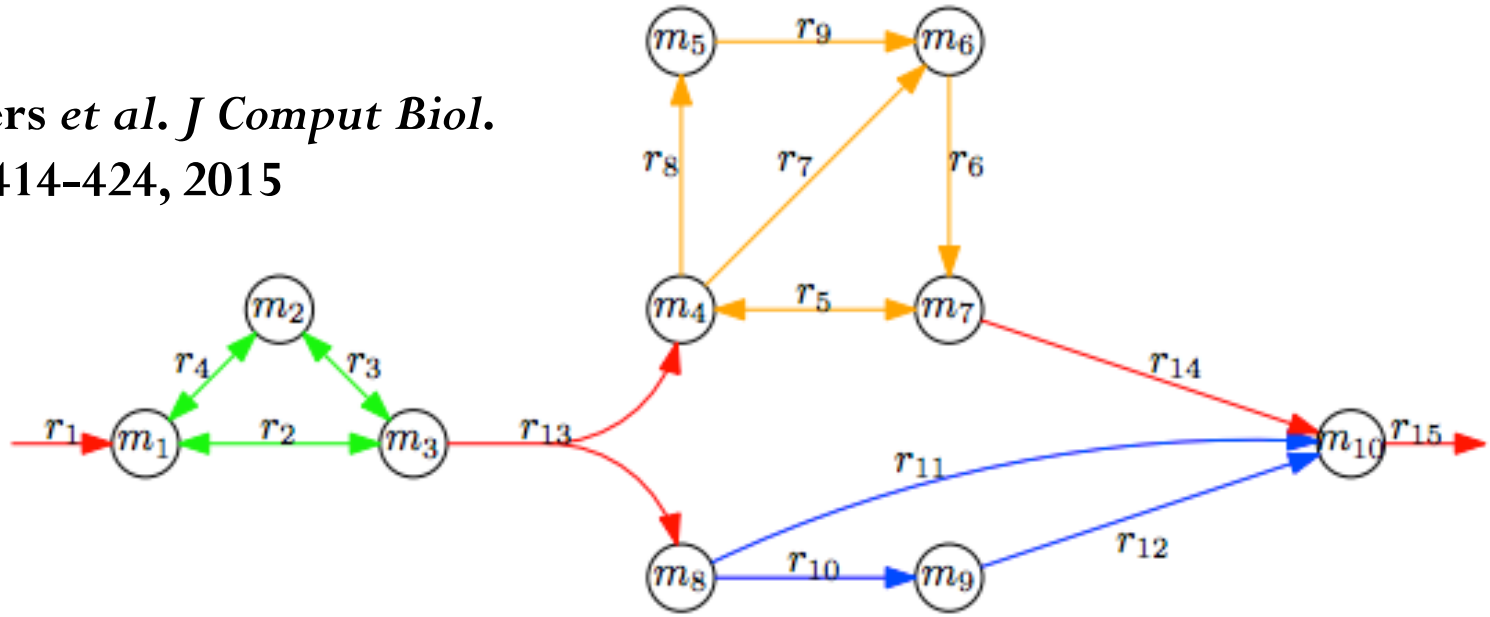
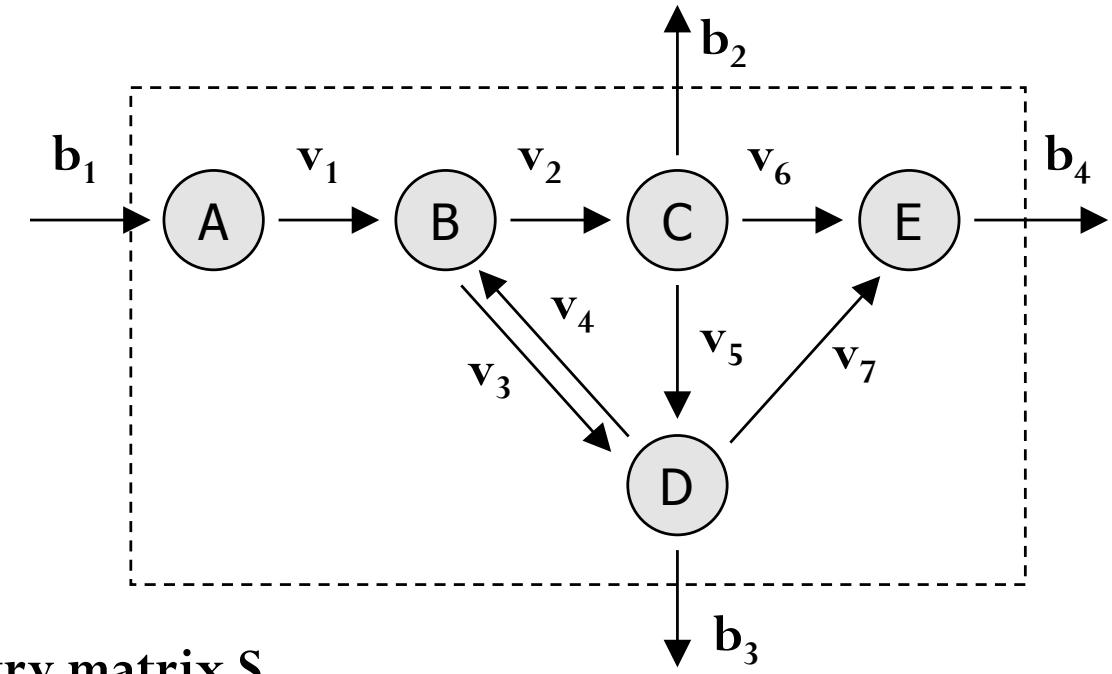



Figure 1: All stoichiometric coefficients in this example are 1. Assume flux through reaction $r_1$ is fixed to 1. Then flux through reactions $(r_1, r_{13}, r_{14}, r_{15})$ is fixed and we get the three modules $(r_2, r_3, r_4)$, $(r_5, r_6, r_7, r_8, r_9)$, and $(r_{10}, r_{11}, r_{12})$.

# Flux modes



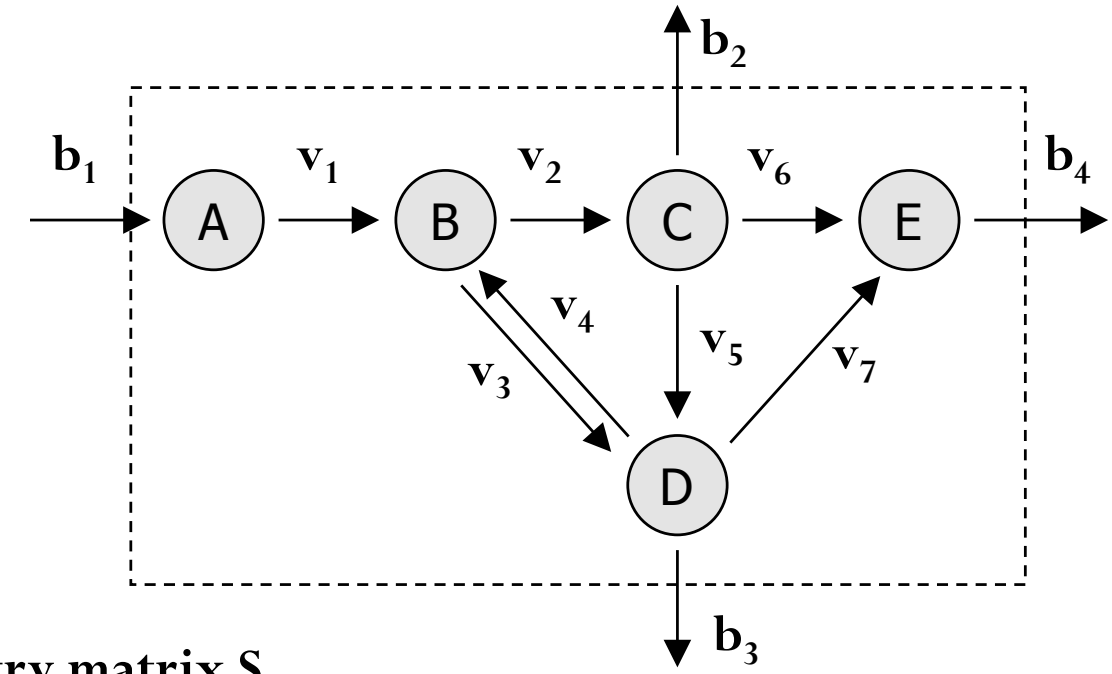

## Stoichiometry matrix S

| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ | $v_6$ | $v_7$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | -1 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 |
| B | 1 | -1 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C | 0 | 1 | 0 | 0 | -1 | -1 | 0 | 0 | -1 | 0 | 0 |
| D | 0 | 0 | 1 | -1 | 1 | 0 | -1 | 0 | 0 | -1 | 0 |
| E | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | -1 |

