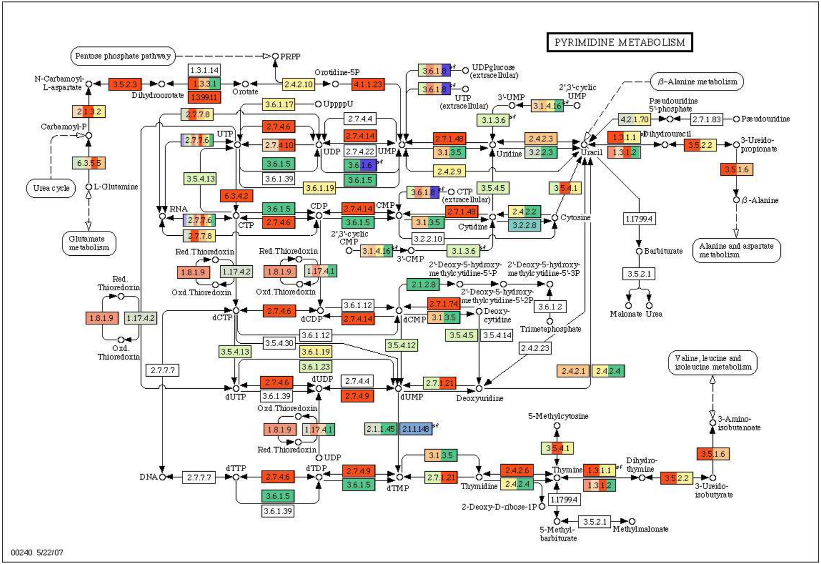
Robustness and Modularity in Metabolic Networks

Alexander Ullrich

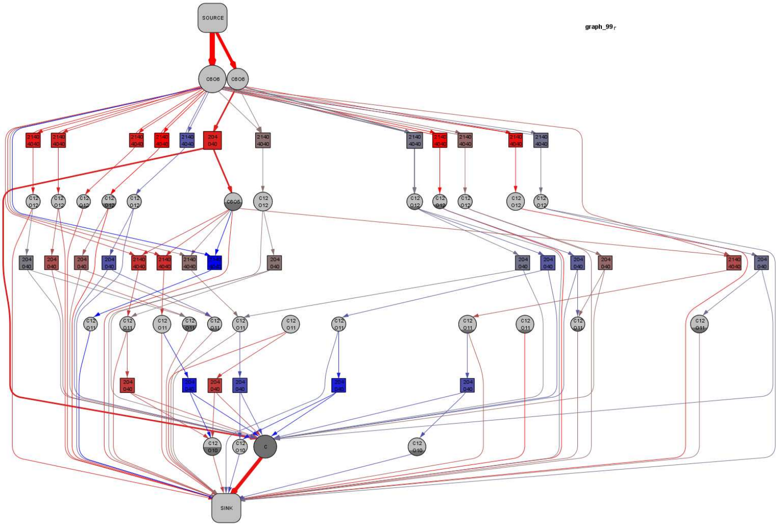
Bioinformatics
University of Leipzig

Bled, February 17

Metabolic Networks

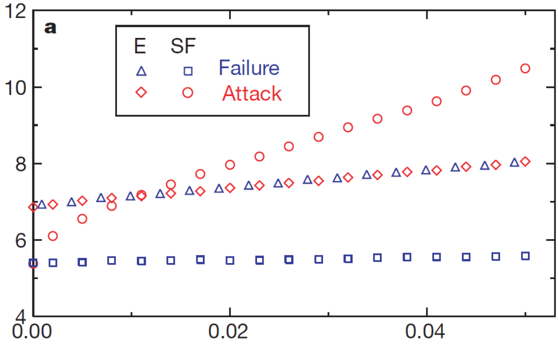


Metabolic Networks



Robustness

- Ability to function despite changes
- Genetic Changes: Mutations,...
- Epigenetic Changes: Fluctuations in Molecule concetrations
- Complex Systems are highly robust
- Scale-free Networks are particularly robust



Modularity

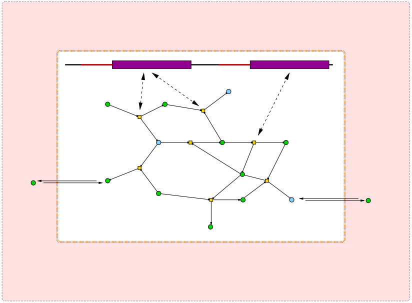
- Module: (structural) subsystem with distinct function
- Key organizing principle of biological networks
- High Clustering Coefficient suggests Modularity
- However, Origin and Preservation of Modularity not understood
- Changing Goals or Environments

Motivation

- Biological systems develop desired properties
  - Robustness, Flexibility, Modularity, Evolvability, ...
  - Properties are connected
- Well studied, but their emergence is less well understood
  - Investigate the evolution of metabolic networks
  - Analyse network structure and metabolic functions
- Answers beyond analyzing real-world data

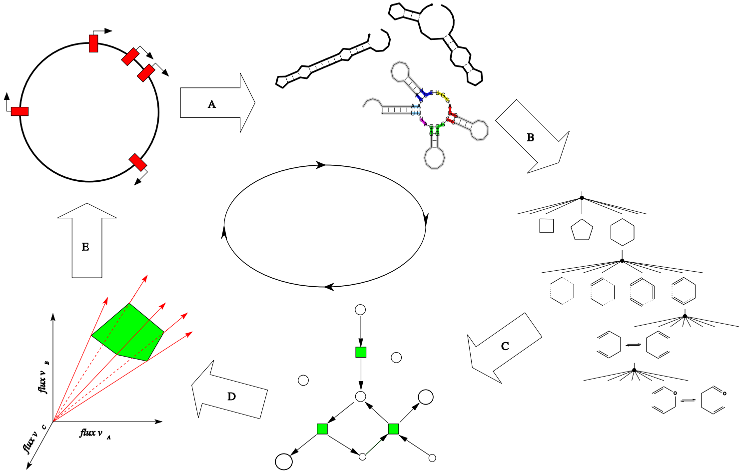
$\rightarrow$ a multi-scale computational model of early metabolism
$\rightarrow$ appropriate measures for network properties

