Network modularity, currency metabolites and graph representations of metabolism

Petter Holme

KTH, CSC, Computational Biology

April 2, 2008, Physics of distributed information systems

http://www.csc.kth.se/~pholme/

why?

Metabolism = is the set of all chemical reactions in an organism? Why study it by networks?

- Details (reaction coeficient, subcellular localisation, etc.) are not known. So to study large-scale structures, a detailed picture makes little sense.
- There are many methods to analyze graphs statistically.

why?

Metabolism = is the set of all chemical reactions in an organism? Why study it by networks?

- Details (reaction coeficient, subcellular localisation, etc.) are not known. So to study large-scale structures, a detailed picture makes little sense.
- There are many methods to analyze graphs statistically.

why?

Metabolism = is the set of all chemical reactions in an organism? Why study it by networks?

- Details (reaction coeficient, subcellular localisation, etc.) are not known. So to study large-scale structures, a detailed picture makes little sense.
- There are many methods to analyze graphs statistically.

complex network studies

. . . of biochemical networks. What questions can we ask?

- how can the large-scale organization be characterized?
- are there any universal features over different species?
- do the differences tell us something about evolution?
- can we identify functional modules?
- . . the functions of molecules?

complex network studies

. . . of biochemical networks. What questions can we ask?

- how can the large-scale organization be characterized?
- are there any universal features over different species?
- do the differences tell us something about evolution?
- can we identify functional modules?
- . . the functions of molecules?

complex network studies

. . . of biochemical networks. What questions can we ask?

- how can the large-scale organization be characterized?
- are there any universal features over different species?
- do the differences tell us something about evolution?
- can we identify functional modules?
- . . the functions of molecules?

complex network studies

. . . of biochemical networks. What questions can we ask?

- how can the large-scale organization be characterized?
- are there any universal features over different species?
- do the differences tell us something about evolution?
- can we identify functional modules?
- . . the functions of molecules?

complex network studies

. . . of biochemical networks. What questions can we ask?

- how can the large-scale organization be characterized?
- are there any universal features over different species?
- do the differences tell us something about evolution?
- can we identify functional modules?
- . . the functions of molecules?

complex network studies

. . . of biochemical networks. What questions can we ask?

- how can the large-scale organization be characterized?
- are there any universal features over different species?
- do the differences tell us something about evolution?
- can we identify functional modules?
- . . the functions of molecules?

wallpapers

Network modularity, currency metabolites and graph representations of metabolism

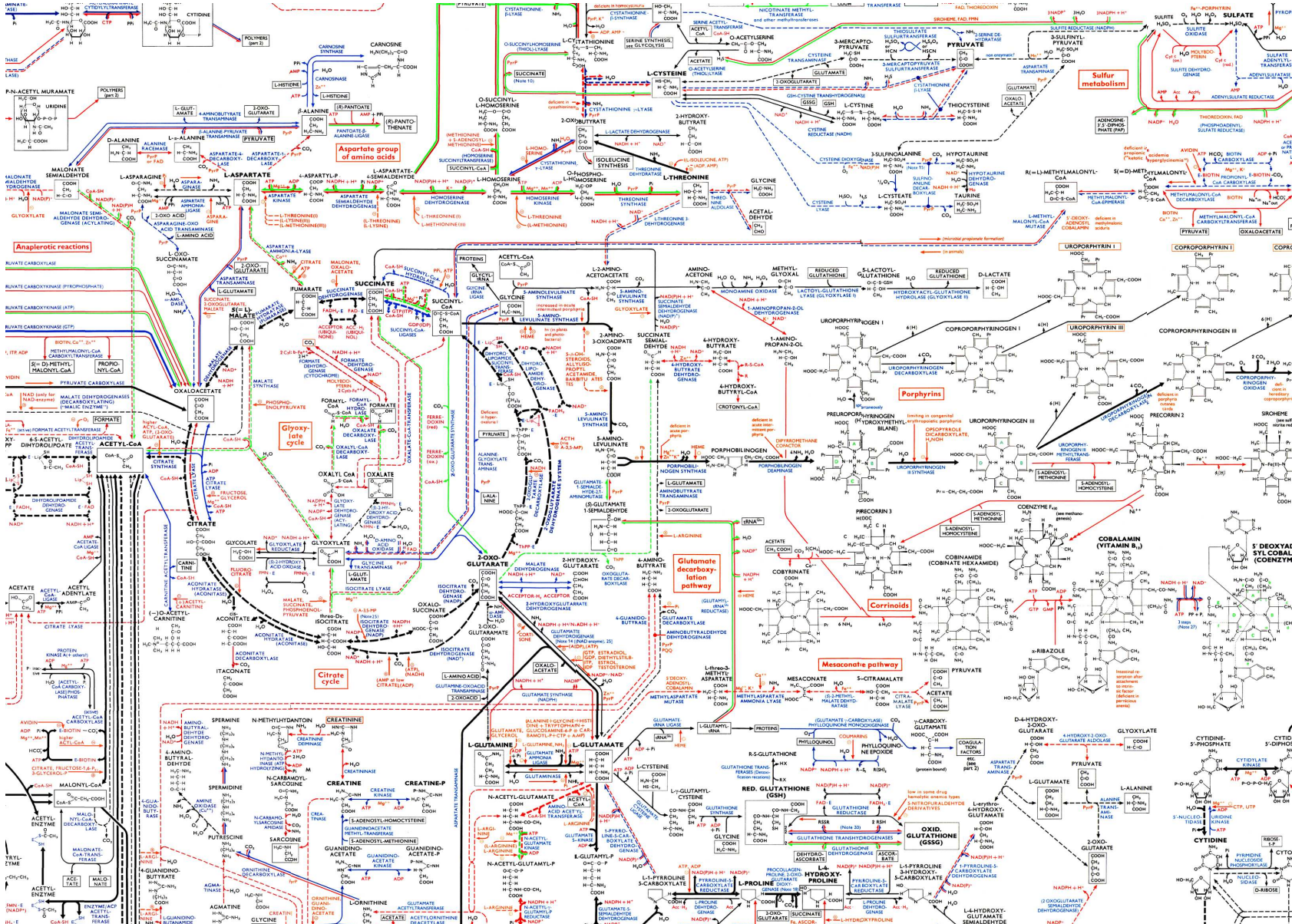
PETTER HOLME

metabolic networks, intro

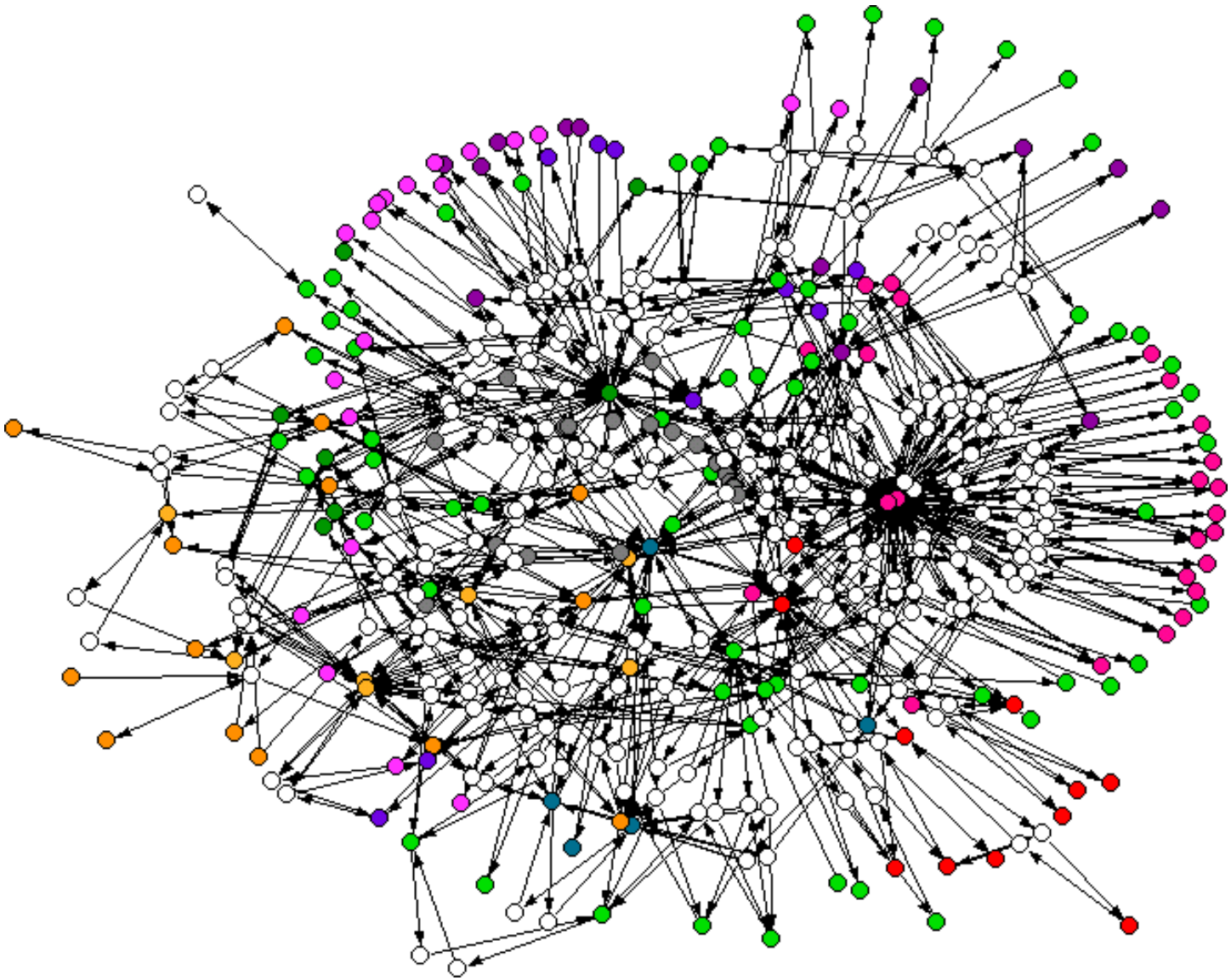
networks (physicist style?)

modularity & currency metabolites

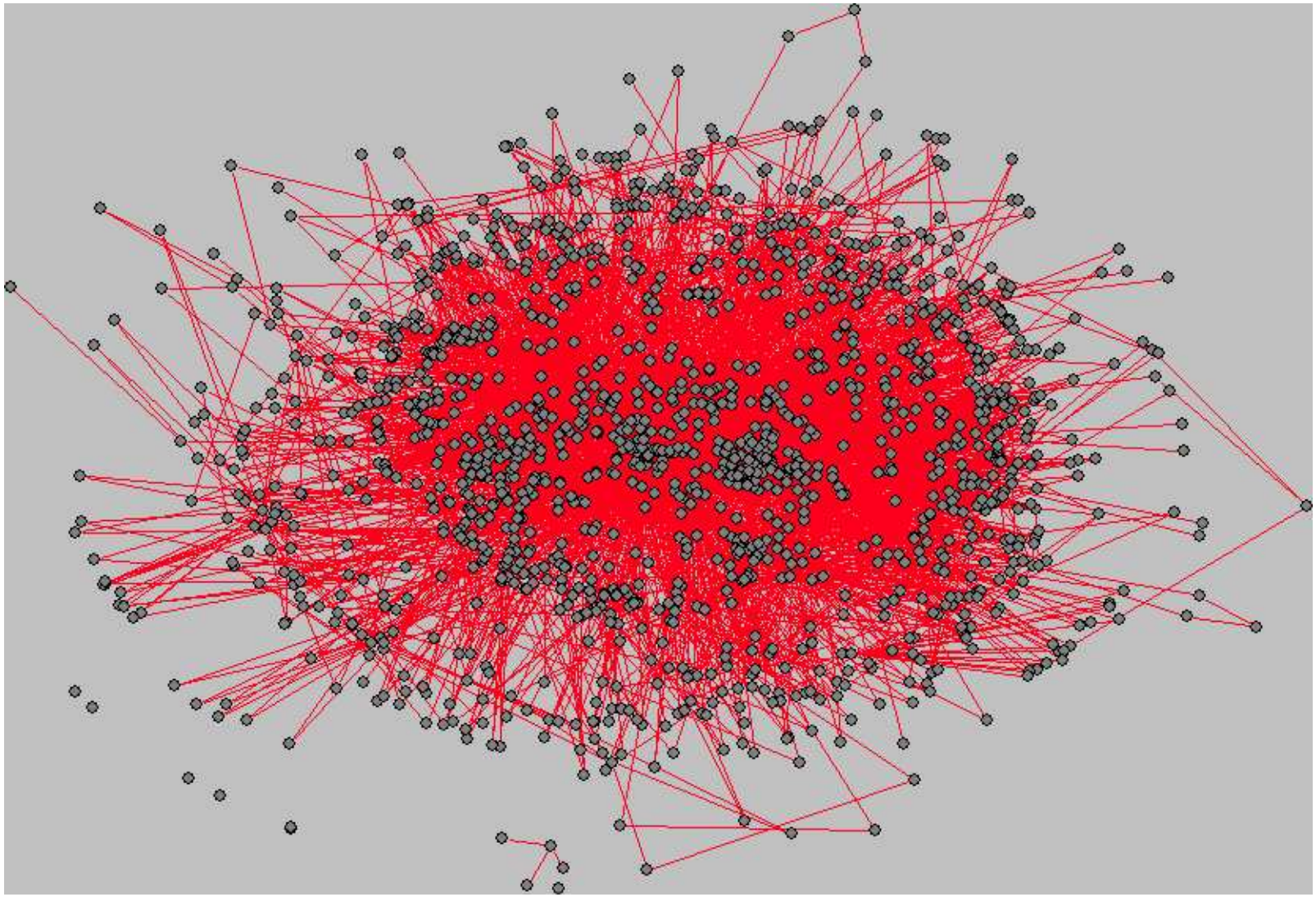
subnetwork hierarchies



B. burgdorferi, 2000



human, 2006



what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise **MAYBE?**
- the network is relatively sparse (the average degree is constant) **ALMOST TRUE**
- there is a dynamic system on the network **TRUE!**
- the time scale of this dynamics is faster than the dynamics of network evolution **TRUE!**

what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise MAYBE?
- the network is relatively sparse (the average degree is constant) ALMOST TRUE
- there is a dynamic system on the network TRUE!
- the time scale of this dynamics is faster than the dynamics of network evolution TRUE!