$S.v = 0$

$v_i \geq 0$ **for all** $i \in \{irrev\}$

# Elementary flux modes




## Stoichiometry matrix S

| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ | $v_6$ | $v_7$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | -1 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 |
| B | 1 | -1 | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C | 0 | 1 | 0 | 0 | -1 | -1 | 0 | 0 | -1 | 0 | 0 |
| D | 0 | 0 | 1 | -1 | 1 | 0 | -1 | 0 | 0 | -1 | 0 |
| E | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | -1 |

$S.v = 0$

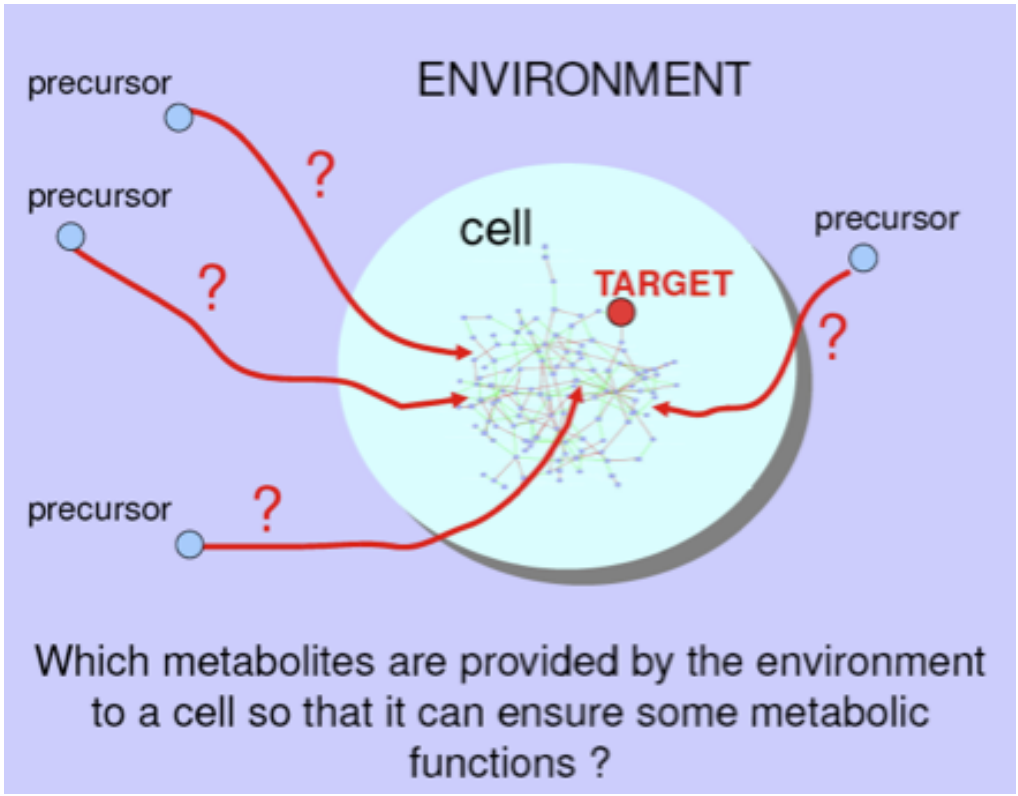
$v_i \geq 0$ **for all** $i \in$ {**irrev**}