Simulation

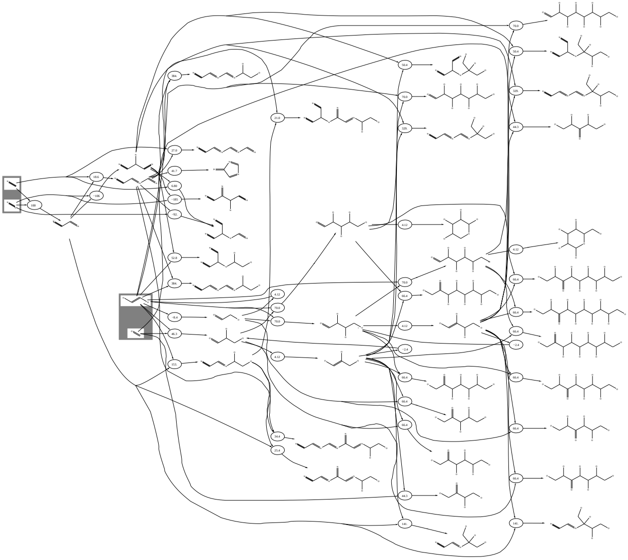


- Protocellular entity
- Bag of ribozymes
- Algebraic chemistry model
- Exchange of molecules with the environment

Simulation - Overview

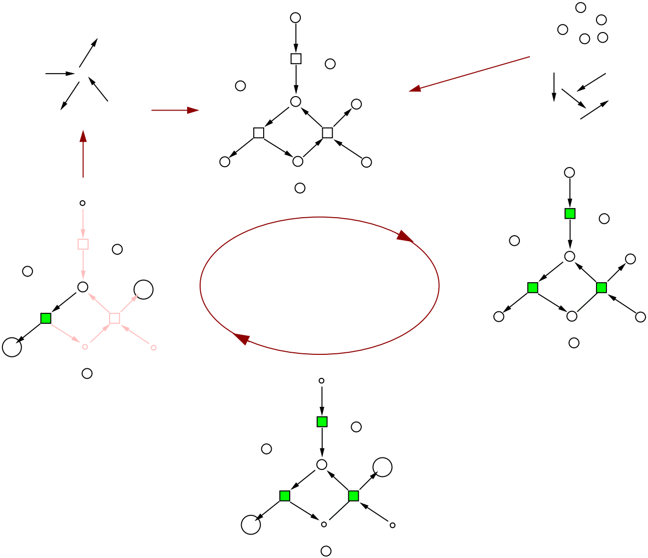


Simulation - Growth



cyanide, formaldehyde glycol; aldolcondensation, tautomerization

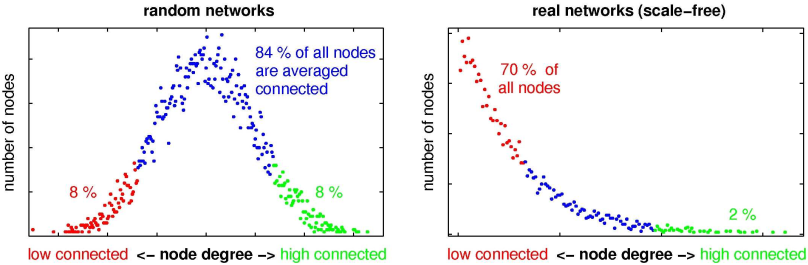
Simulation - Stochastic Network Generator



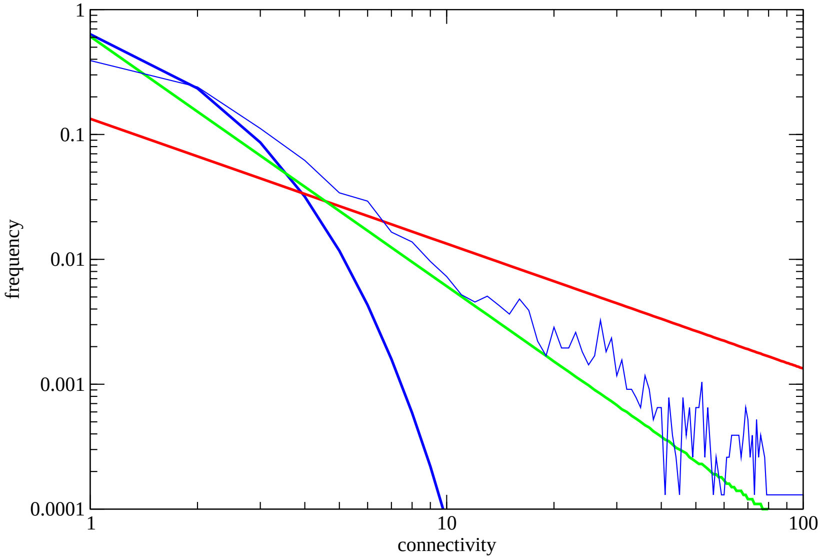
Faulon, J-L, (2001) J Chem Inf Comput Sci 41:894-908