what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise MAYBE?
- the network is relatively sparse (the average degree is constant) ALMOST TRUE
- there is a dynamic system on the network TRUE!
- the time scale of this dynamics is faster than the dynamics of network evolution TRUE!

what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise MAYBE?
- the network is relatively sparse (the average degree is constant) ALMOST TRUE
- there is a dynamic system on the network TRUE!
- the time scale of this dynamics is faster than the dynamics of network evolution TRUE!

what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise MAYBE?
- the network is relatively sparse (the average degree is constant) ALMOST TRUE
- there is a dynamic system on the network TRUE!
- the time scale of this dynamics is faster than the dynamics of network evolution TRUE!

what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise MAYBE?
- the network is relatively sparse (the average degree is constant) ALMOST TRUE
- there is a dynamic system on the network TRUE!
- the time scale of this dynamics is faster than the dynamics of network evolution **TRUE!**

what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise MAYBE?
- the network is relatively sparse (the average degree is constant) ALMOST TRUE
- there is a dynamic system on the network **TRUE!**
- the time scale of this dynamics is faster than the dynamics of network evolution **TRUE!**

what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise **MAYBE?**
- the network is relatively sparse (the average degree is constant) **ALMOST TRUE**
- there is a dynamic system on the network **TRUE!**
- the time scale of this dynamics is faster than the dynamics of network evolution **TRUE!**

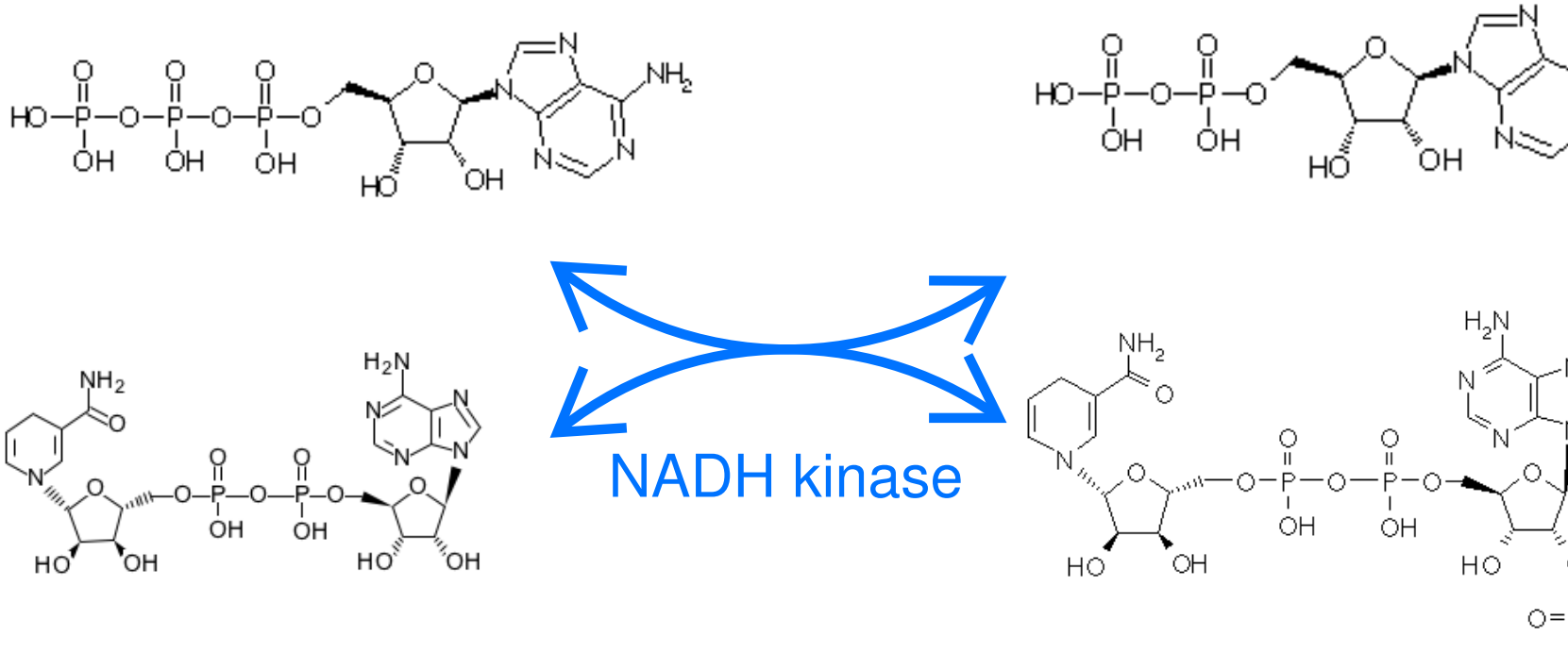
what systems?

What systems can be analyzed with complex network methods?

- items are, naturally, coupled pairwise **MAYBE?**
- the network is relatively sparse (the average degree is constant) **ALMOST TRUE**
- there is a dynamic system on the network **TRUE!**
- the time scale of this dynamics is faster than the dynamics of network evolution **TRUE!**