**no vector** *w* **such that:**

$support(w) \subset support(v)$

# Minimal precursor sets

## Biological motivation




**$X$ precursor set of $T$ with $Z$ iff:**

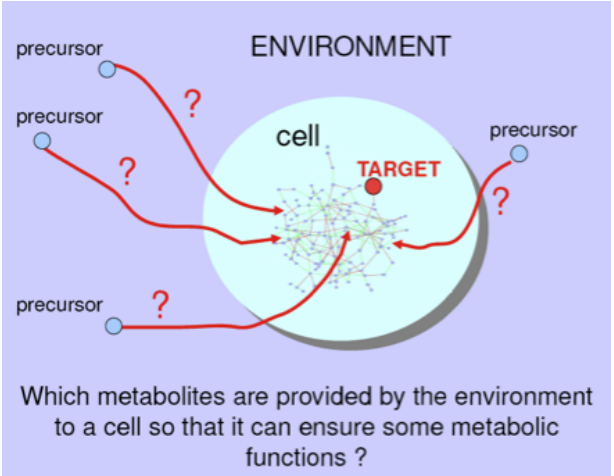
$$Scope_Z(X) \subseteq Z \cup T$$

**(plus stoichiometry)**

**May lead to another type of biological networks…**

# Minimal precursor sets

## More on precursor sets




## And on some other related topics with me later in the year

## Environment may be other species
## Species interactions, including “symbiosis”

### Main “symbiotic” relations

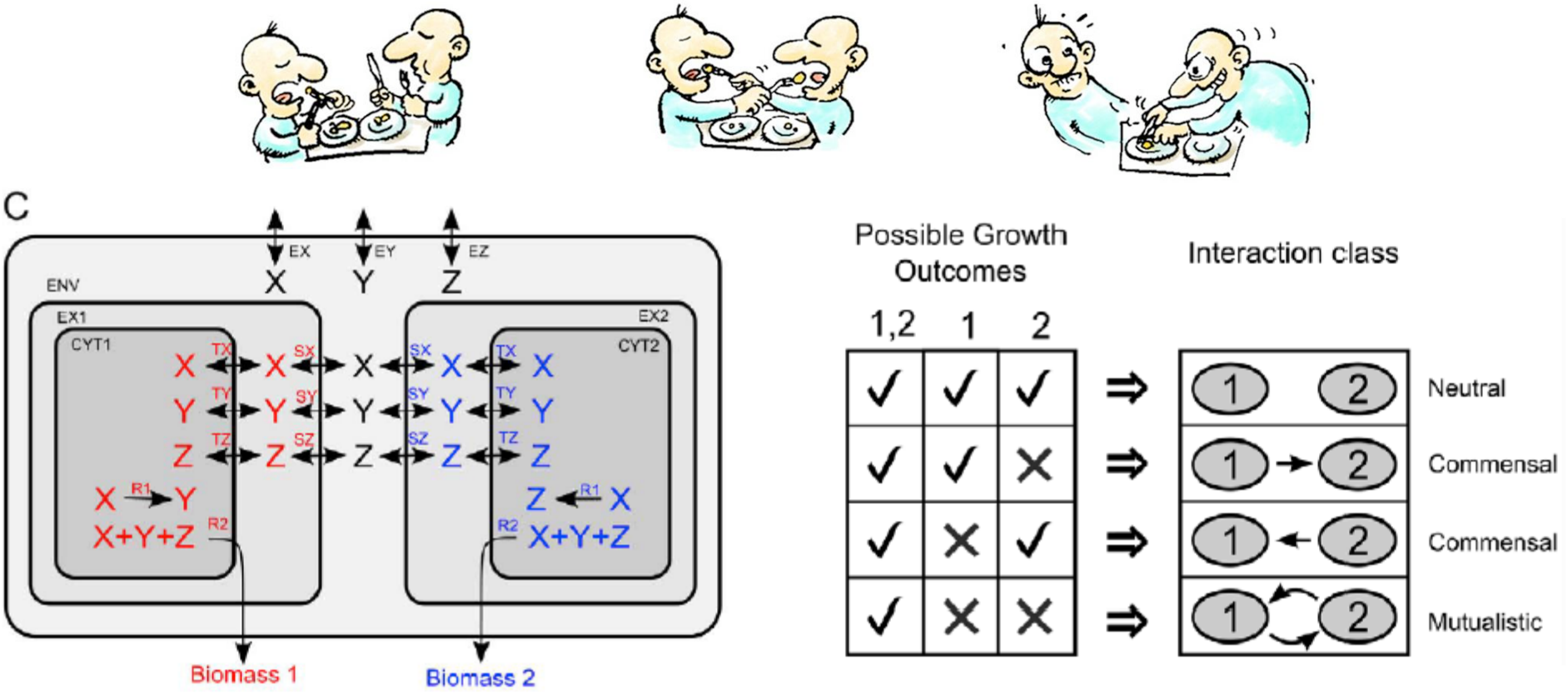
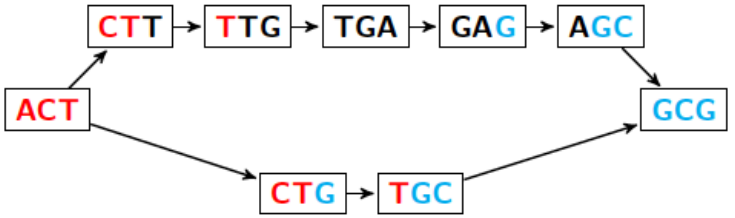


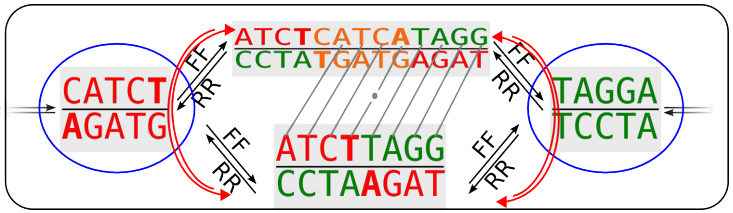

Figure: Klitgord, Segrè, PLOS, 2010

# Another aspect of symbiosis
# Example of graphs as tools

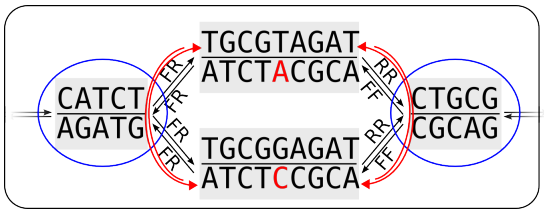
**Next Generation Sequencing (NGS), especially in the context of no reference genome**

**Vertex-disjoint *st*-paths in de Bruijn (di)graphs**

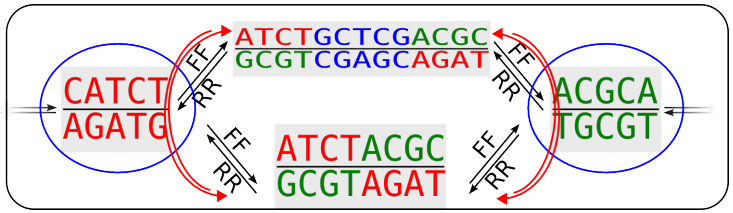



**Repeats: 1 path of length at most 2k-2, the two paths align**



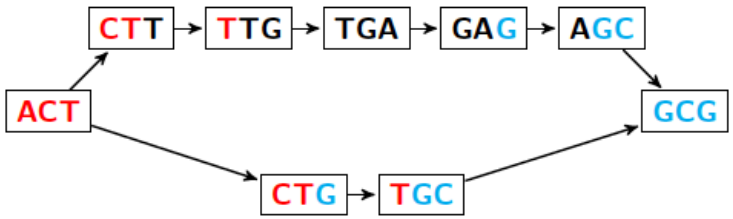

**Single Nucleotide Polymorphism (SNP): 2 paths of length 2k-1**




**Alternative Splicing (AS): 1 path of length <= 2k-2**

# Graphs as tools – NGS data analysis

**Vertex-disjoint *st*-paths in de Bruijn (di)graphs**




**And other related topics with Blerina Sinaimeri**

## Another type of approach of "species interactions"
## Game theory

$$A + X \rightarrow bio_1$$
$$C + X \rightarrow bio_1 + D$$

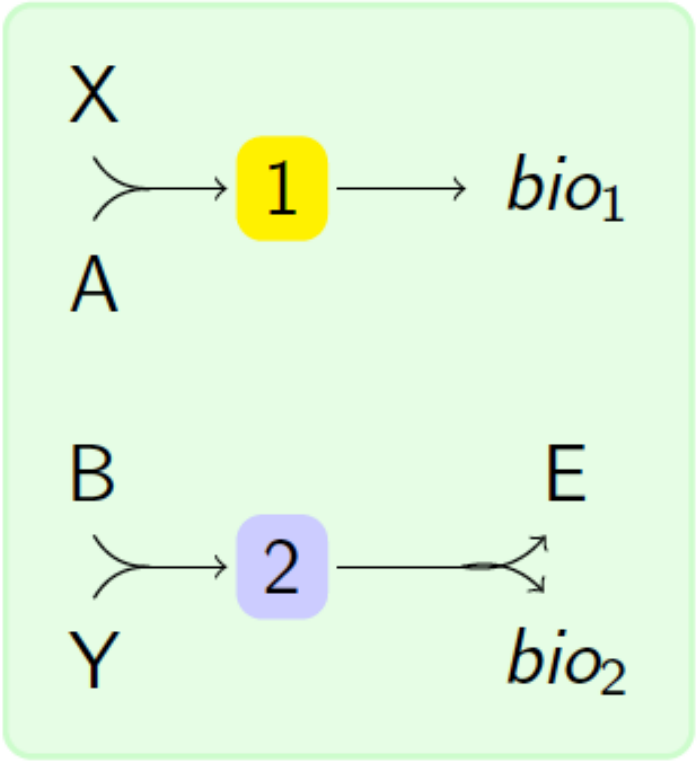
$$B + Y \rightarrow bio_2 + E$$
$$D + Y \rightarrow bio_2 + C$$