General network analysis

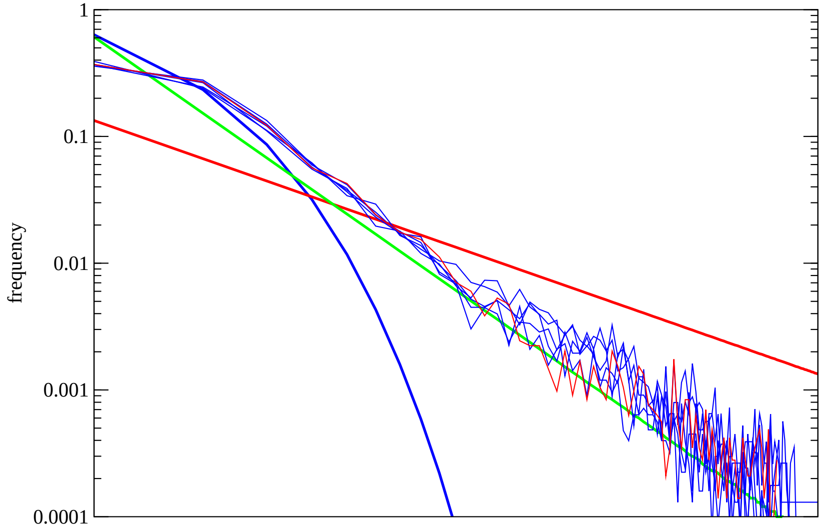
- Connectivity Distribution
  - small vs big
  - early vs evolved
- Clustering Coefficient, Centrality, Entropy, ...
  - simulated vs real world