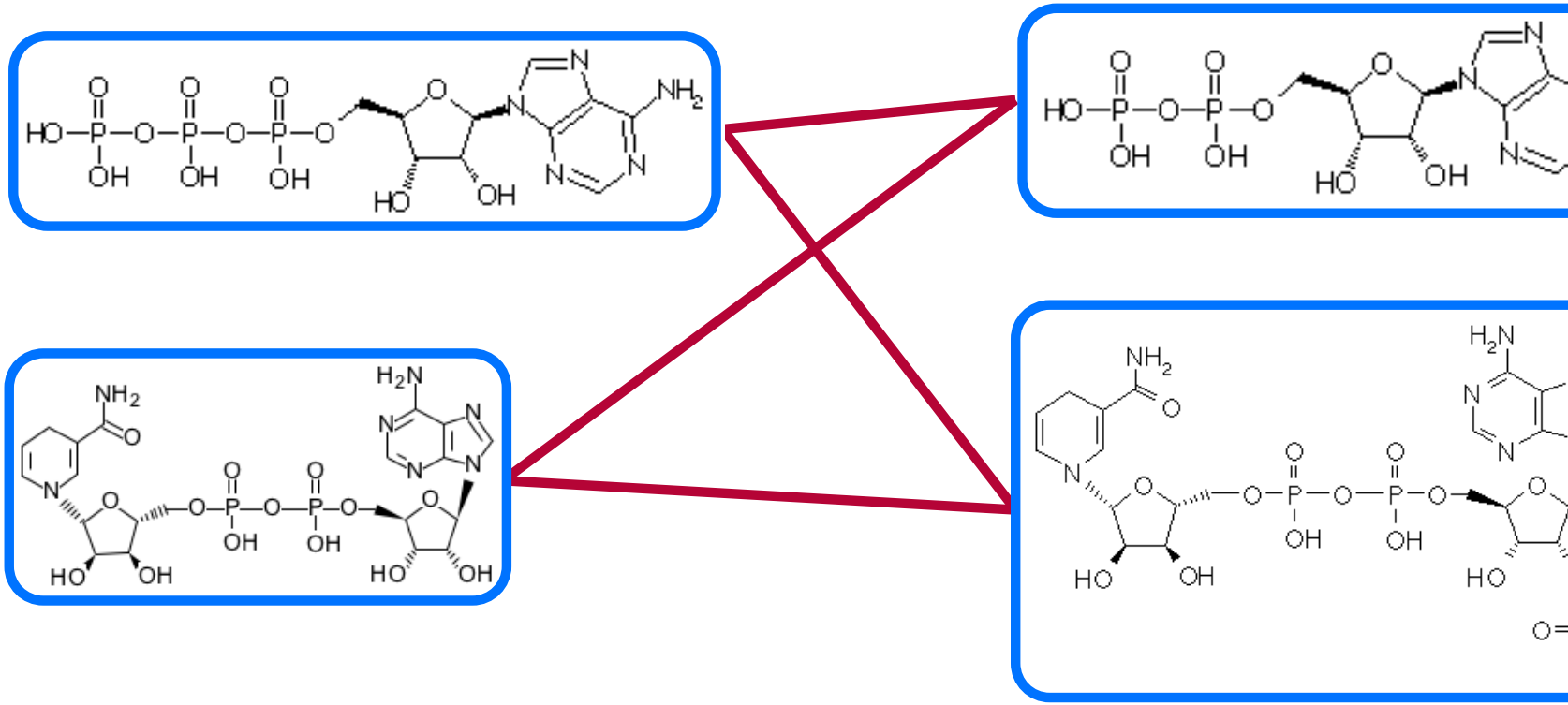
representations

ATP + NADH ↔ ADP + NADPH



representations

ATP + NADH ↔ ADP + NADPH

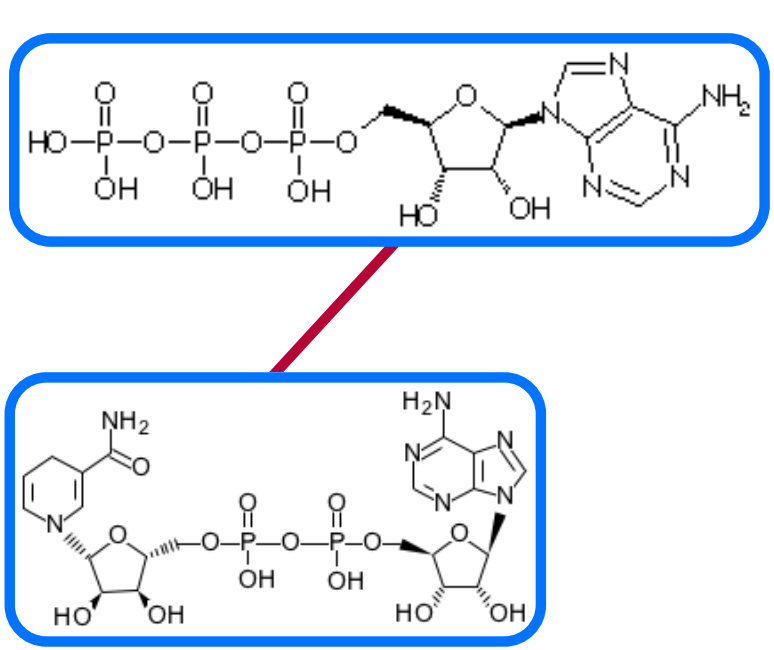


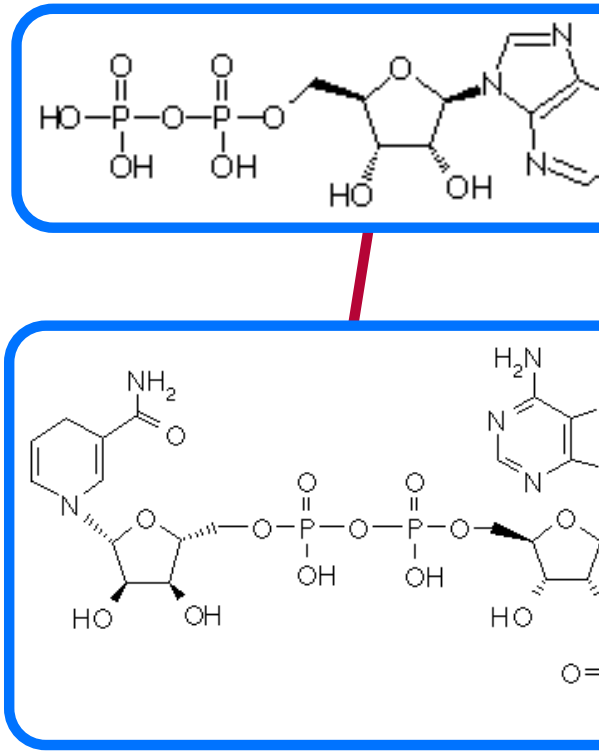
vertices: substrates

edges: between products / substrates on different sides

representations

ATP + NADH ↔ ADP + NADPH



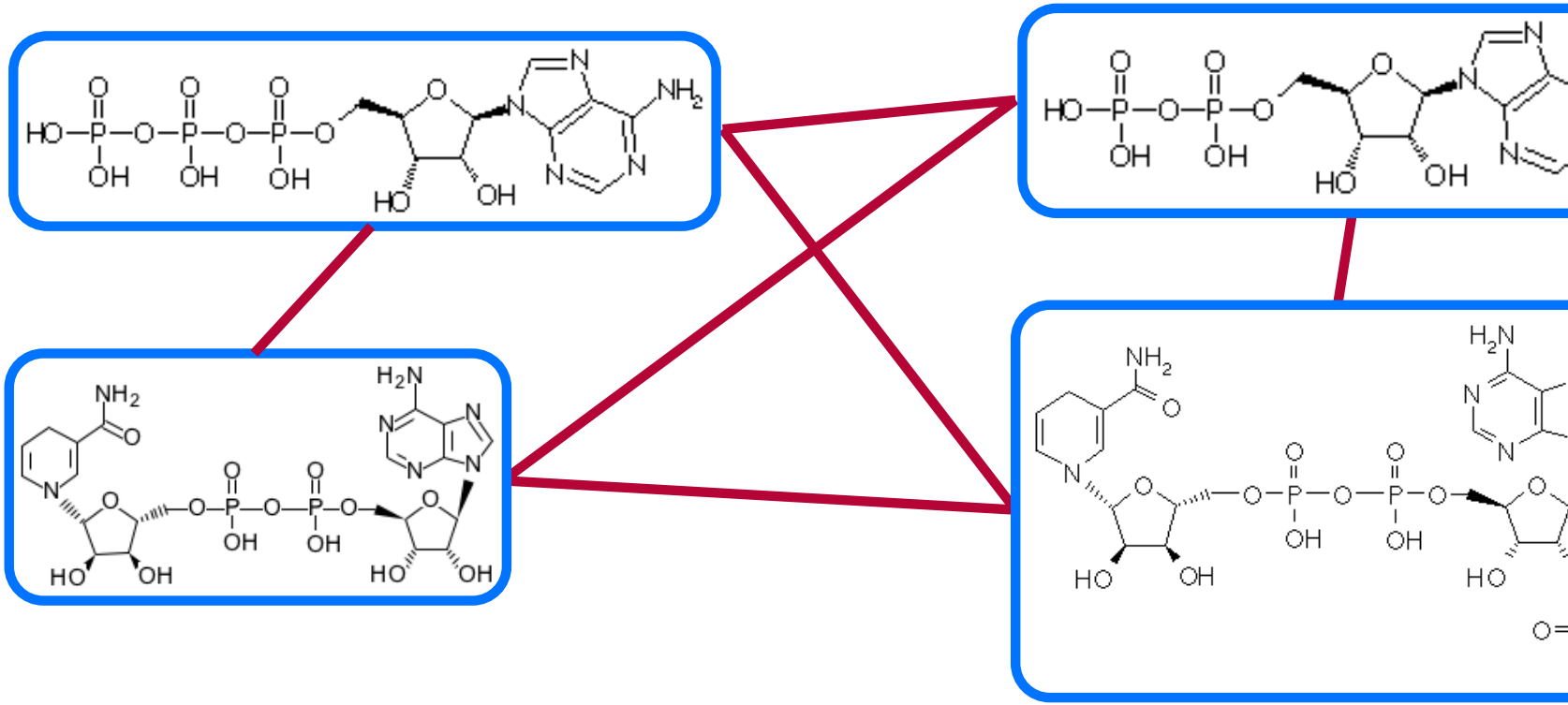


vertices: substrates

edges: between products / substrates on the same side

representations

ATP + NADH ↔ ADP + NADPH

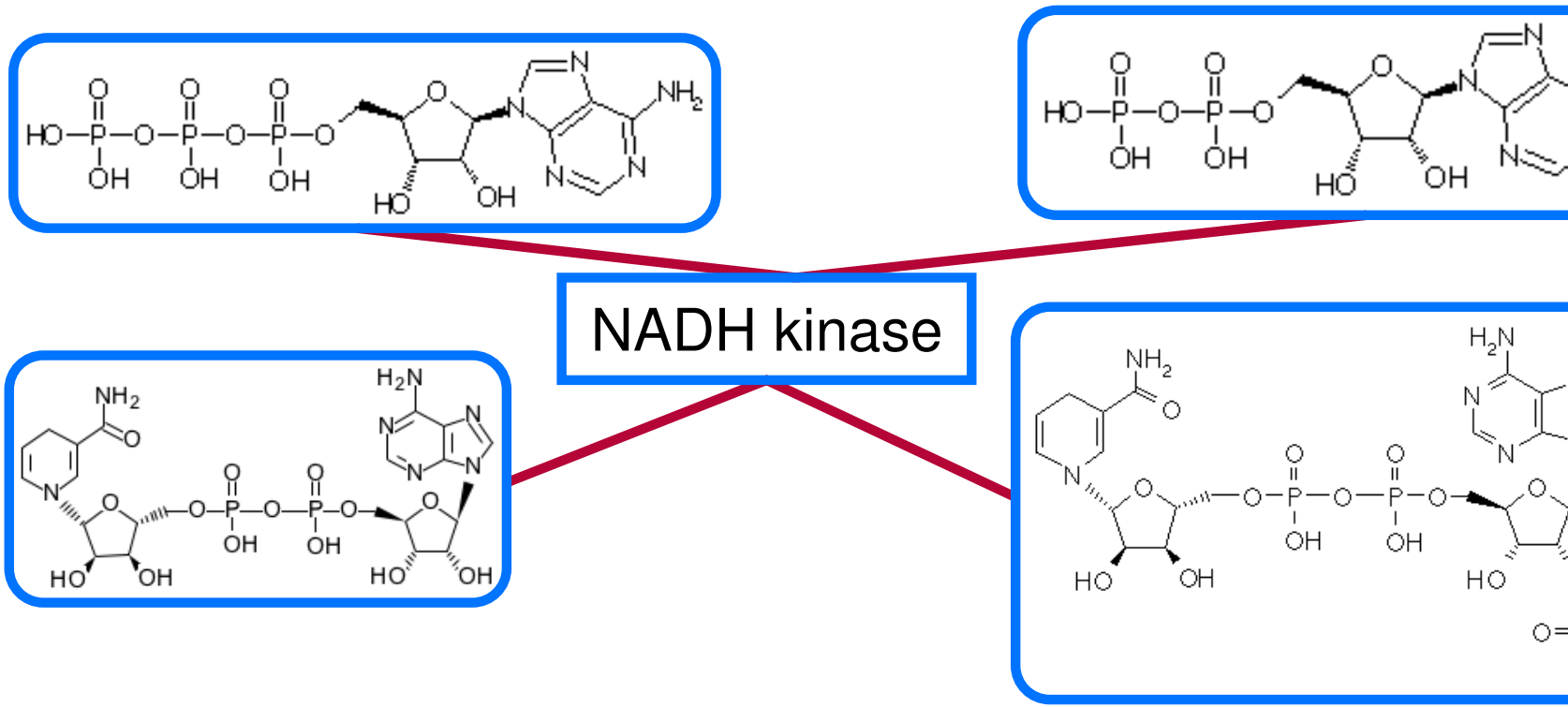


vertices: substrates

edges: between all products / substrates

representations

ATP + NADH ↔ ADP + NADPH



vertices: substrates + enzymes (reactions)

edges: between substance / reaction vertices

representations

- different representations, give different information
- most common representation = substance graphs—following atoms, the number of conversions between two molecules are small $\Leftrightarrow$ the graph distance is small

representations

- different representations, give different information
- most common representation = substance graphs—following atoms, the number of conversions between two molecules are small ⇔ the graph distance is small

representations

- different representations, give different information
- most common representation = substance graphs—following atoms, the number of conversions between two molecules are small $\Leftrightarrow$ the graph distance is small

the dogmas of network science

- real networks have both structure and randomness
- the network structure relates to the function of the network

the dogmas of network science

- real networks have both structure and randomness
- the network structure relates to the function of the network

the dogmas of network science

- real networks have both structure and randomness
- the network structure relates to the function of the network

what is network structure?

- how the network differs a random network
- **to be more precise:** how the network differs from a null model

what is network structure?

- how the network differs a random network
- **to be more precise:** how the network differs from a null model

what is network structure?

- how the network differs a random network
- **to be more precise:** how the network differs from a null model

network null-models

- Network structures are relative . . .
- . . . so one has to be clear about what to compare with . . . a null model
- *Null model 1:* random graphs (Poisson random graphs, Erdős-Rényi graphs)
- *Null model 2:* random graphs constrained to the set of degrees of the original graph

network null-models

- Network structures are relative . . .
- . . . so one has to be clear about what to compare with . . . a null model
- *Null model 1:* random graphs (Poisson random graphs, Erdős-Rényi graphs)
- *Null model 2:* random graphs constrained to the set of degrees of the original graph

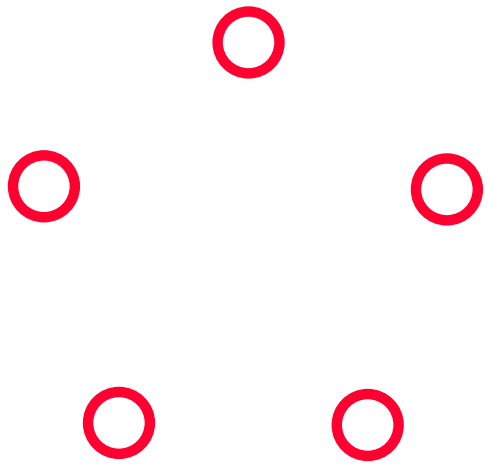
network null-models

- Network structures are relative . . .
- . . . so one has to be clear about what to compare with . . . a null model
- *Null model 1:* random graphs (Poisson random graphs, Erdős-Rényi graphs)
- *Null model 2:* random graphs constrained to the set of degrees of the original graph

network null-models

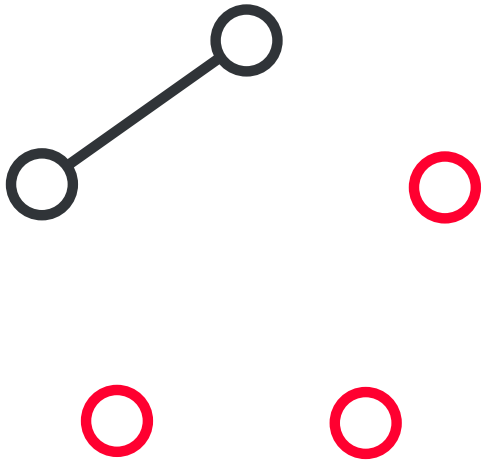
- Network structures are relative . . .
- . . . so one has to be clear about what to compare with . . . a null model
- *Null model 1:* random graphs (Poisson random graphs, Erdős-Rényi graphs)
- *Null model 2:* random graphs constrained to the set of degrees of the original graph

random graphs



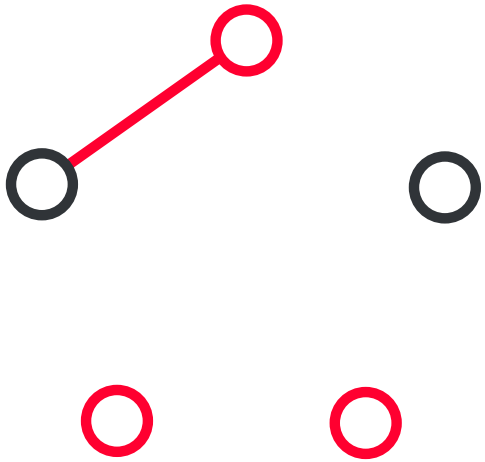
for each pair of vertices,
with probability p, add an edge

random graphs



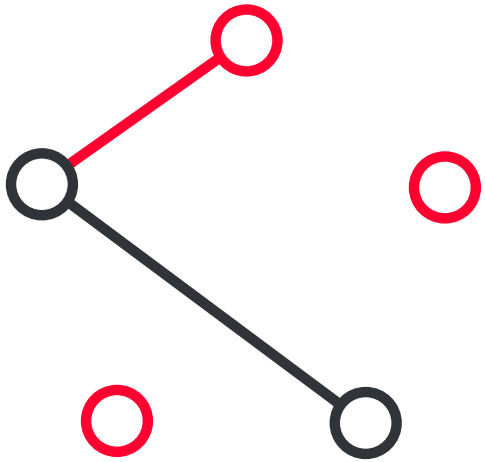
for each pair of vertices,
with probability p, add an edge

random graphs



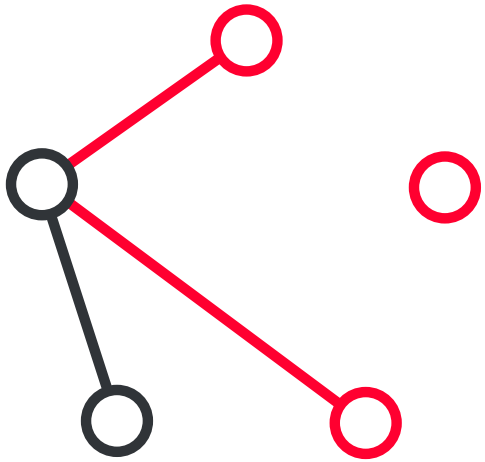
for each pair of vertices,
with probability p, add an edge

random graphs



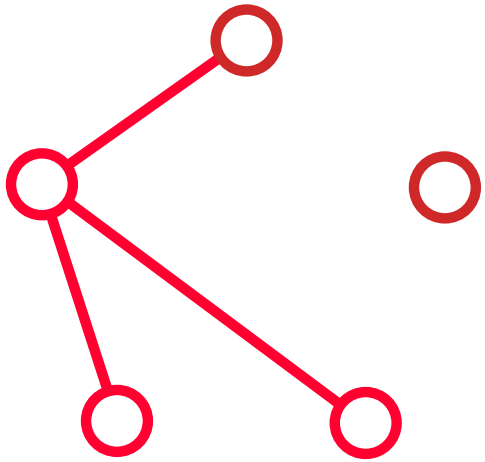
for each pair of vertices,
with probability p, add an edge

random graphs



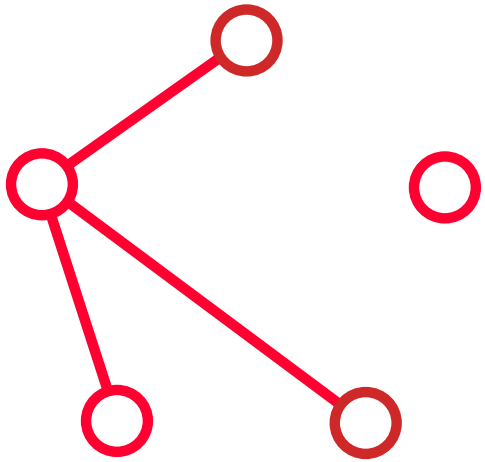
for each pair of vertices,
with probability p, add an edge

random graphs



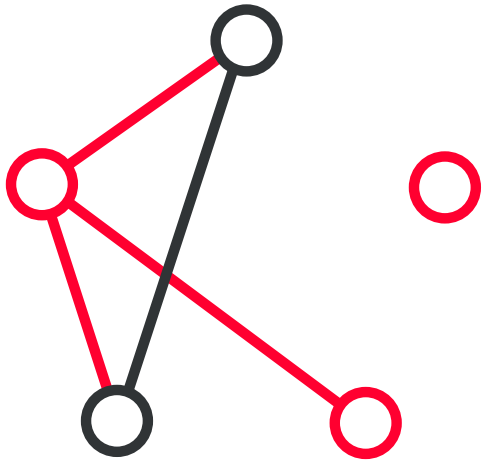
for each pair of vertices,
with probability p, add an edge

random graphs



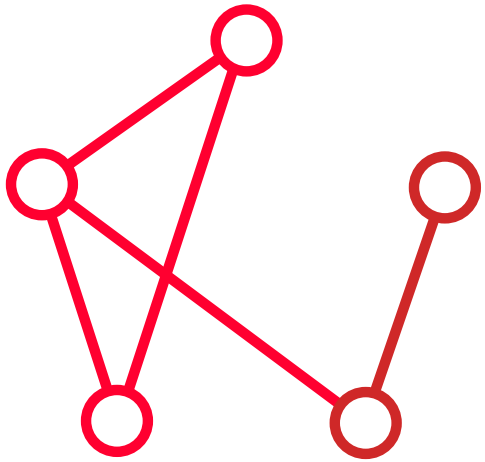
for each pair of vertices,
with probability p, add an edge