$$A + X \rightarrow bio_3 + D$$
$$E \rightarrow bio_3$$

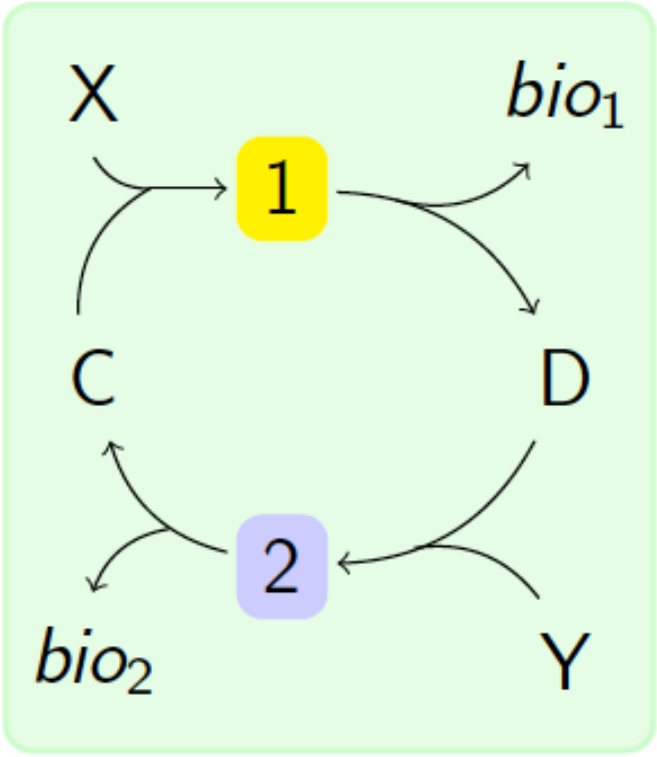
## Change of available sources

$A + X \rightarrow bio_1$
$C + X \rightarrow bio_1 + D$

$B + Y \rightarrow bio_2 + E$
$D + Y \rightarrow bio_2 + C$




A, B x

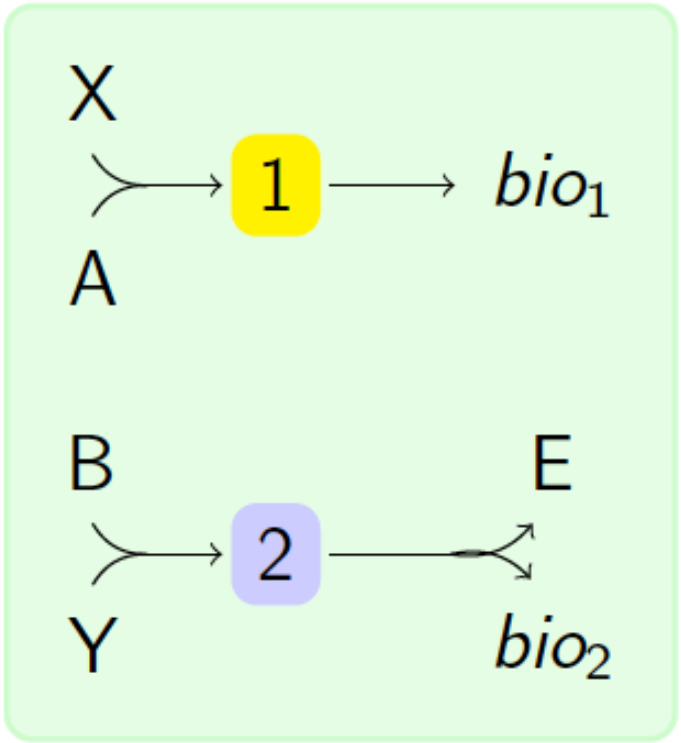


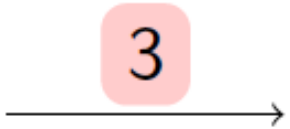
## Change of player composition

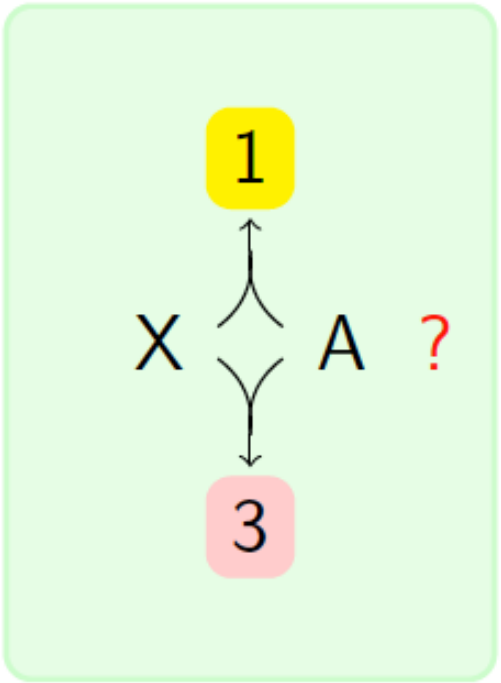
$A + X \rightarrow bio_1$
$C + X \rightarrow bio_1 + D$