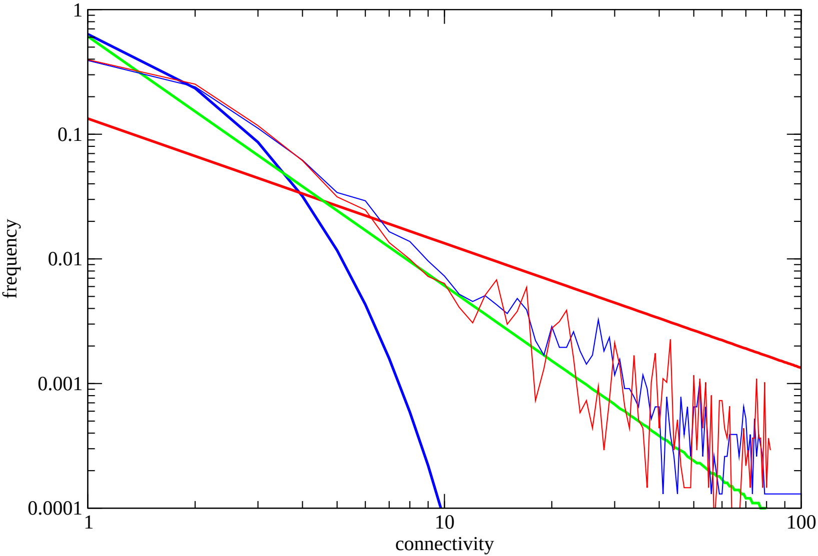
after 10 Generations



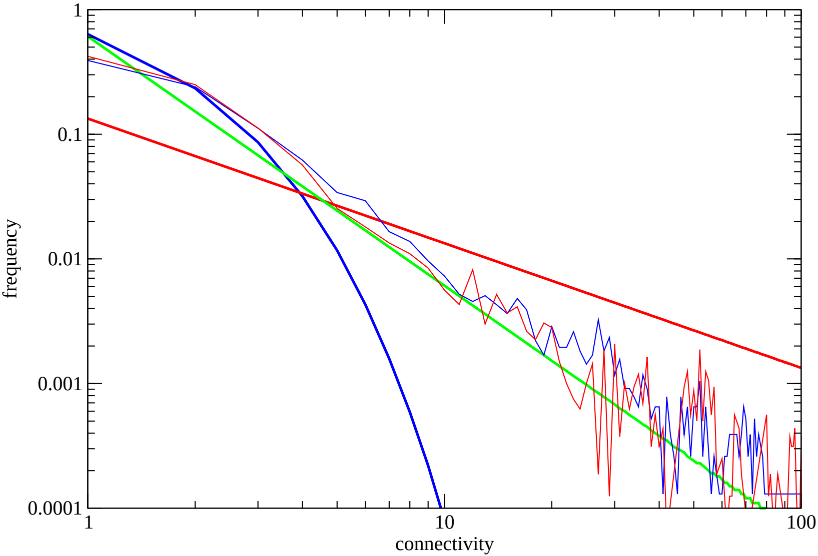
after 50 Generations



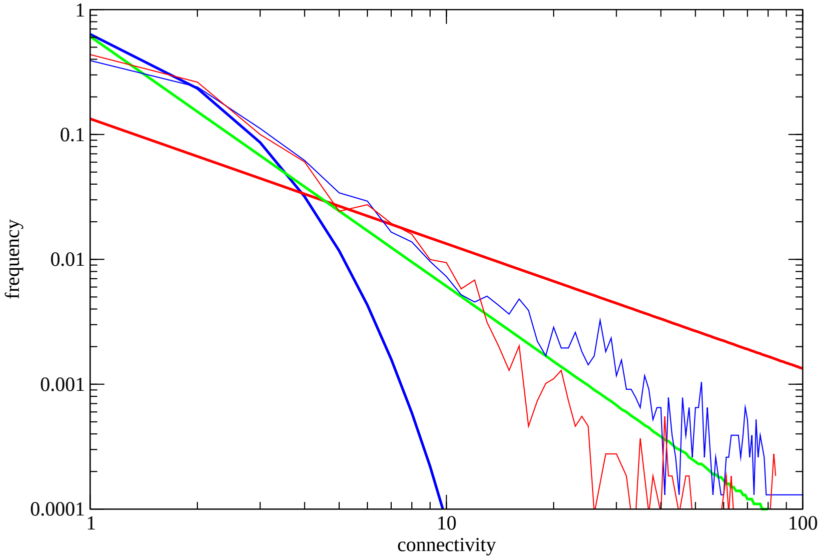
after 100 generations



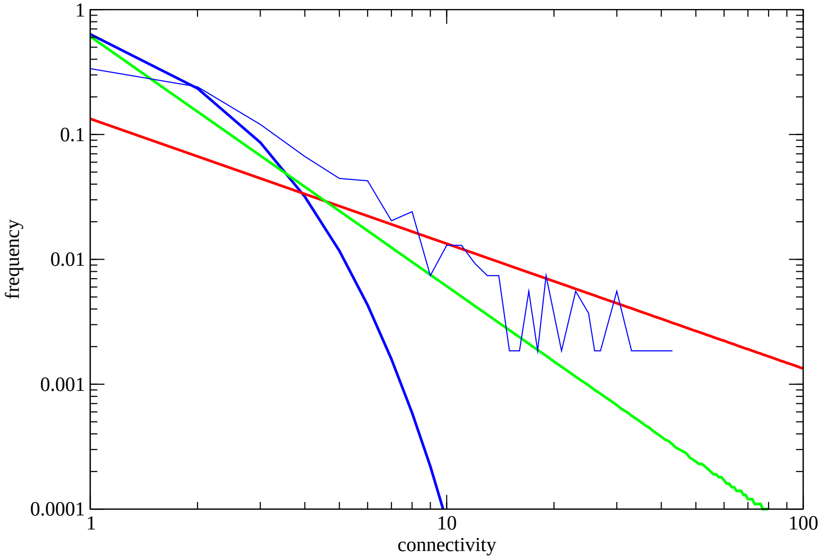
after 250 generations



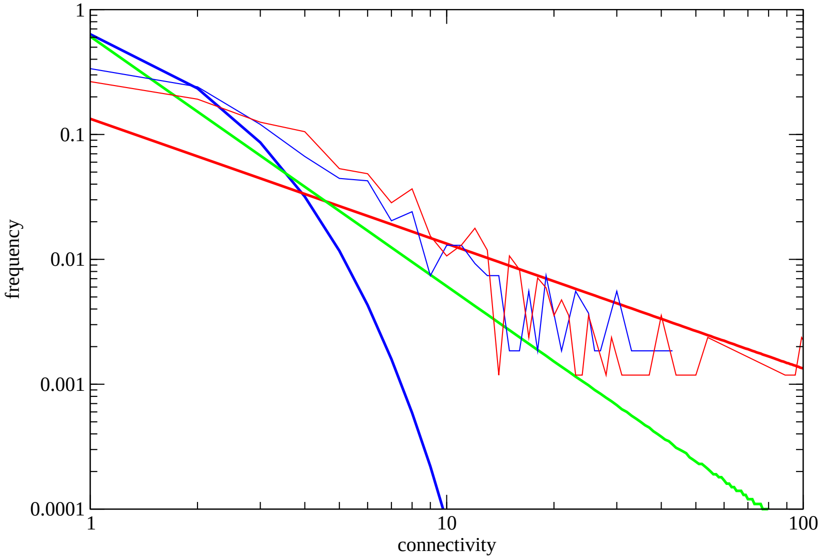
after 500 generations



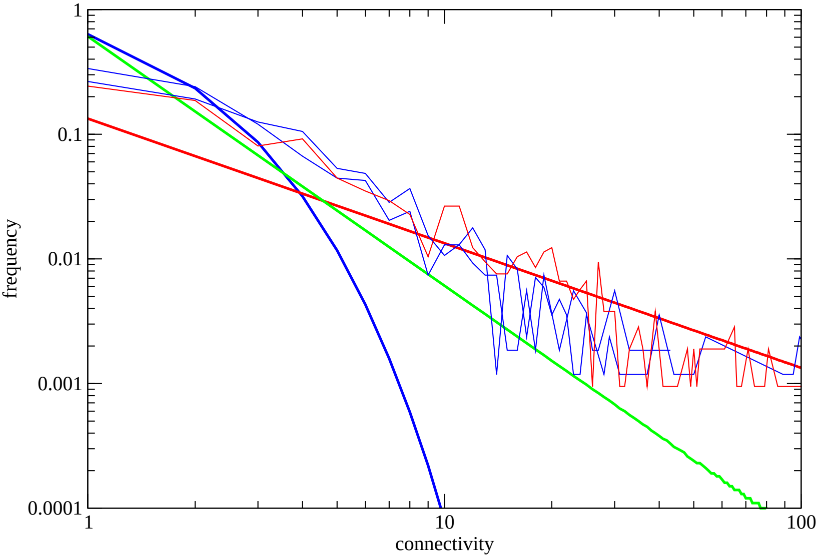
Changing Environment - after 100 generations