random graphs



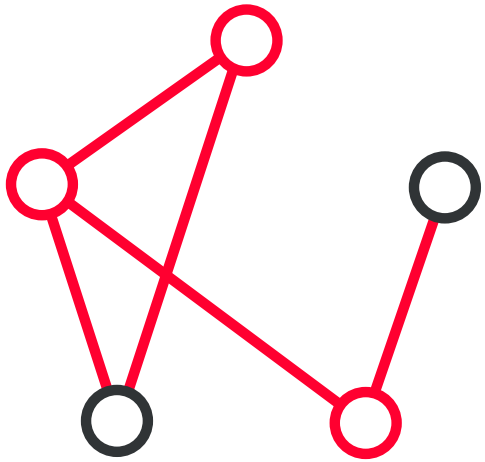
for each pair of vertices,
with probability p, add an edge

random graphs



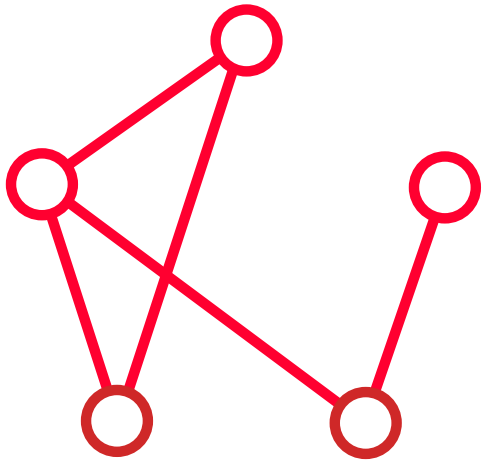
for each pair of vertices,
with probability p, add an edge

random graphs



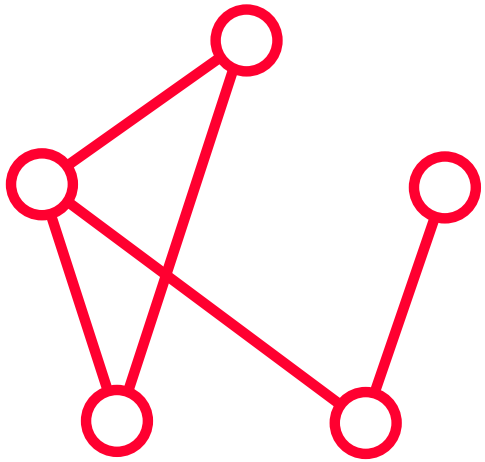
for each pair of vertices,
with probability p, add an edge

random graphs



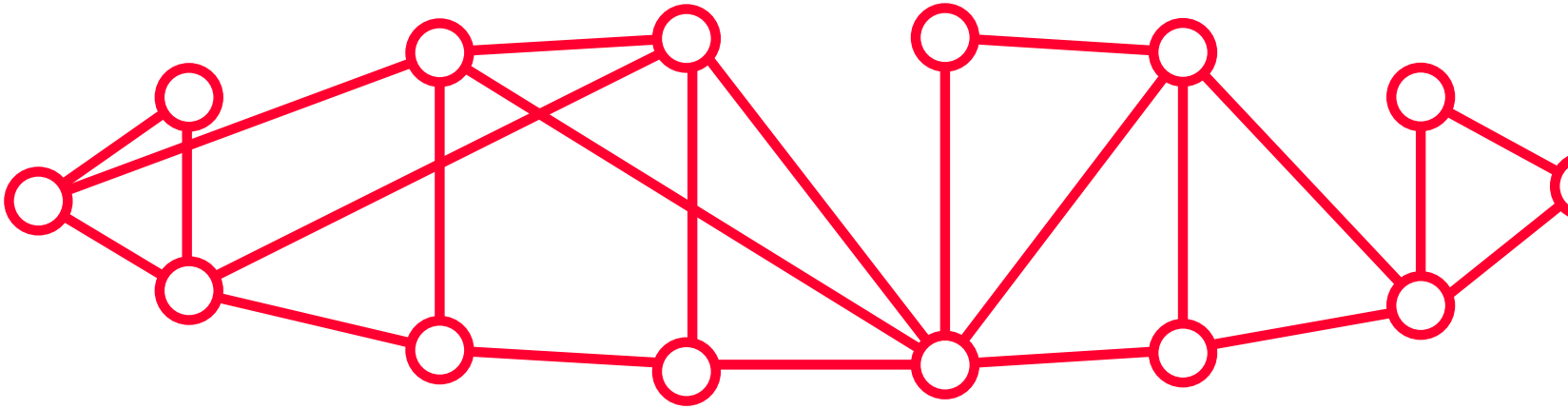
for each pair of vertices,
with probability p, add an edge

random graphs



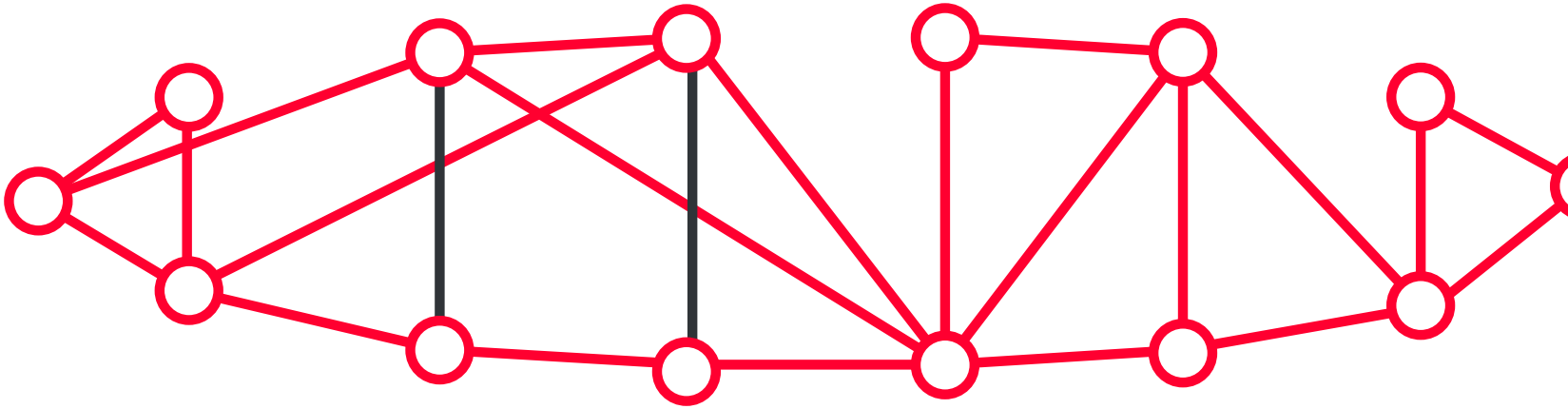
for each pair of vertices,
with probability p, add an edge

random rewiring



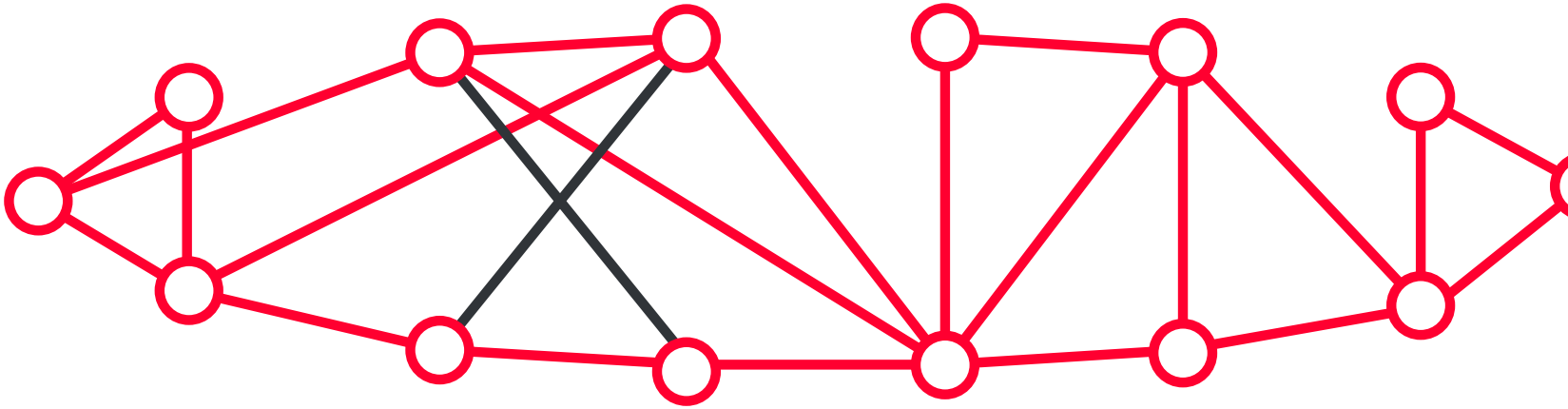
start from the original graph
choose edge pairs, and swap them

random rewiring



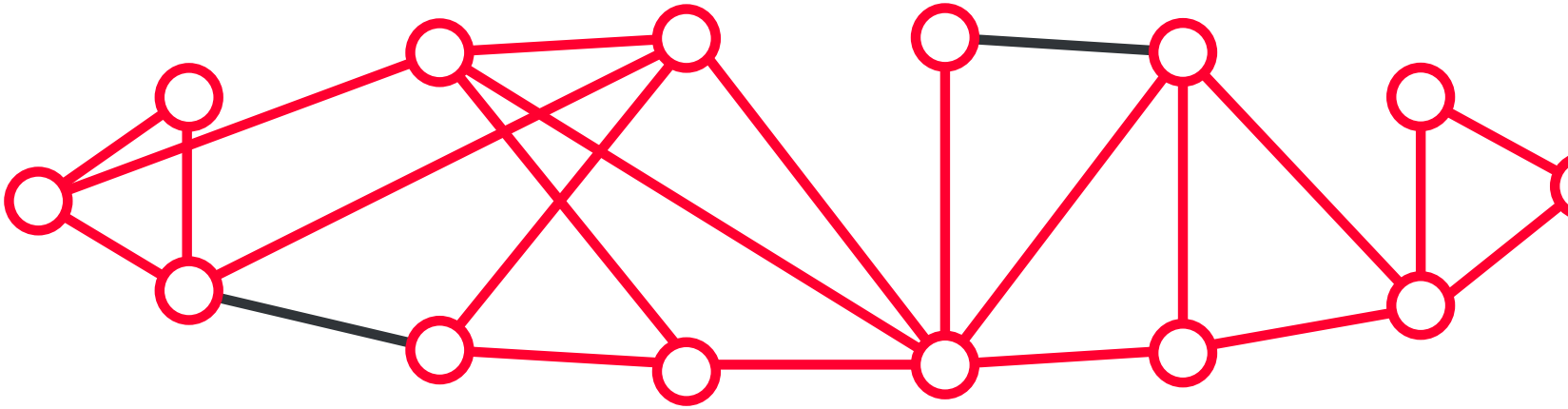
start from the original graph
choose edge pairs, and swap them

random rewiring



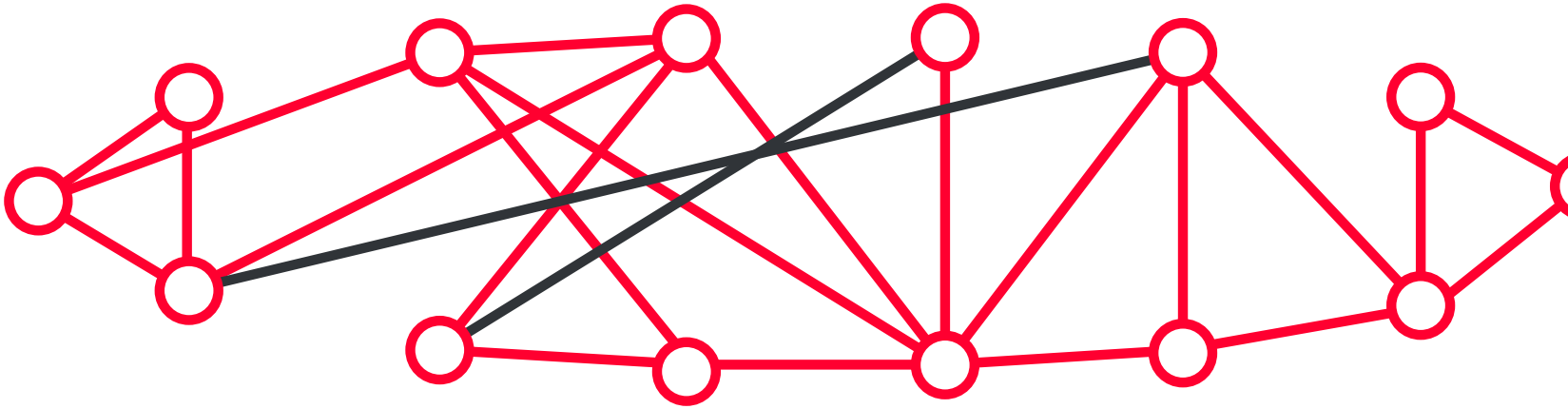
start from the original graph
choose edge pairs, and swap them

random rewiring



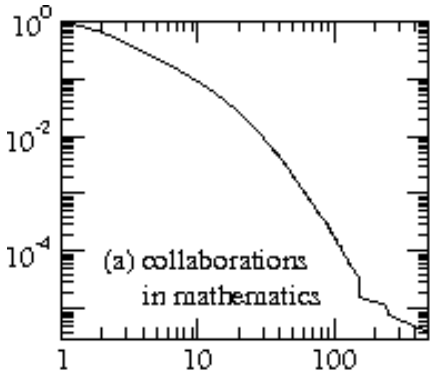
start from the original graph
choose edge pairs, and swap them

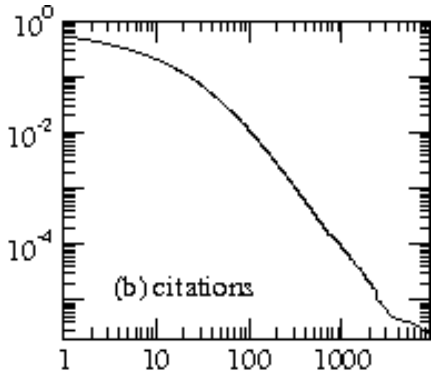
random rewiring

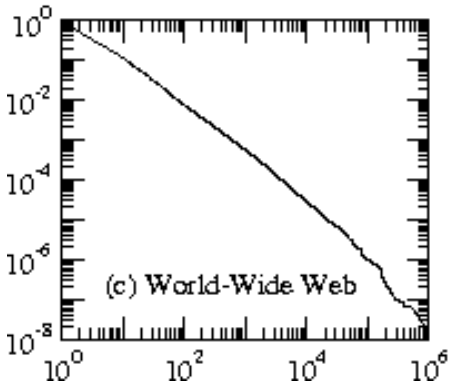


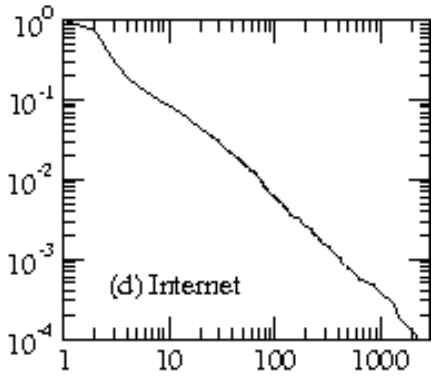
start from the original graph
choose edge pairs, and swap them

degree distribution









network structure of metabolism

- Power-law degree distributions.
- Increasing average degree.
- Network modularity?

network structure of metabolism

- Power-law degree distributions.
- Increasing average degree.
- Network modularity?