$B + Y \rightarrow bio_2 + E$
$D + Y \rightarrow bio_2 + C$

$A + X \rightarrow bio_3 + D$
$E \rightarrow bio_3$



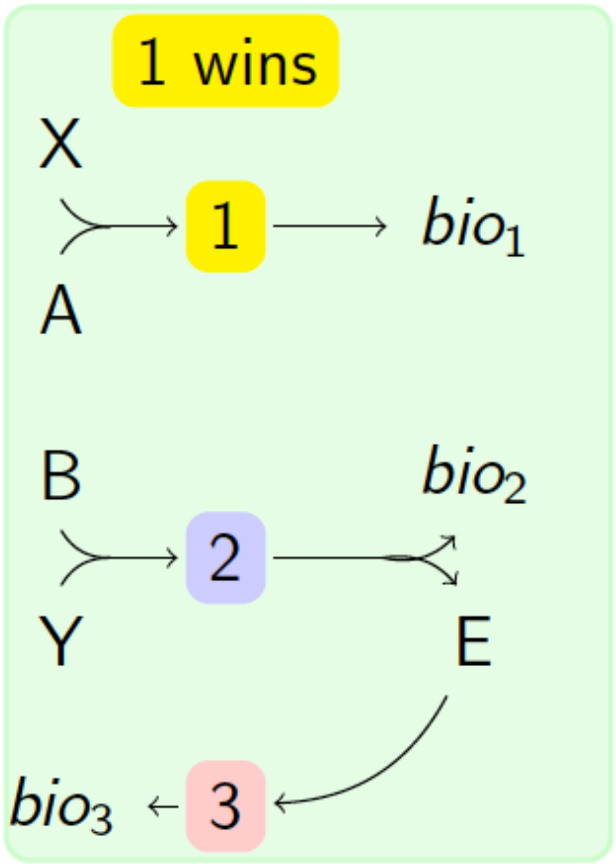



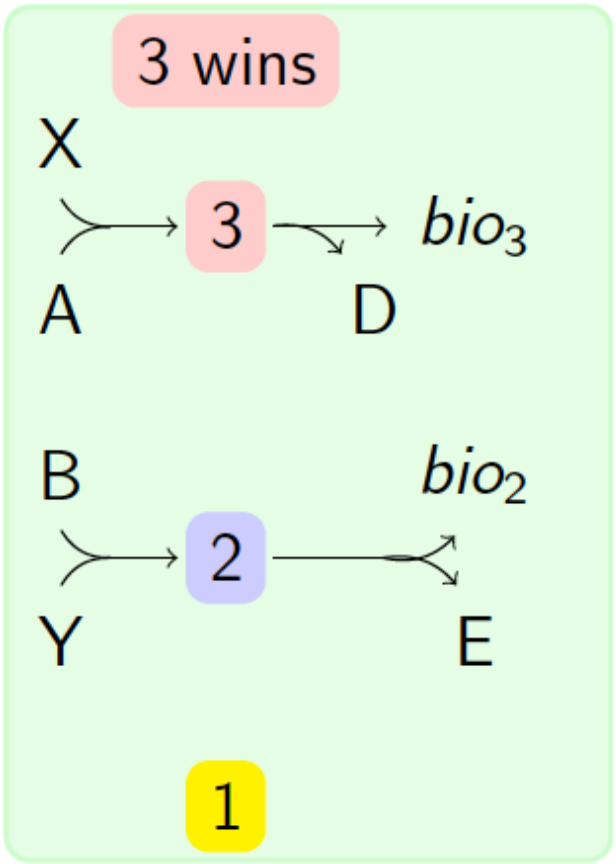
## Change of player composition

$A + X \rightarrow bio_1$
$C + X \rightarrow bio_1 + D$

$B + Y \rightarrow bio_2 + E$
$D + Y \rightarrow bio_2 + C$

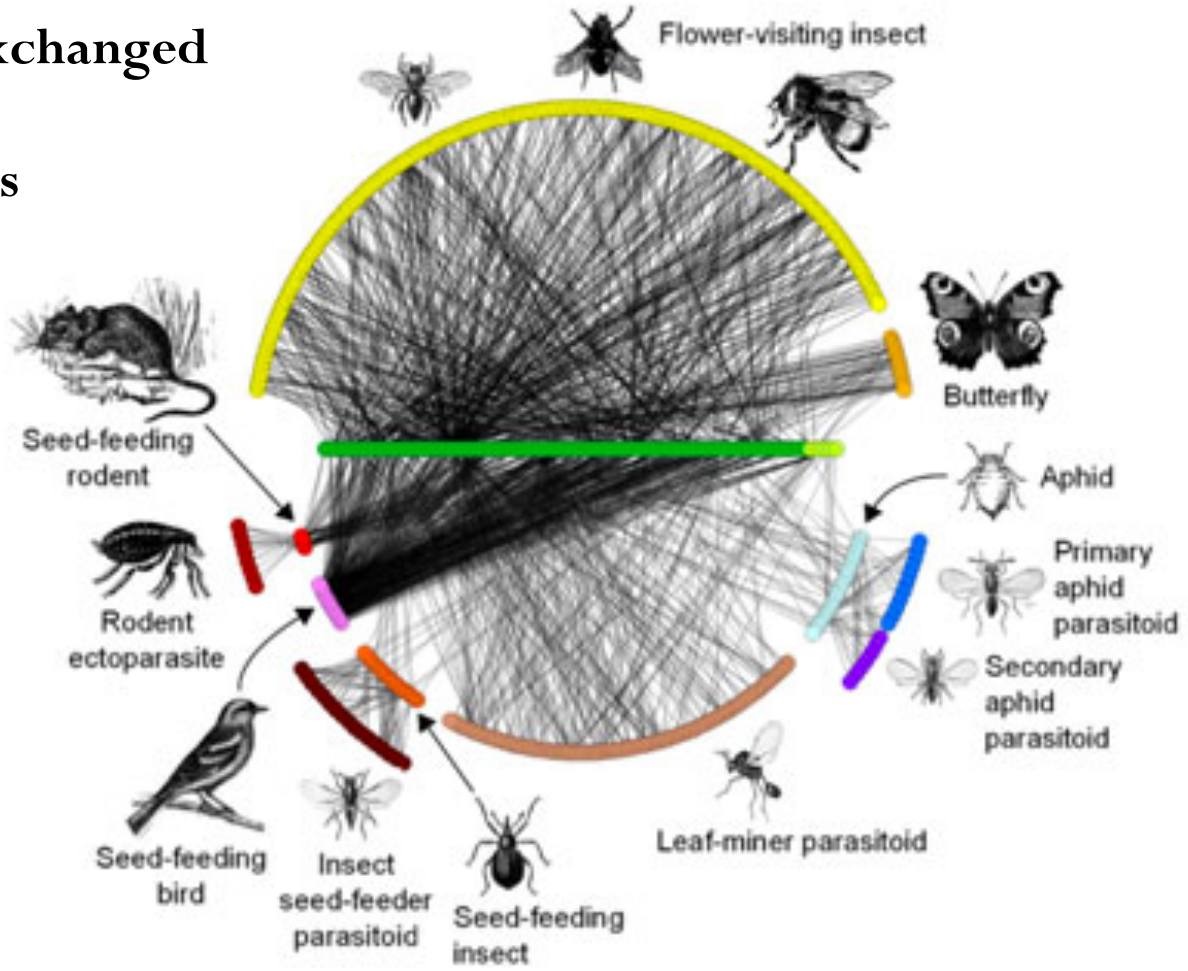
$A + X \rightarrow bio_3 + D$
$E \rightarrow bio_3$



# Molecular ecological networks

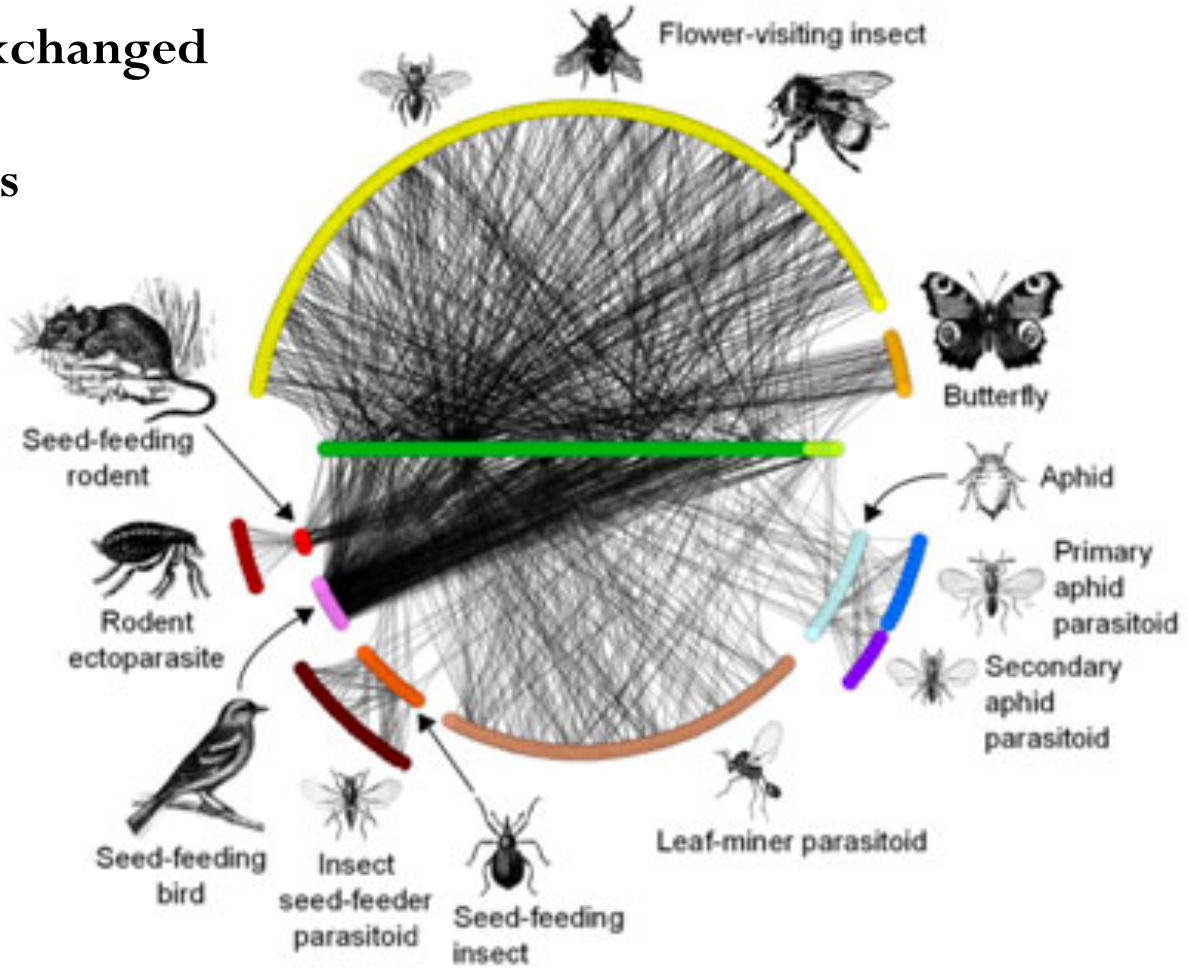
**Not only metabolites exchanged but also possibly other (macro)molecules play a role in the interaction**