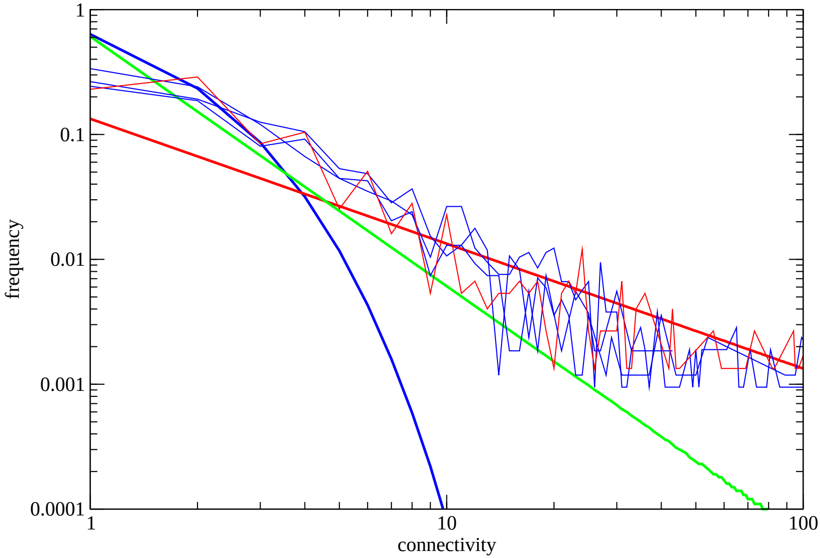
after 250 generations



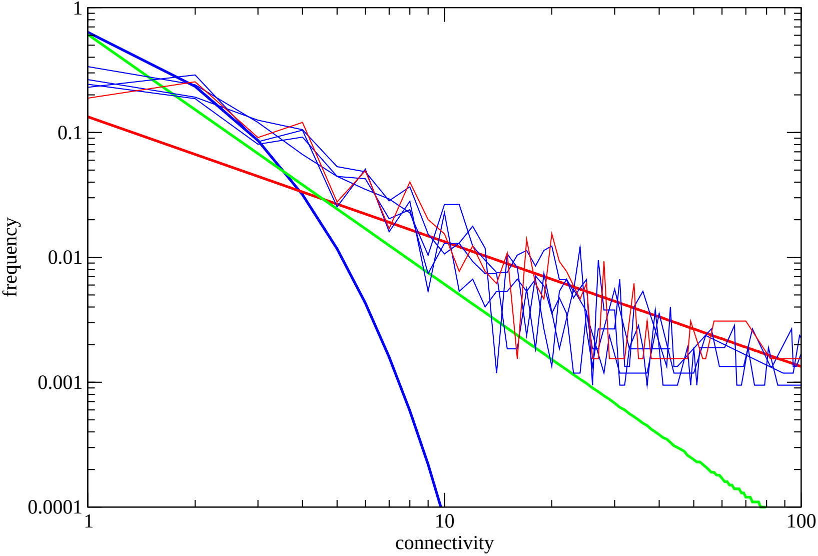
after 500 generations



after 750 generations



after 1000 generations



Metabolic network analysis

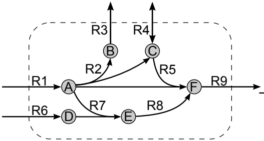
We have sets of edges forming meaningful complex entities

$\downarrow$

pathways

- number of pathways $\rightarrow$ flexibilty
- change in case of single/multiple knockouts $\rightarrow$ robustness
- number of acceptable knockouts $\rightarrow$ robustness

Metabolic Pathway Analysis



$\longrightarrow$

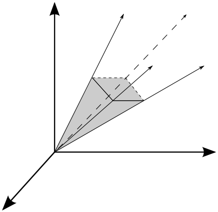
	$R1$	$R2$	$R3$	$R4$	$R5$	$R6$	$R7$	$R8$	$R9$
A	1	−1	0	0	−1	0	−1	0	0
B	0	1	−1	0	0	0	0	0	0
C	0	1	0	1	−1	0	0	0	0
D	0	0	0	0	0	1	−1	0	0
E	0	0	0	0	0	0	1	−1	0
F	0	0	0	0	1	0	0	1	−1

$\downarrow$

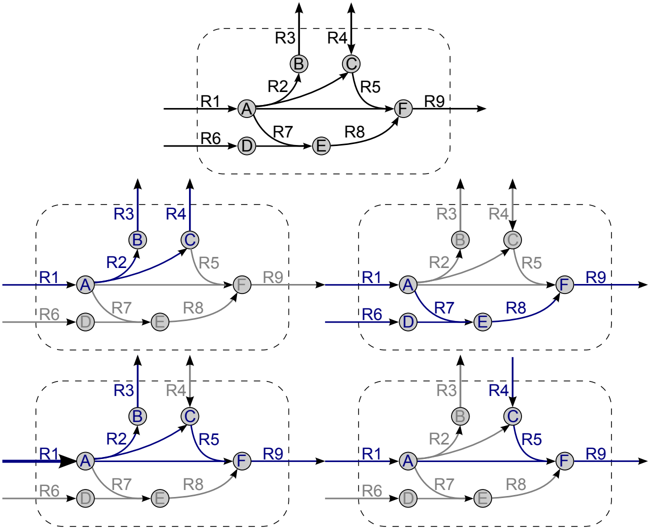
$$S * v_i = 0$$

	v_1	v_2	...	v_{n-1}	v_n
$R1$	2	−1	...	0	0
$R2$	1	−1	...	−1	−1
$R3$	1	−1	...	−1	−1
$R4$	0	1	...	1	0
$R5$	1	0	...	0	−1
$R6$	0	0	...	1	2
$R7$	0	0	...	1	2
$R8$	0	0	...	1	2
$R9$	1	0	...	1	1

$\longleftarrow$



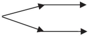
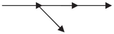
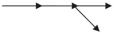
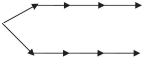
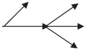
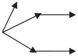
Metabolic Pathway Analysis



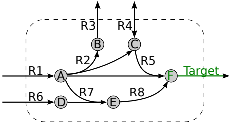
Knockout effects

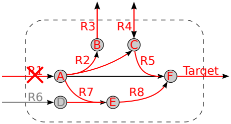
single $R_1 = \frac{\sum_{i=1}^{r} z^i}{r \ * \ z}$ depletion $R_2 = \frac{\sum_{i=1}^{n} R_1^i}{n}$

multiple $R_3(k) = \frac{\sum_{i=1}^{s(k)} z^i}{s(k) \ * \ z}$ overall $R_3(\leq K) = \sum_{k=1}^{K} R_3(k) p_k$

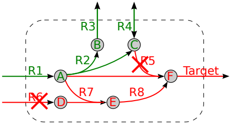
Example system		Number of reactions	Number of elementary modes	$R_1(1)$	$R_1(2)$	$R_1(3)$	$R_1(\leqslant 3)$
1		4	2	$1/2 = 0.5$	$1/6 \approx 0.167$	0	0.414
2		4	2	$1/2 = 0.5$	$1/4 = 0.25$	$1/8 = 0.125$	0.436
3		4	2	$3/8 = 0.375$	$1/12 \approx 0.083$	0	0.302
4		4	2	$1/4 = 0.25$	0	0	0.189
5		8	2	$7/16 \approx 0.438$	$3/8 = 0.375$	$5/16 \approx 0.313$	0.418
6		8	2	$1/2 = 0.5$	$3/14 \approx 0.214$	$1/14 \approx 0.071$	0.416
7		5	4	$13/20 = 0.65$	$3/8 = 0.375$	$7/40 = 0.175$	0.573
8		5	3	$2/3 \approx 0.667$	$2/5 = 0.4$	$1/5 = 0.2$	0.592