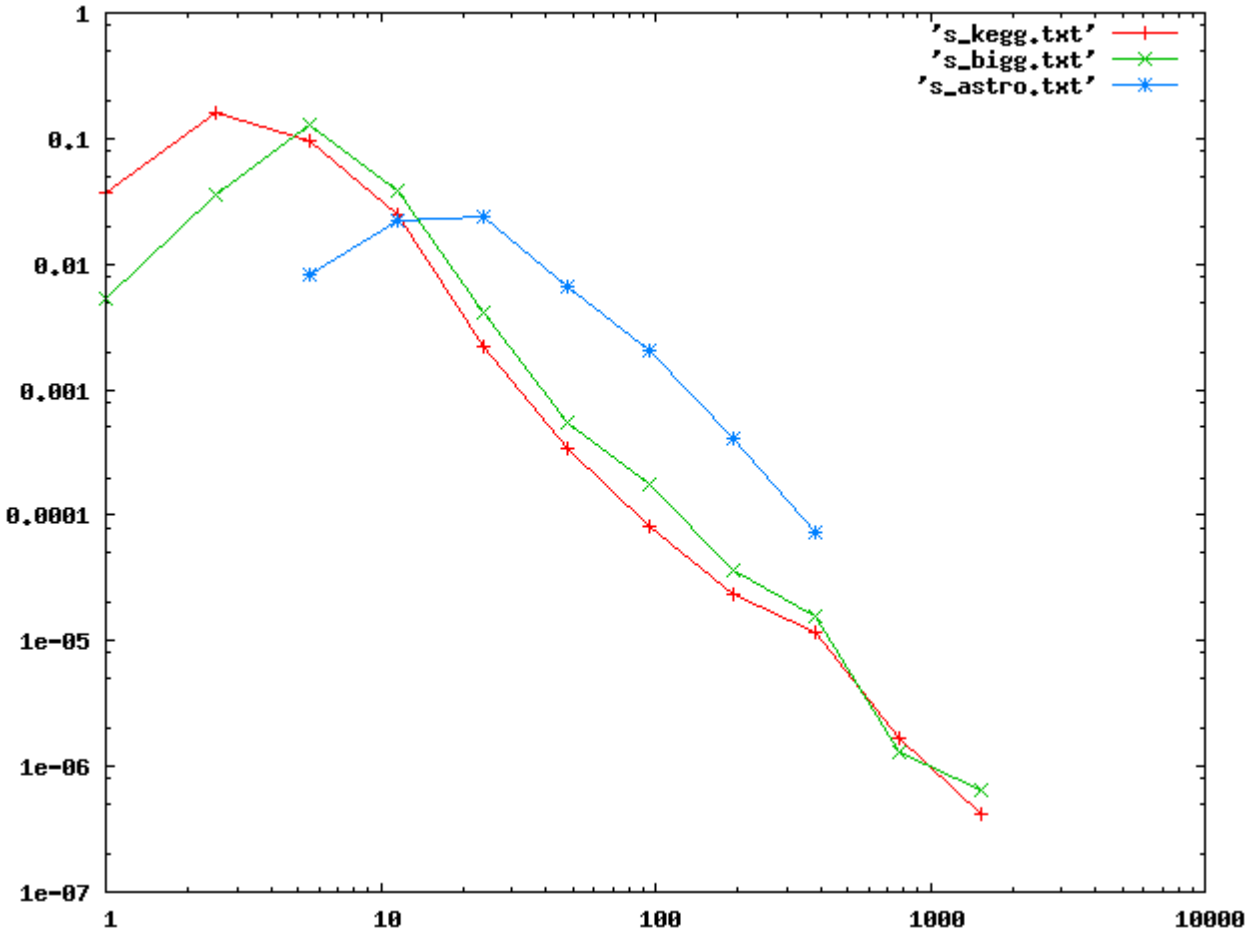
network structure of metabolism

- Power-law degree distributions.
- Increasing average degree.
- Network modularity?

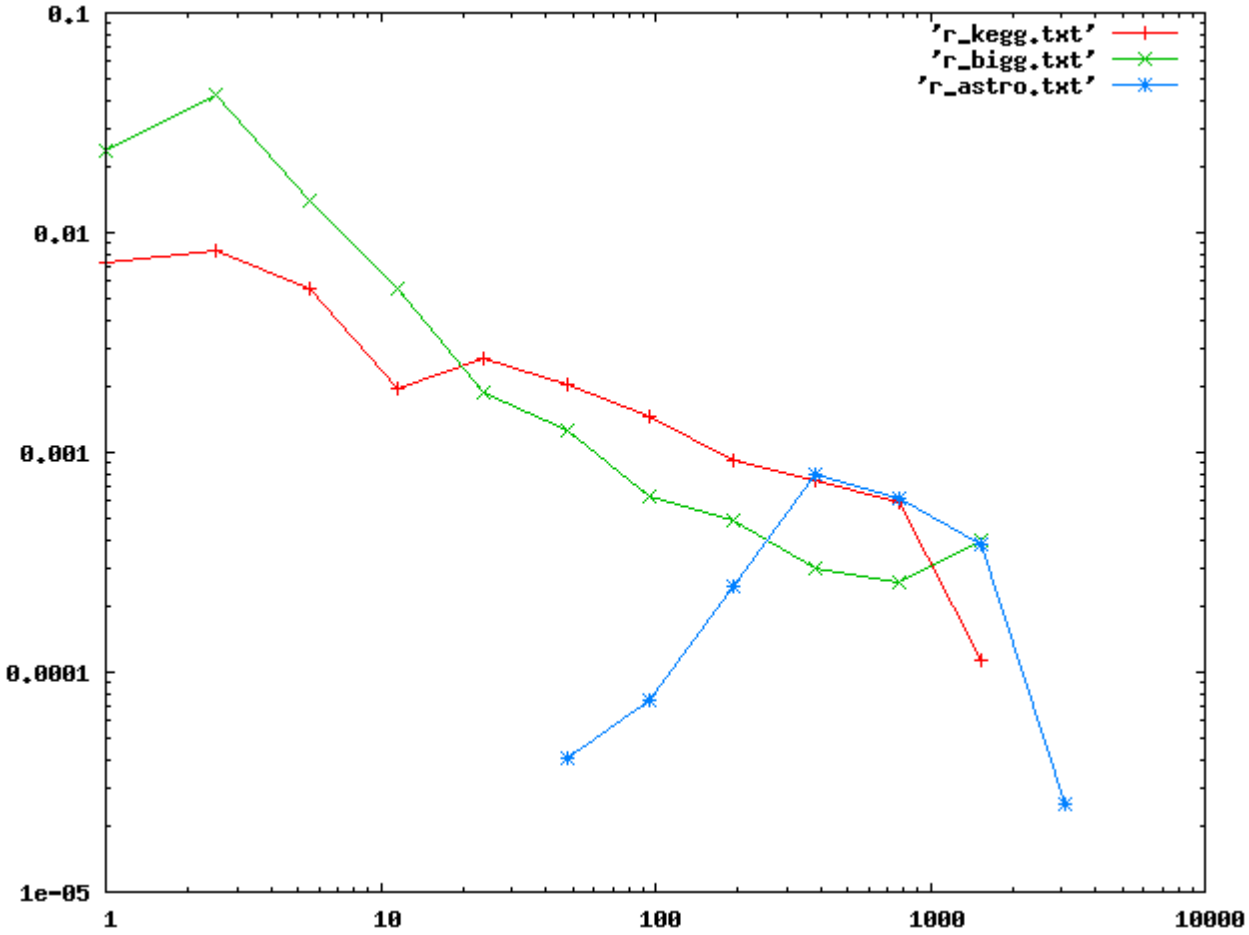
network structure of metabolism

- Power-law degree distributions.
- Increasing average degree.
- Network modularity?

network structure of metabolism



network structure of metabolism

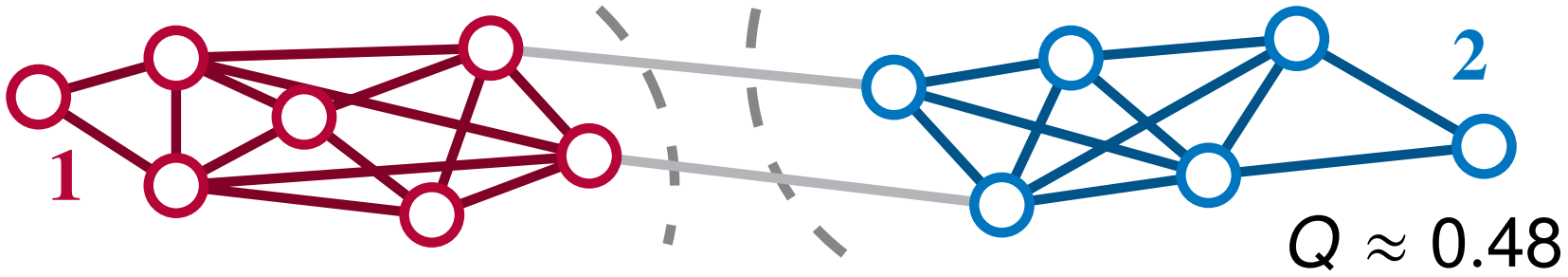


modularity (of a partition)

Definition

$$Q = \sum_i \left[e_{ii} - \left(\sum_j e_{ij} \right)^2 \right] \quad (1)$$

the sum is over the a partition into clusters and e_{ij} is the fraction of edges that leads between vertices of cluster i and j



modularity (of a graph)

Definition

$$\hat{Q}(G) = \max_{\text{partitions}} Q \tag{2}$$

(null model: random graphs)

Definition

$$\hat{Q}(G) = \max_{\text{partitions}} Q - \left\langle \max_{\text{partitions}} Q \right\rangle \tag{3}$$

(null model: random graphs with the same degree sequence)

modularity (of a graph)

Definition

$$\hat{Q}(G) = \max_{\text{partitions}} Q \tag{2}$$

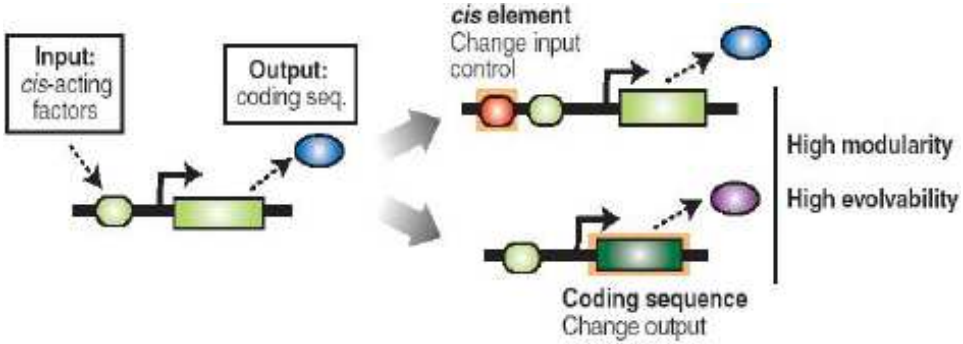
(null model: random graphs)

Definition

$$\hat{Q}(G) = \max_{\text{partitions}} Q - \left\langle \max_{\text{partitions}} Q \right\rangle \tag{3}$$

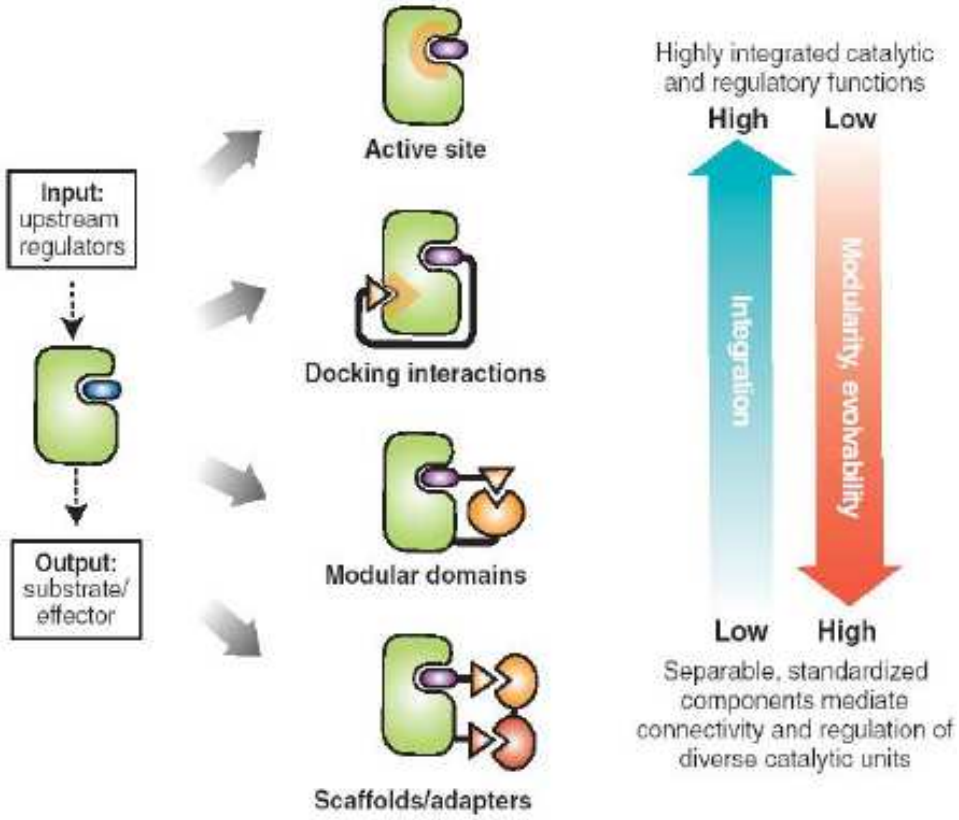
(null model: random graphs with the same degree sequence)

modularity (the biological idea)