# Molecular ecological networks

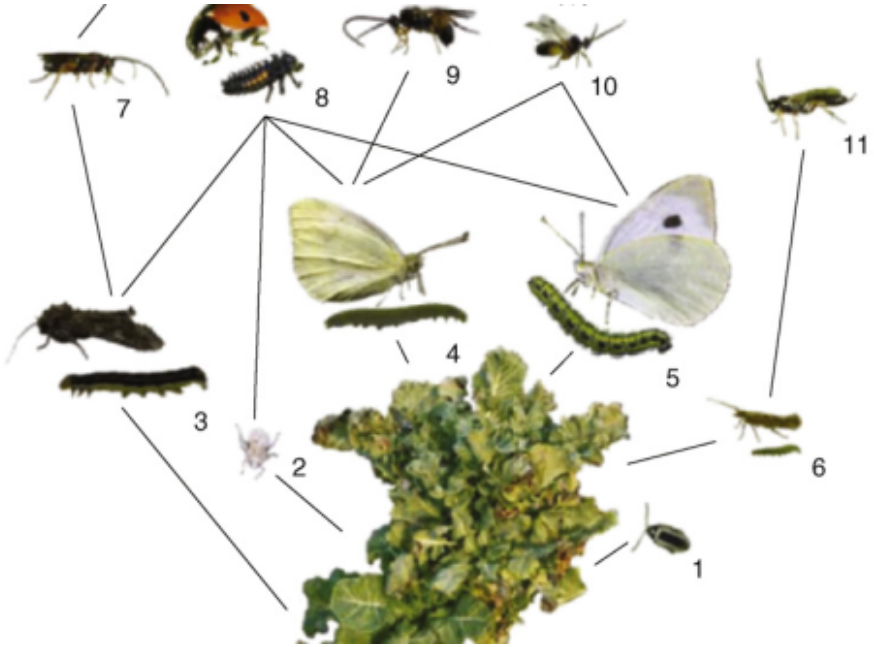
**Not only metabolites exchanged but also possibly other (macro)molecules play a role in the interaction**