Minimal Knockout sets

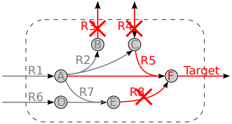




{$R1$}

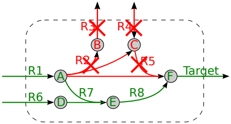


{$R5$, $R6$}

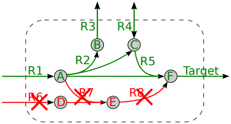


{$R3$, $R4$, $R8$}

NO knockout sets



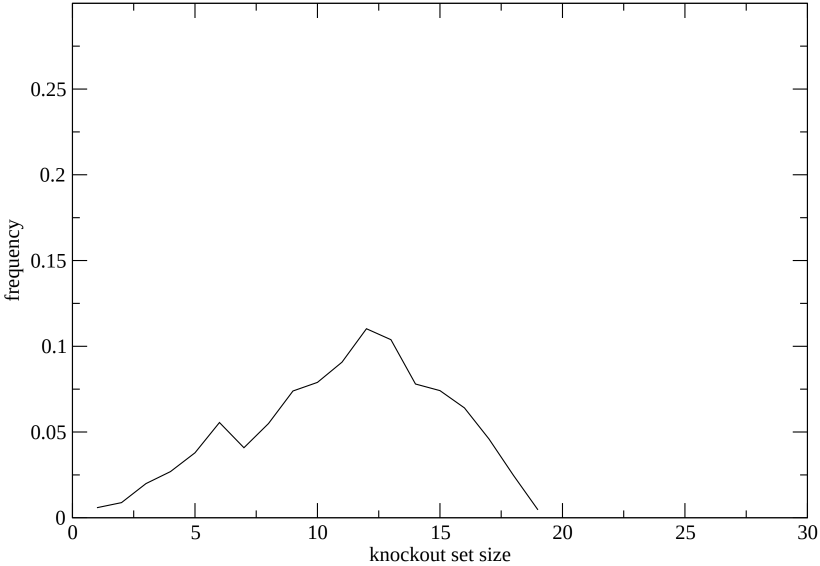
{$R2$, $R3$, $R4$, $R5$}



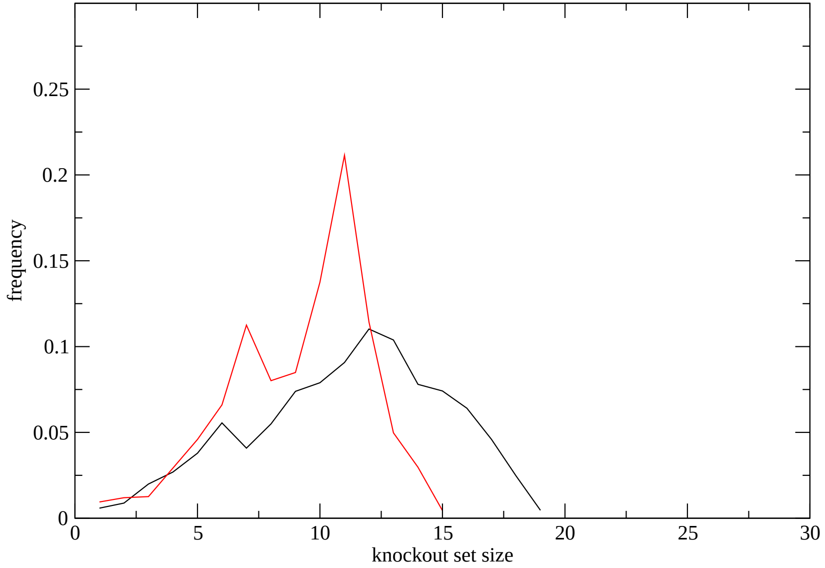
{$R6$, $R7$, $R8$}

Knockout set size distribution → Robustness (bigger is better)