(Bhattacharyya *et al.*, 2006, Annu. Rev. Biochem. **75**, pp. 655–80)

modularity (the biological idea)



(Bhattacharyya *et al.*, 2006, Annu. Rev. Biochem. **75**, pp. 655–80)

modularity and robustness

- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

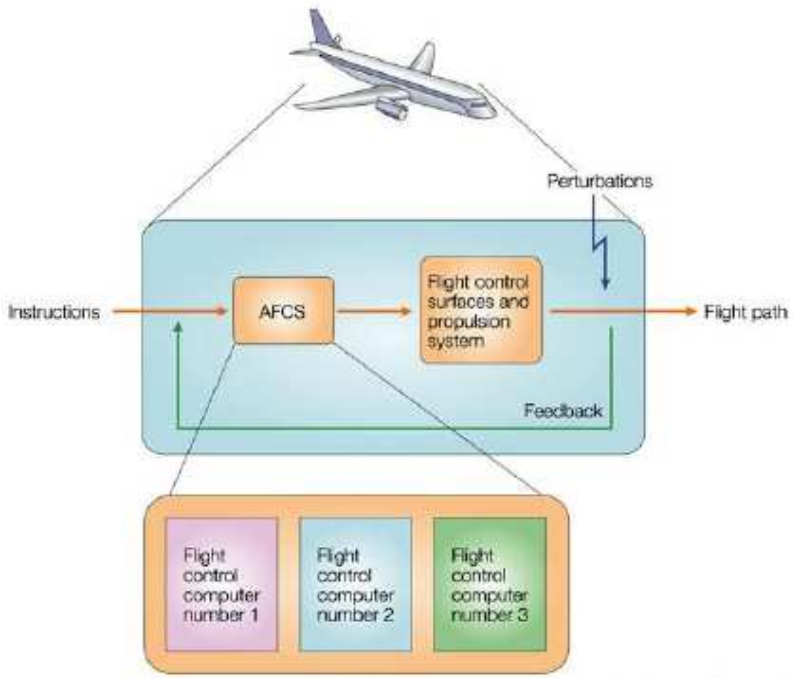
- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

- robustness—sensitivity *of a relevant dynamic system on the network* to perturbations
- perturbations = changes in:
  - network topology
  - concentrations
  - delays in periodic input
- in spreading of harmful things (like disease), modularity increases robustness
- in other systems where a flow is needed throughout the networks, modularity might decrease robustness

modularity and robustness

modular redundancy



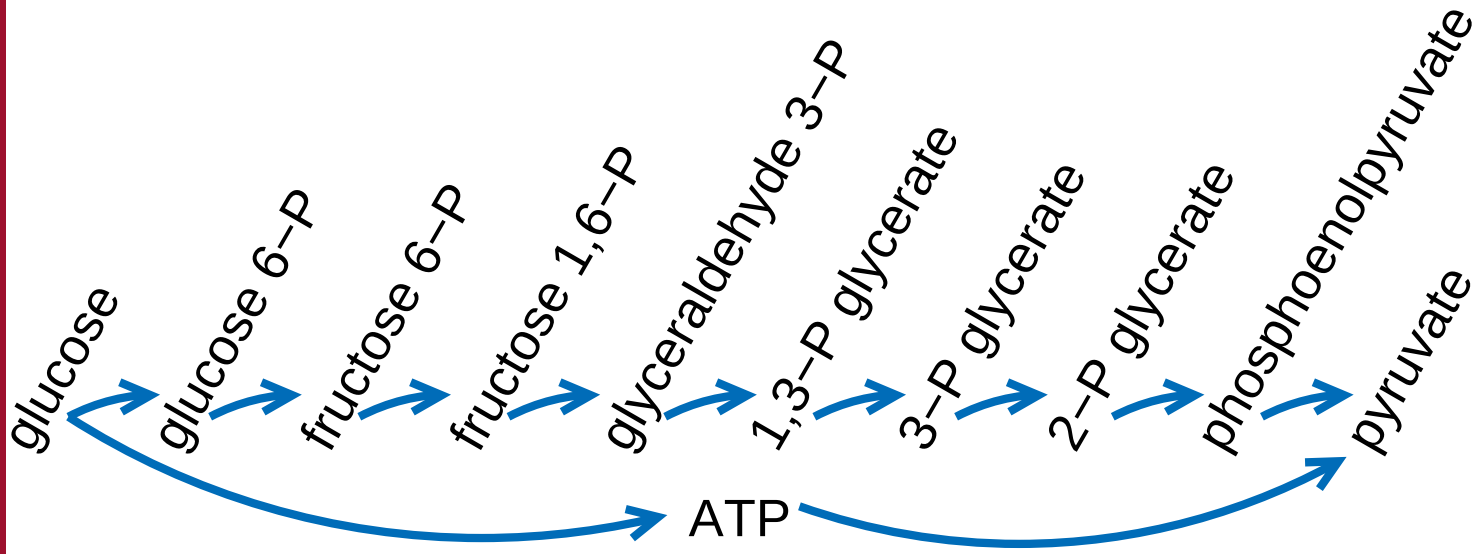
(Kitano, 2004, Nat. Rev. Genet. **5**, pp. 826–837.)

modularity and robustness

distributed redundancy (A. Wagner):

> *In distributed robustness, many parts of a system contribute to its function, but all of these parts have different roles. When one part fails or is changed through mutations, other parts can compensate for this failure, but not simply by standing in for the failed part.*

currency metabolites



currency metabolites

Wagner & Fell, 2001	Schuster *et al.*, 2002	Ma & Zeng, 2003
ATP	ATP	ATP
ADP	ADP	ADP
NADP	NADP	NADP
NADPH	NADPH	
NAD		NAD
NADH		
	P_i	P_i
	H_2O	H_2O
	H^+	
	PP_i	
	CMP	
		CO_2
		O_2
		NH_3

currency metabolites (a definition)

1. currency metabolites have high degree
2. they make not meaningful shortcuts
3. i.e. tie together distant parts of the network
4. i.e. tie different modules together

. . . let's turn this around to a definition . . .

currency metabolites (a definition)

1. currency metabolites have high degree
2. they make not meaningful shortcuts
3. i.e. tie together distant parts of the network
4. i.e. tie different modules together

. . . let's turn this around to a definition . . .

currency metabolites (a definition)

1. currency metabolites have high degree
2. they make not meaningful shortcuts
3. i.e. tie together distant parts of the network
4. i.e. tie different modules together

. . . let's turn this around to a definition . . .

currency metabolites (a definition)

1. currency metabolites have high degree
2. they make not meaningful shortcuts
3. i.e. tie together distant parts of the network
4. i.e. tie different modules together

. . . let's turn this around to a definition . . .

currency metabolites (a definition)

1. currency metabolites have high degree
2. they make not meaningful shortcuts
3. i.e. tie together distant parts of the network
4. i.e. tie different modules together

. . . let's turn this around to a definition . . .

currency metabolites (a definition)

1. currency metabolites have high degree
2. they make not meaningful shortcuts
3. i.e. tie together distant parts of the network
4. i.e. tie different modules together

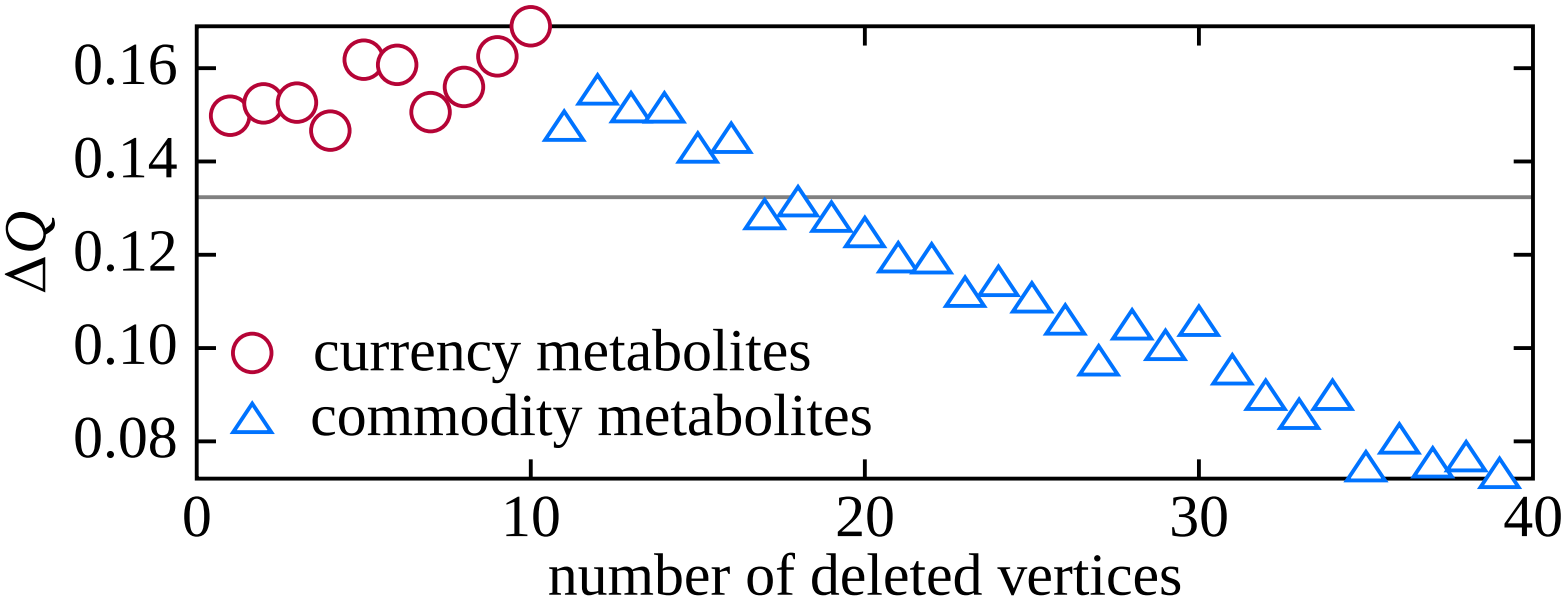
. . . let's turn this around to a definition . . .

currency metabolites (a definition)

Definition

Remove vertices in order of (currently) highest degree. The set of removed vertices that gives the network the highest modularity is the set of currency metabolites.

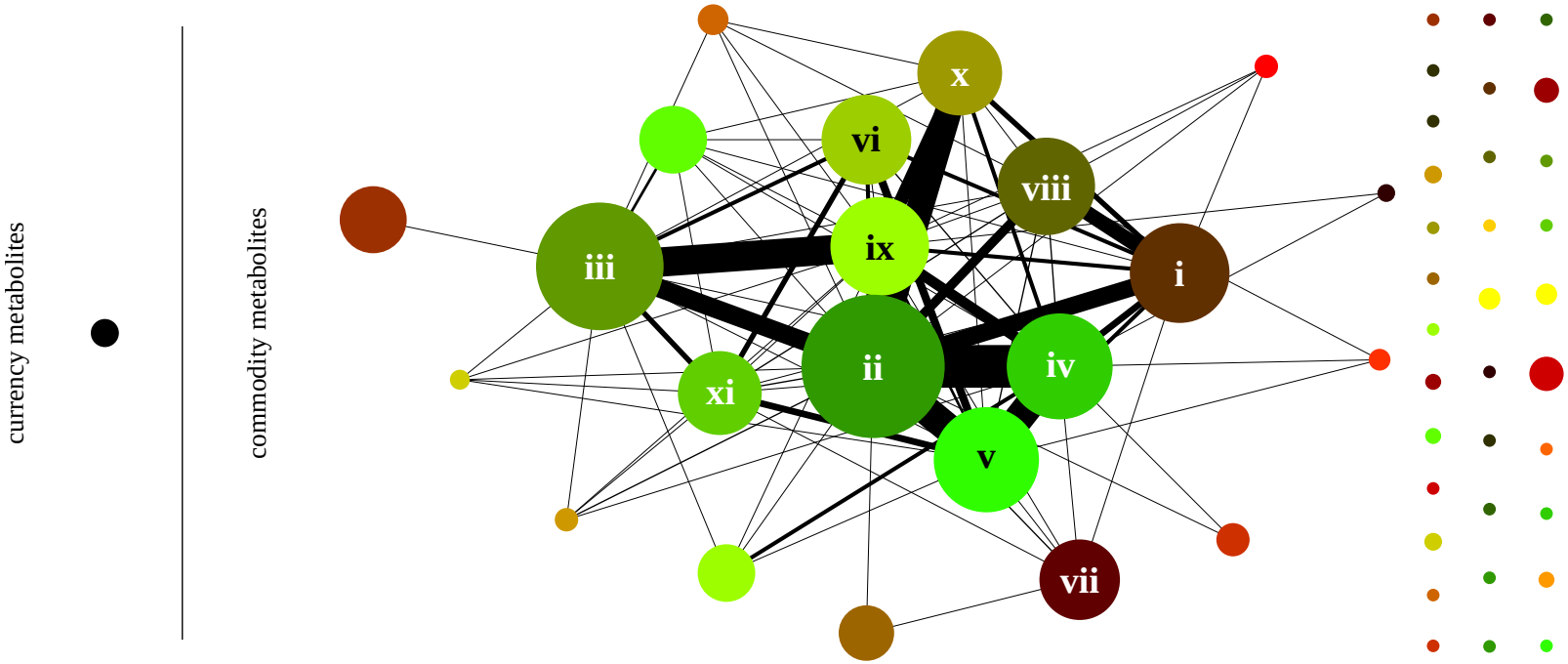
human currency metabolites



detected currency metabolites

Wagner & Fell, 2001	Schuster *et al.*, 2002	Ma & Zeng, 2003
ATP	**ATP**	**ATP**
ADP	**ADP**	**ADP**
NADP	**NADP**	**NADP**
NADPH	**NADPH**	
NAD		**NAD**
NADH		
	P_i	P_i
	H_2O	**H_2O**
	H^+	
	PP_i	
	CMP	
		CO_2
		O_2
		NH_3

modularity and robustness



i. nitrogen metabolism
ii. amino acid metabolism / protein synthesis
iii. sugar metabolism
iv. citric acid cycle / porphyrin synthesis
v. fatty acid metabolism
vi. steroid metabolism
vii. metabolism of xenobiotics
viii. tyrosine and tryptamine metabolism
ix. DNA metabolism / glycolysis
x. nucleoside / nucleotide metabolism
xi. glycerophospholipid and arachidonic acid metabolism

different organisms

organism	samples	nodes	edges	c. m.	modularity
animals	5	1621	4662	6.2	0.157
plants	1	1561	4302	1	0.144
fungi	2	1281	3654	1.5	0.150
bacteria	99	1050	2739	1.7	0.140

conclusions

- proposed graph based definition of currency metabolites
- metabolic networks *are* modular, but not so much
- the reason they are not more modular might be give robustness