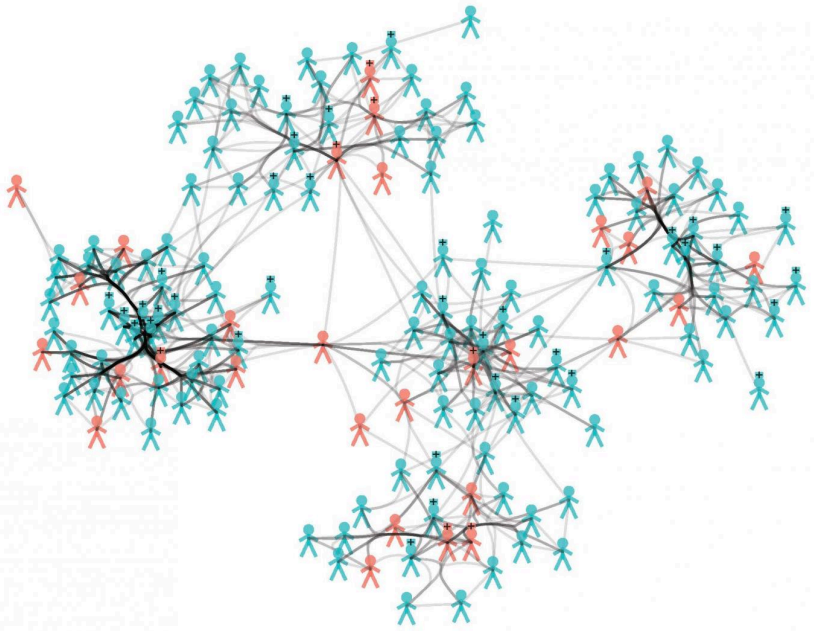
**Plus environment**



## Leads to more general ecological networks as well as to “infection” networks

**The dynamic aspect of such networks is important (but not same kind of “dynamics” as mentioned earlier!)**



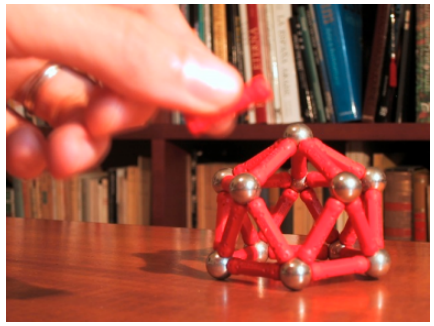



# Dynamic graph algorithms

**Some typical updates:**

- **insert**
- **delete**
- **change weight**

**But there are many others which might be relevant!**

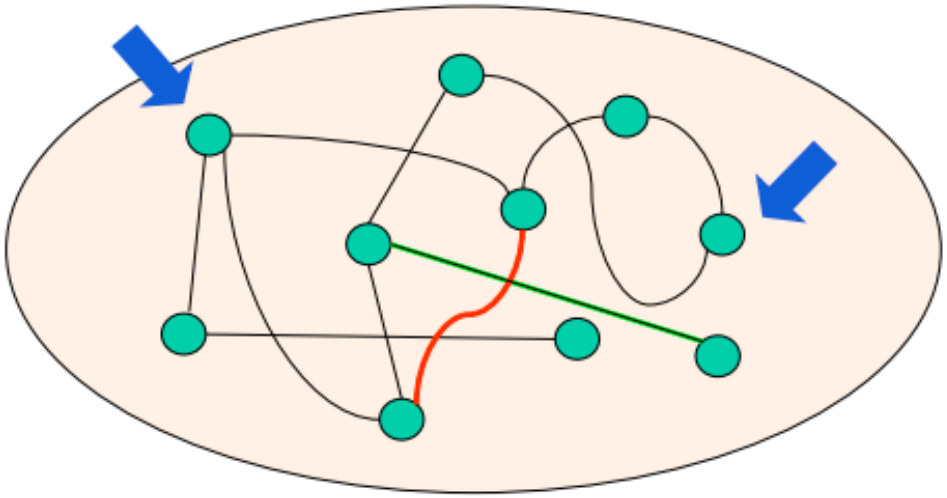


Initialize

Insert

Delete

Query

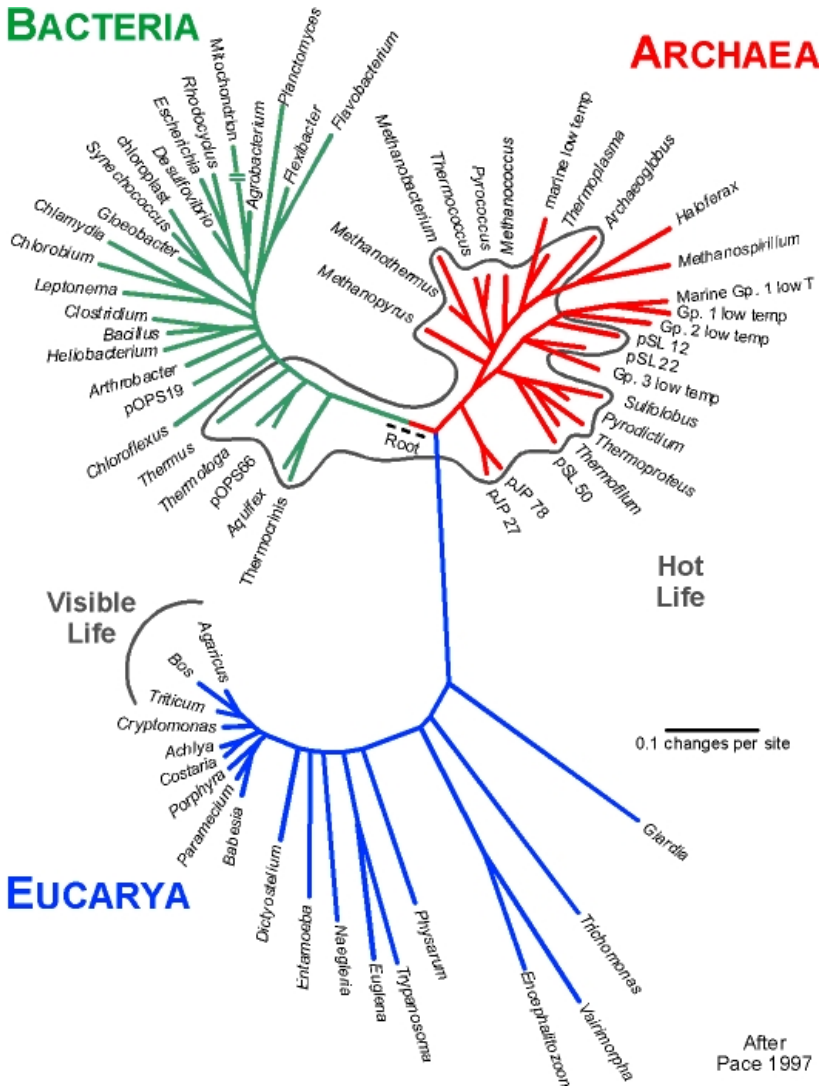


A graph

# This was “fast dynamics”
# Things can however change much more slowly leading to Evolutionary networks

---

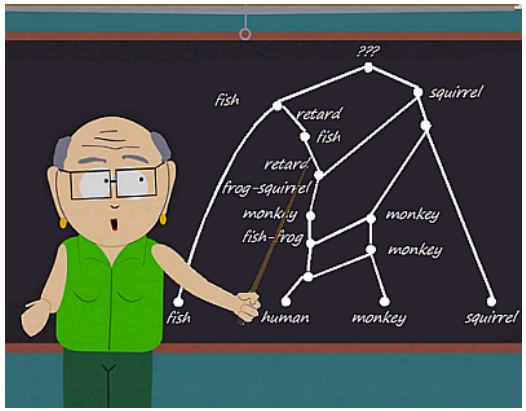
## Phylogenetic trees as a way to study evolution

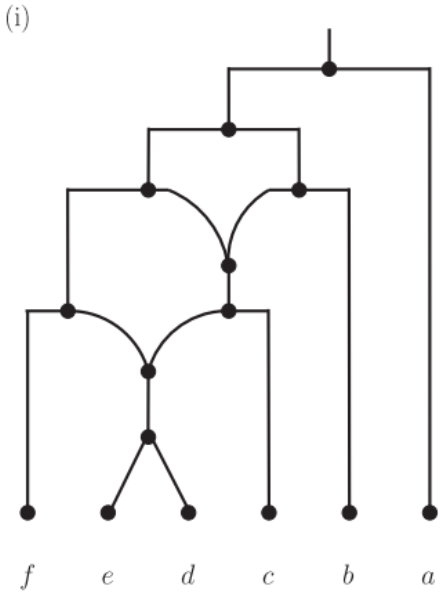


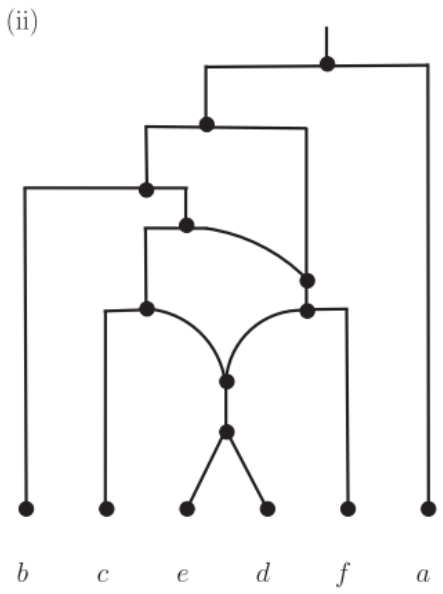
# From phylogenetic trees to networks

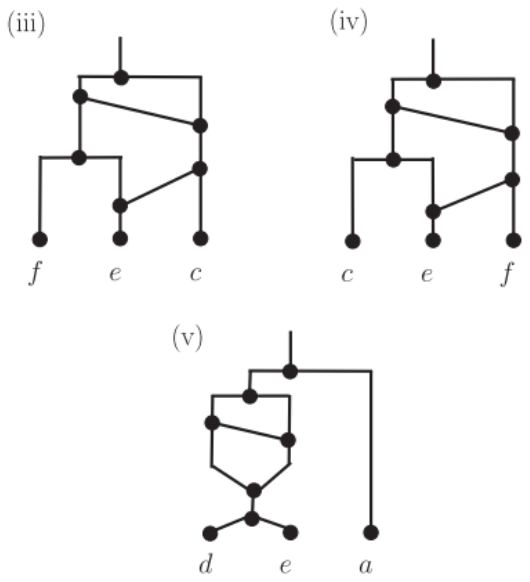
**Two main reasons:**

- **Contradictory relationships**
- **Reticulation**
  - **Hybridisation**
  - **Recombination**