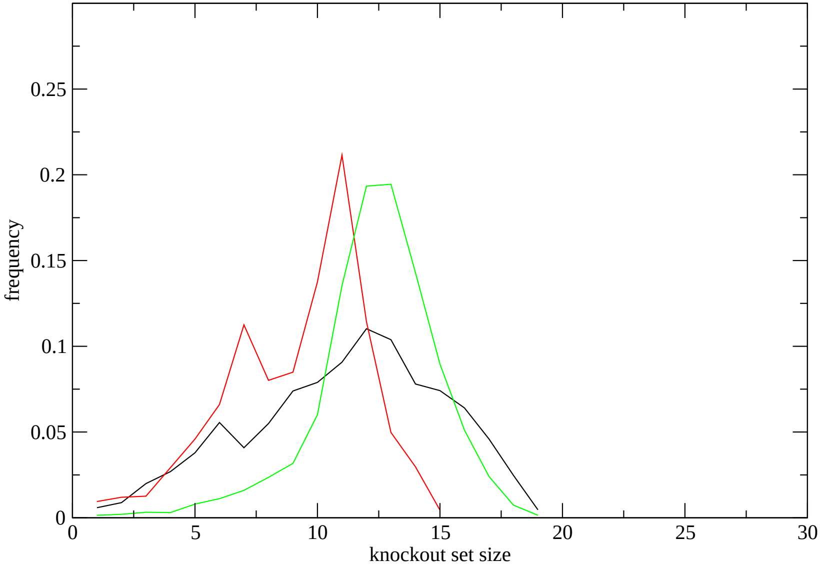
after 10 generations



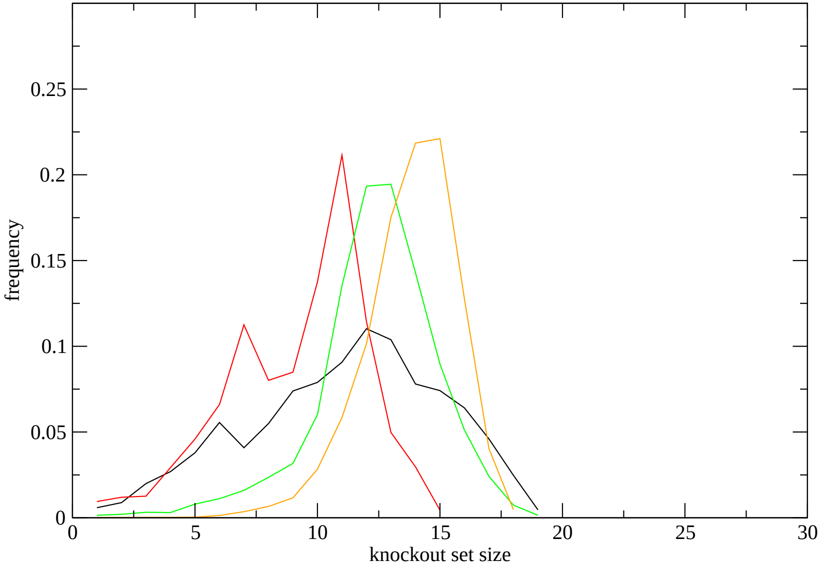
after 20 generations



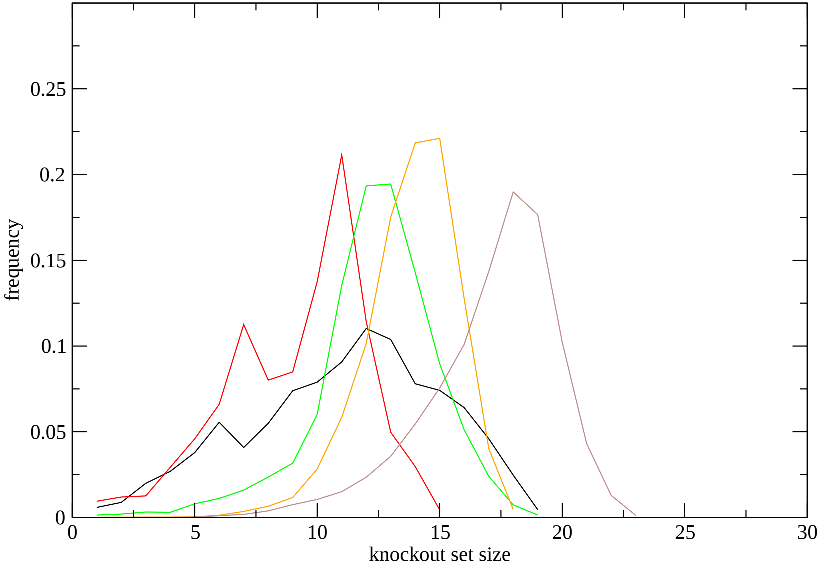
after 50 generations



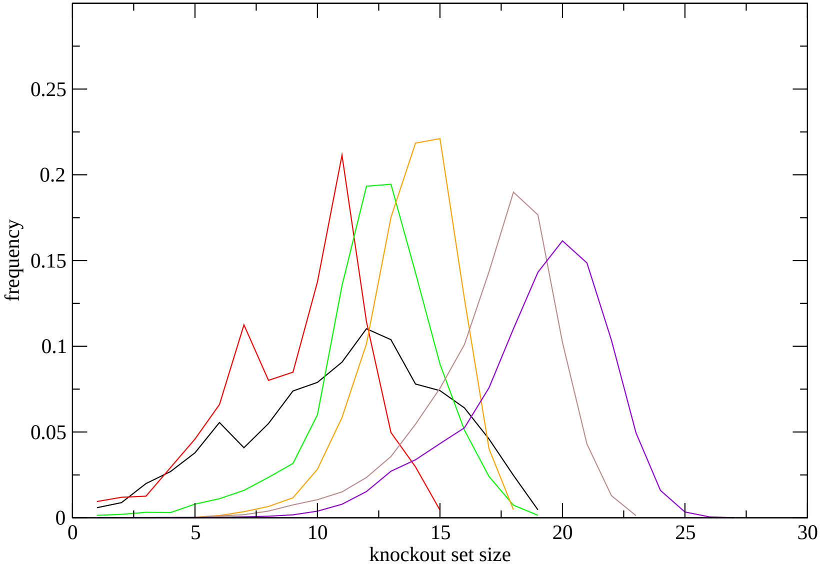
after 100 generations



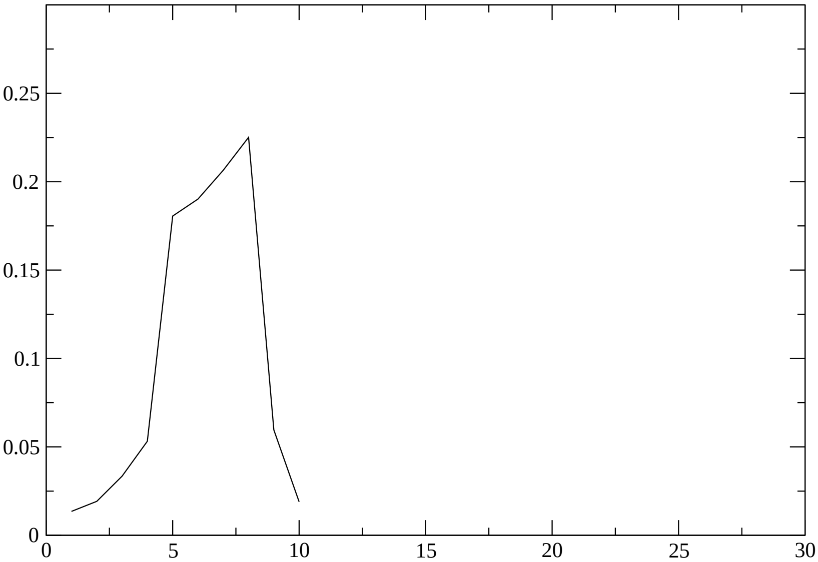
after 250 generations