(Huss & Holme, 2007, IET Syst. Biol. **1**, pp. 280–5.)

conclusions

- proposed graph based definition of currency metabolites
- metabolic networks *are* modular, but not so much
- the reason they are not more modular might be give robustness

(Huss & Holme, 2007, IET Syst. Biol. **1**, pp. 280–5.)

conclusions

- proposed graph based definition of currency metabolites
- metabolic networks *are* modular, but not so much
- the reason they are not more modular might be give robustness

(Huss & Holme, 2007, IET Syst. Biol. **1**, pp. 280–5.)

conclusions

- proposed graph based definition of currency metabolites
- metabolic networks *are* modular, but not so much
- the reason they are not more modular might be give robustness

(Huss & Holme, 2007, IET Syst. Biol. **1**, pp. 280–5.)

conclusions

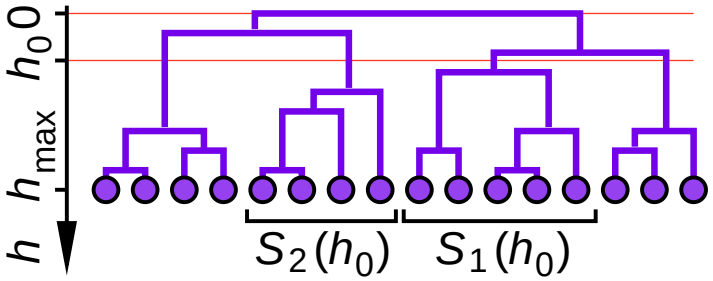
- proposed graph based definition of currency metabolites
- metabolic networks *are* modular, but not so much
- the reason they are not more modular might be give robustness

(Huss & Holme, 2007, IET Syst. Biol. **1**, pp. 280–5.)

subnetwork hierarchies

P. Holme, M. Huss & H. Jeong, 2003, Bioinformatics **19**, pp. 532–8.

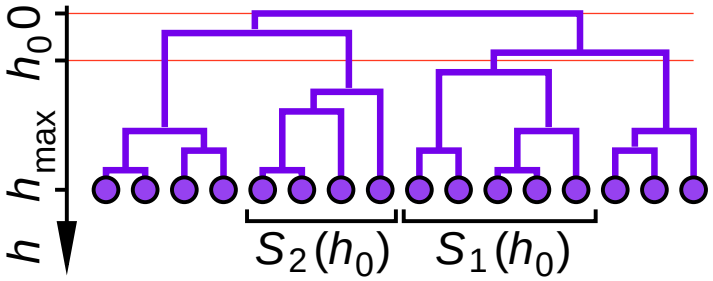
- Start with a directed bipartite networks, with both reaction and substance vertices (keep the currency metabolites).
- Iteratively remove the reaction vertices with the highest *betweenness* (fraction of shortest paths passing through a vertex).
- Study the dendrogram of this process.



subnetwork hierarchies

P. Holme, M. Huss & H. Jeong, 2003, Bioinformatics **19**, pp. 532–8.

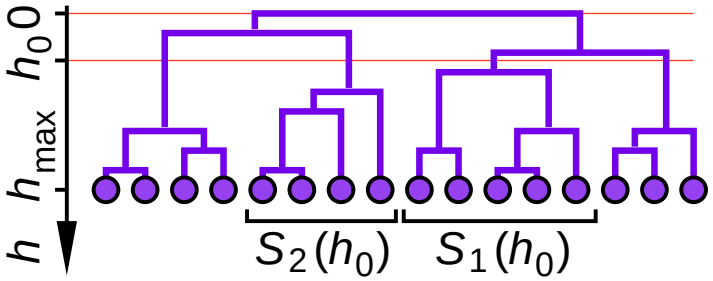
- Start with a directed bipartite networks, with both reaction and substance vertices (keep the currency metabolites).
- Iteratively remove the reaction vertices with the highest *betweenness* (fraction of shortest paths passing through a vertex).
- Study the dendrogram of this process.



subnetwork hierarchies

P. Holme, M. Huss & H. Jeong, 2003, Bioinformatics **19**, pp. 532–8.

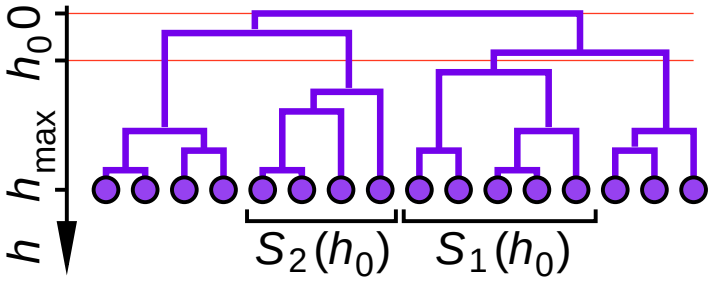
- Start with a directed bipartite networks, with both reaction and substance vertices (keep the currency metabolites).
- Iteratively remove the reaction vertices with the highest *betweenness* (fraction of shortest paths passing through a vertex).
- Study the dendrogram of this process.



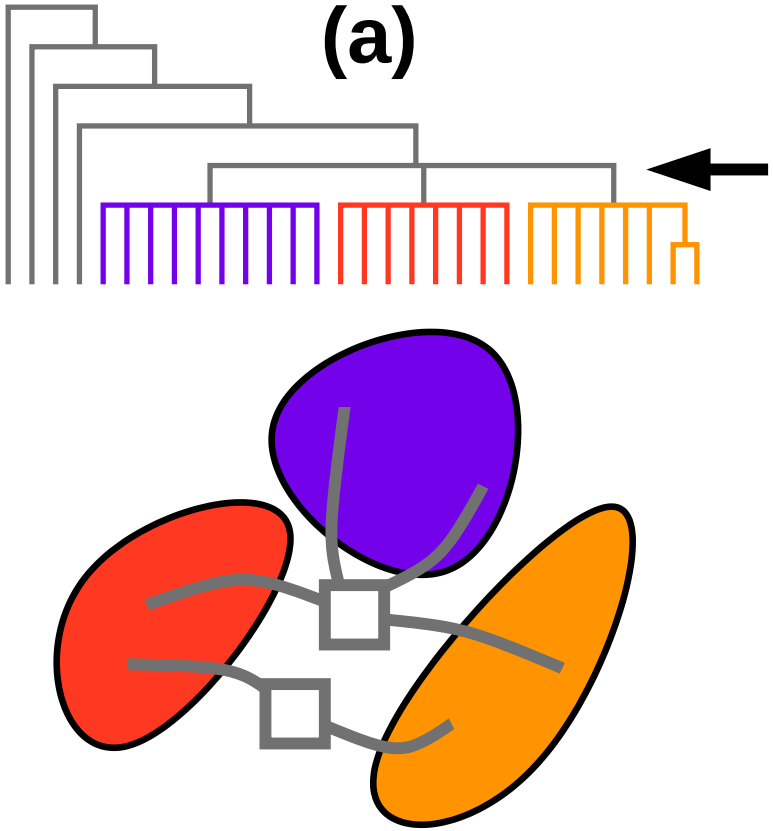
subnetwork hierarchies

P. Holme, M. Huss & H. Jeong, 2003, Bioinformatics **19**, pp. 532–8.

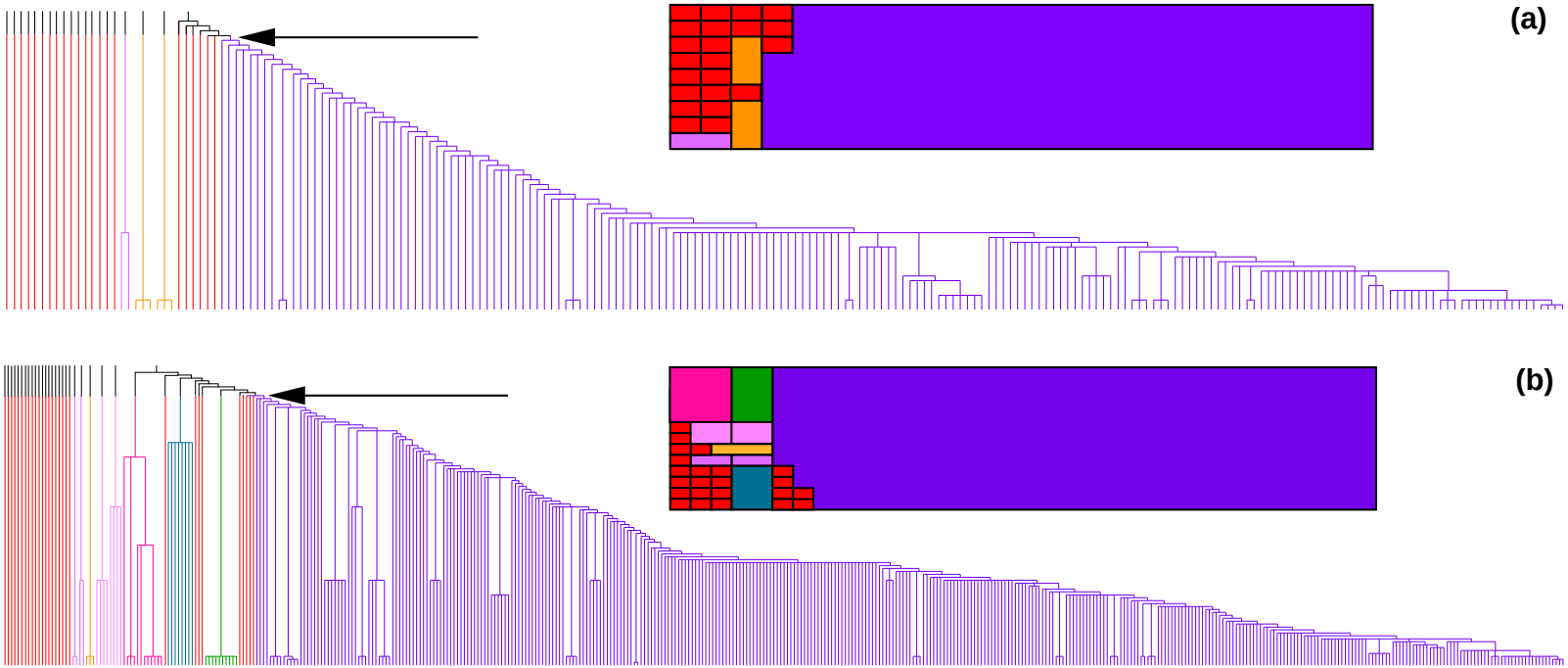
- Start with a directed bipartite networks, with both reaction and substance vertices (keep the currency metabolites).
- Iteratively remove the reaction vertices with the highest *betweenness* (fraction of shortest paths passing through a vertex).
- Study the dendrogram of this process.