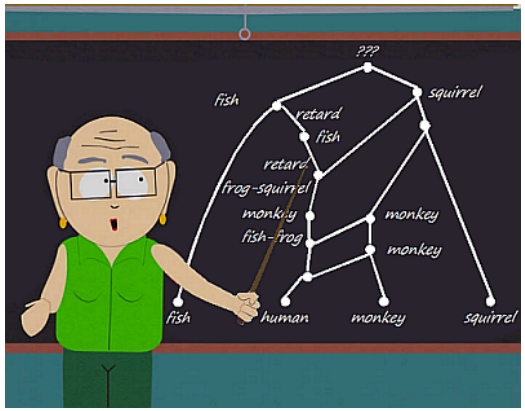



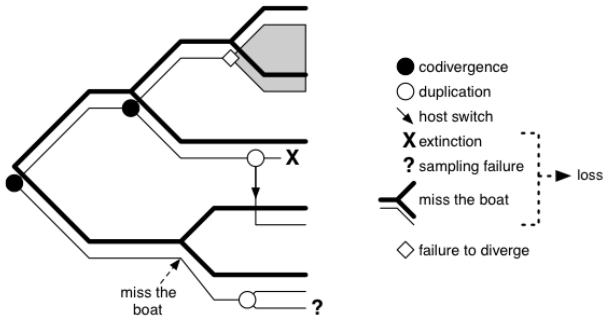



# Phylogenetic networks

## More on phylogenetic networks



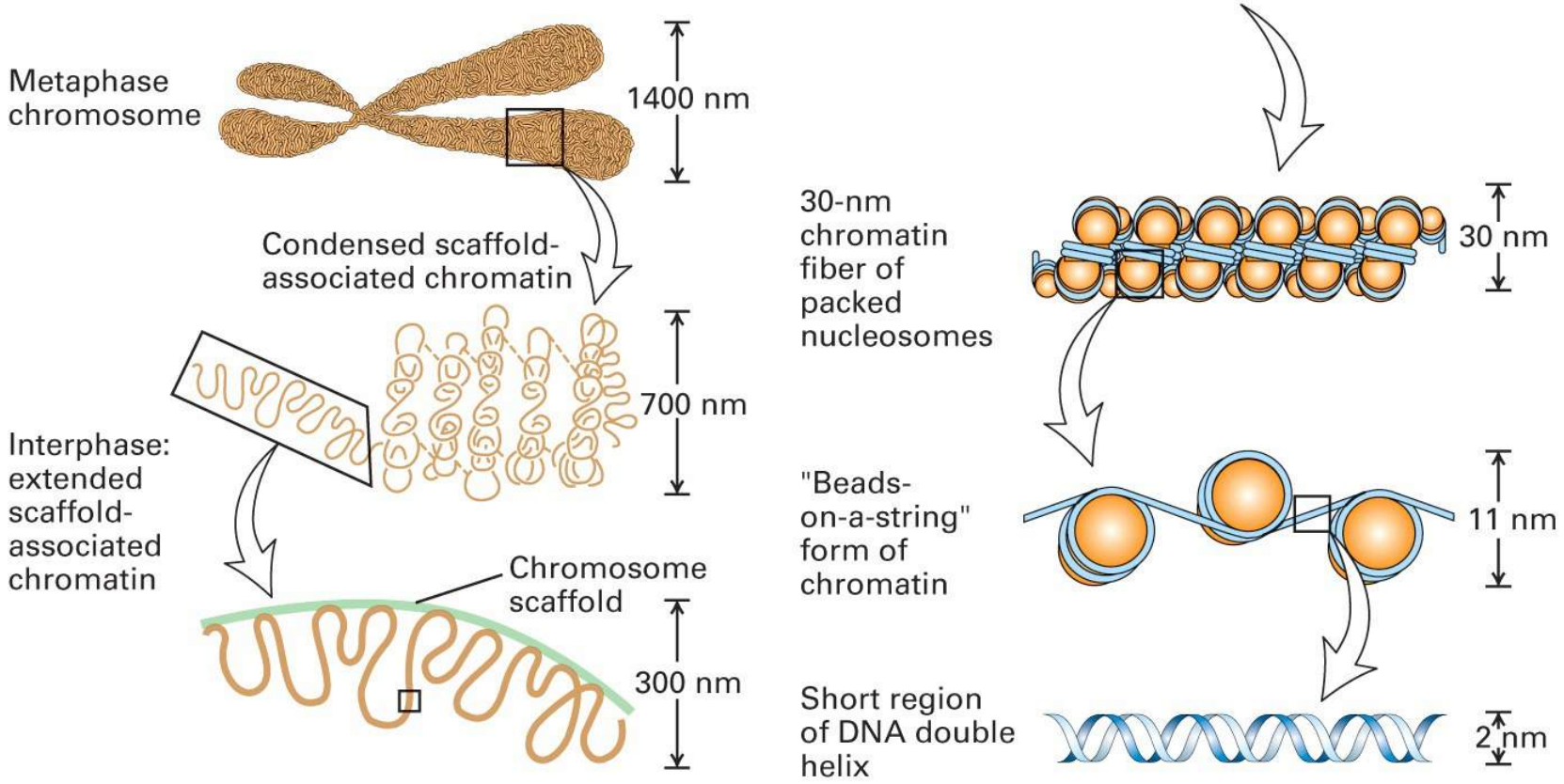

## And on co-phylogeny with Blerina Sinaimeri



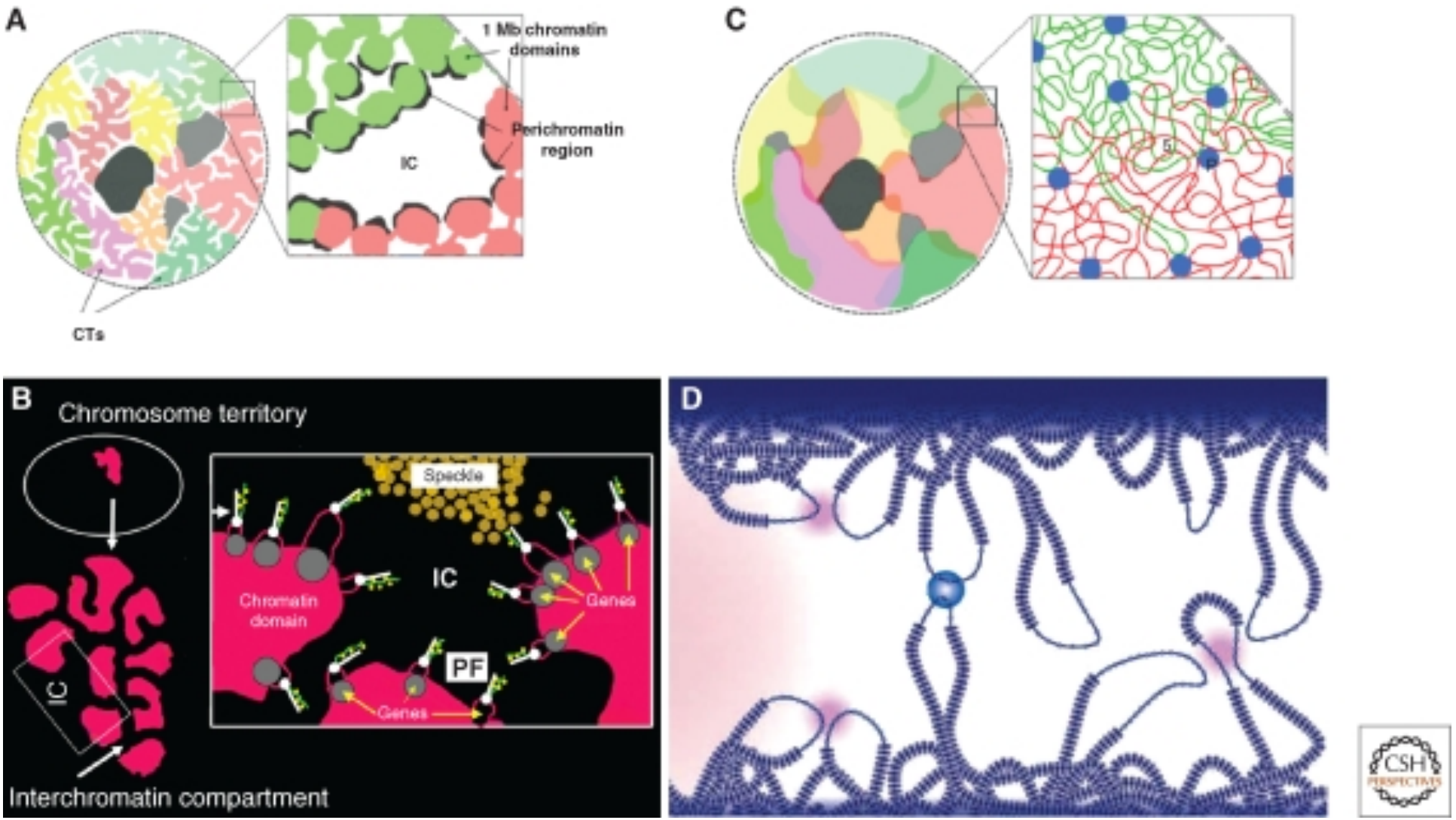

**Finally, remember the conclusion of my (brief) biological introduction:**
**Chromosomes are not spaghetti!**

---

**DNA in a living cell is in a highly compacted and structured**
**Transcription is dependent on such structural state – SEQUENCE alone does not tell the whole story!**



# The big question here is how to infer the network from some (noisy) interaction data



## A few references for those curious to know more

**Systems biology: Property of reconstructed networks, Bernhard Palsson**

**Systems biology: Simulation of dynamic network states, Bernhard Palsson**

**An introduction to systems biology:De`sign principles of biological circuits, Uri Alon**

**Algebraic statistics for molecular biology, Lior Pachter and Bernd Sturmfels**

**Hypergraphs and cellular networks, Steffen Klamt**

**Papers by Jörg Stelling**

**And many, many more**
**If interested in having more references, contact us!**