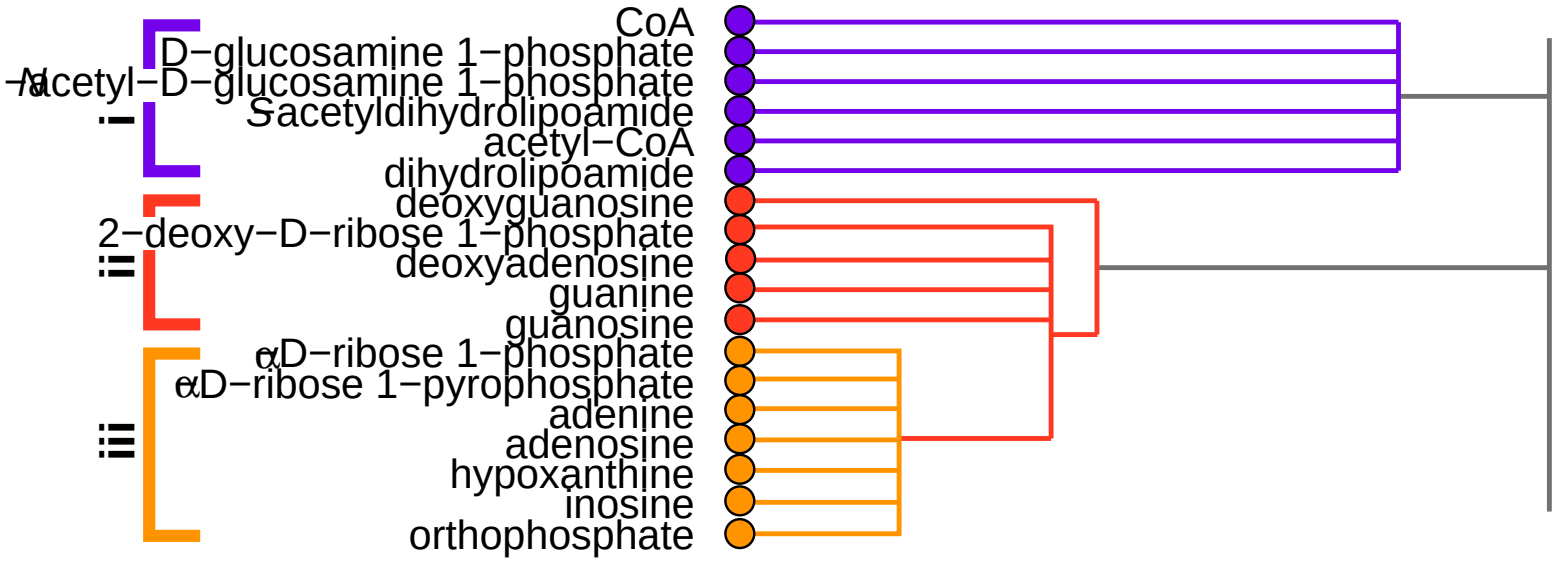
shell- vs community-type ordering



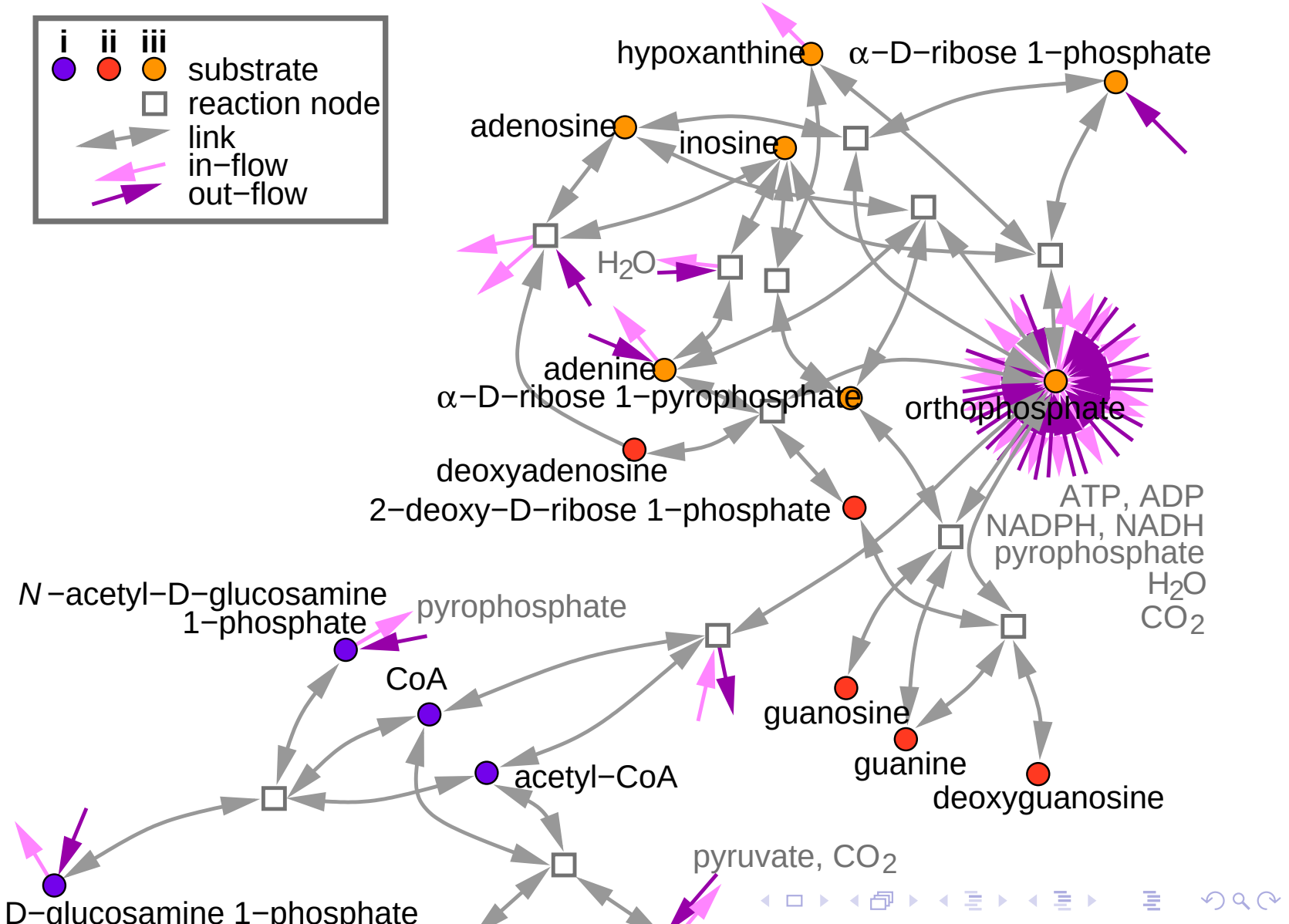
dendrogram



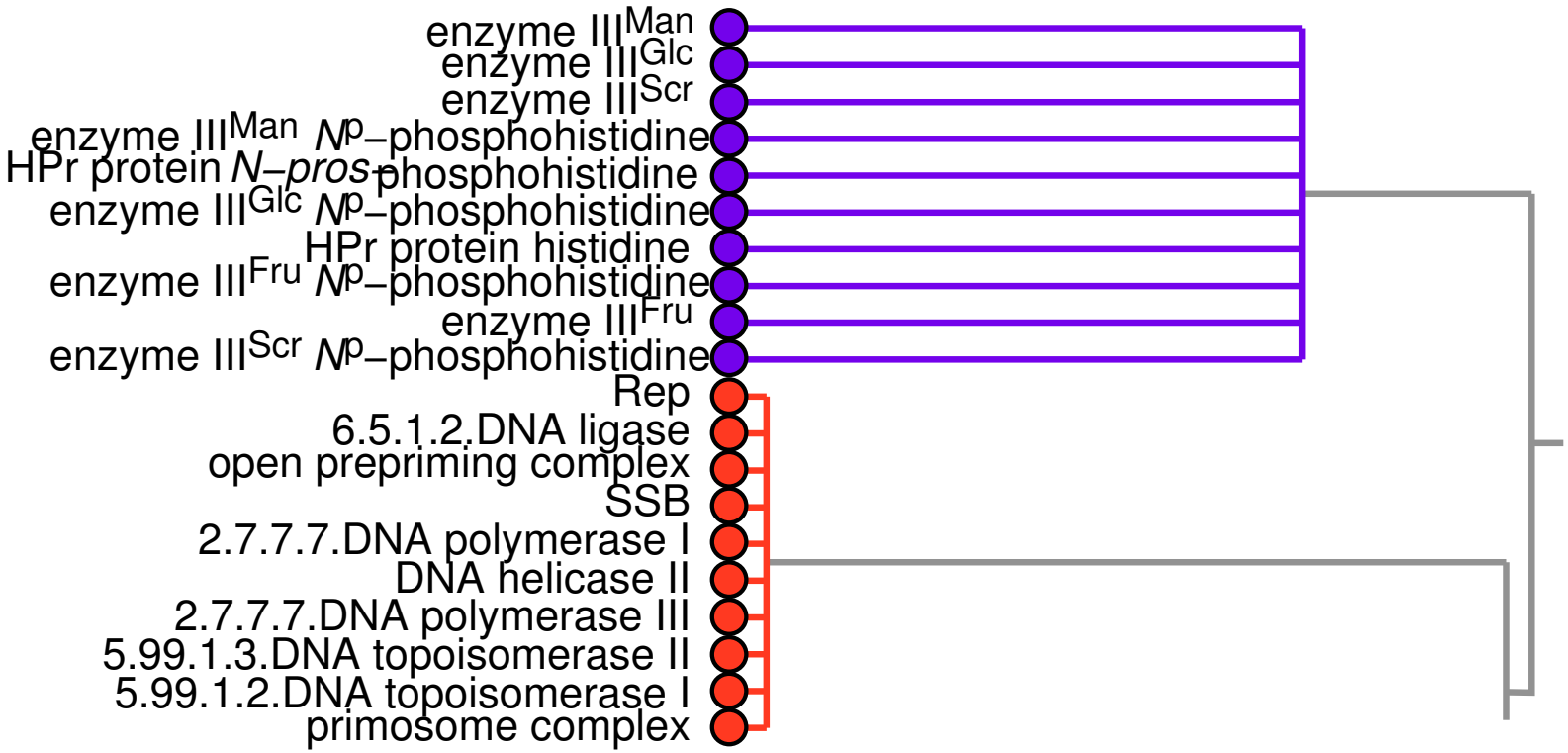
modules



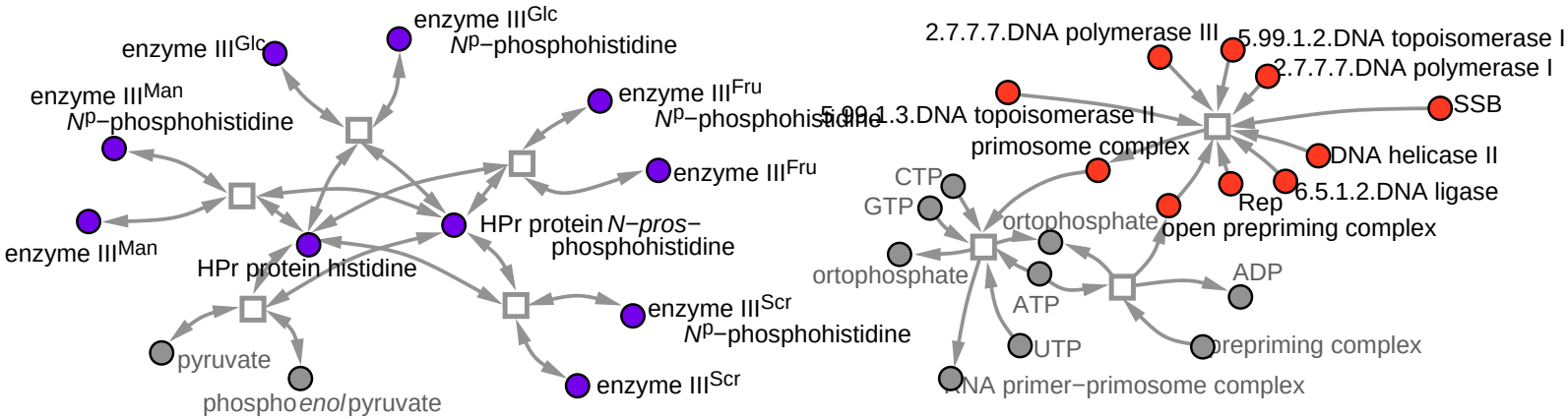
modules



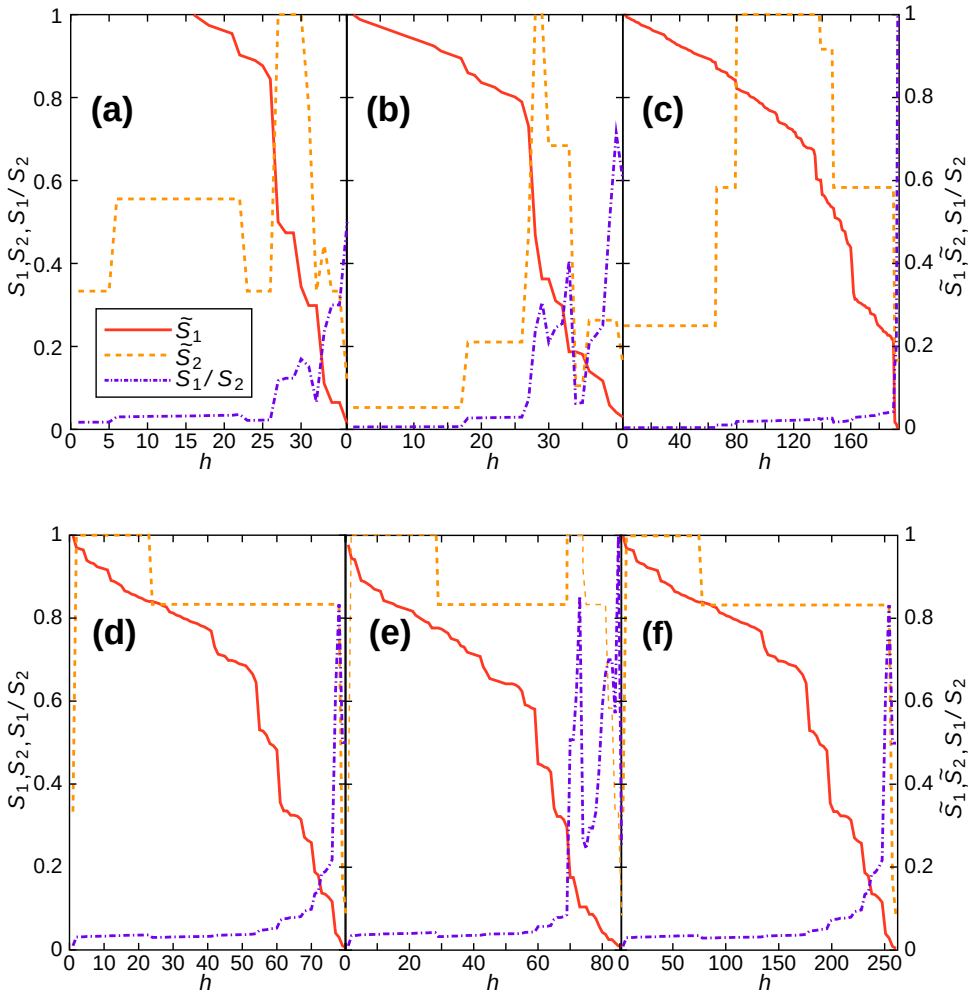
modules



modules



dentrogram statistics



conclusions

- Shell- rather than community-type ordering
- A few, seemingly clear, modules

conclusions

- Shell- rather than community-type ordering
- A few, seemingly clear, modules

conclusions

- Shell- rather than community-type ordering
- A few, seemingly clear, modules