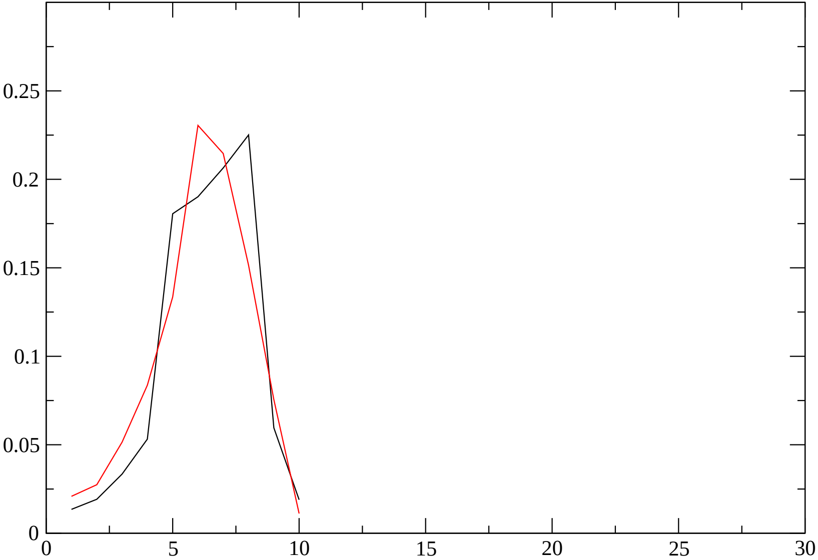
after 500 generations



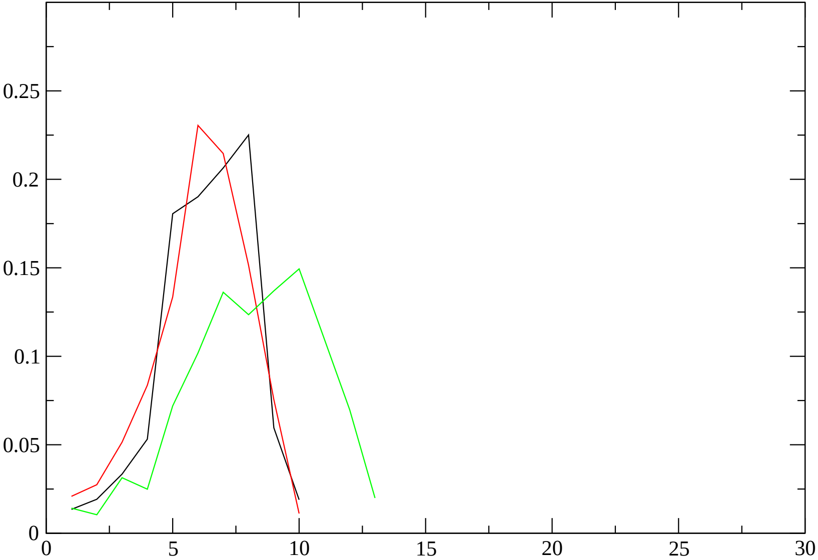
Changing Environment - after 10 generations



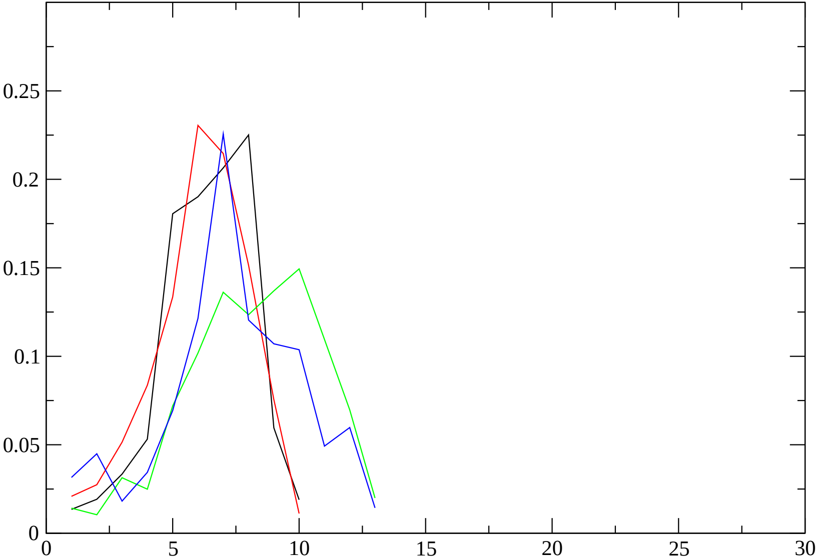
after 50 generations



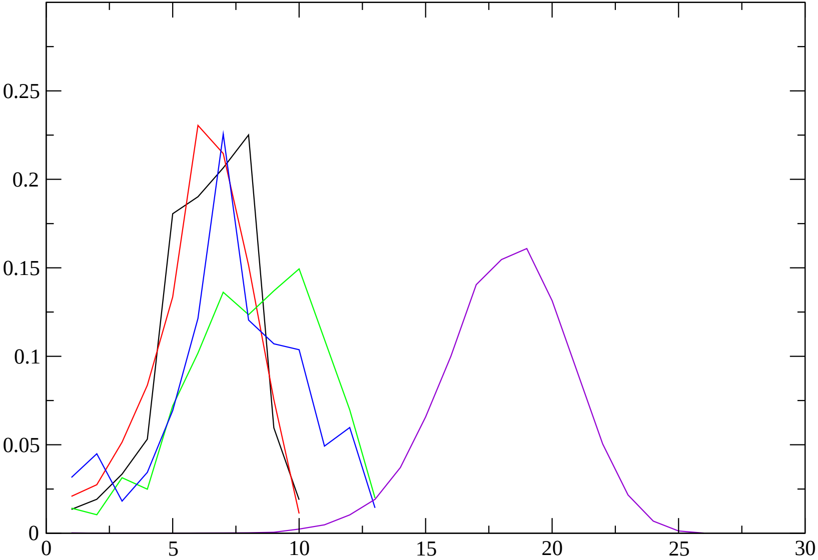
after 250 generations



after 500 generations



after 1000 generations



Chemical Organizations

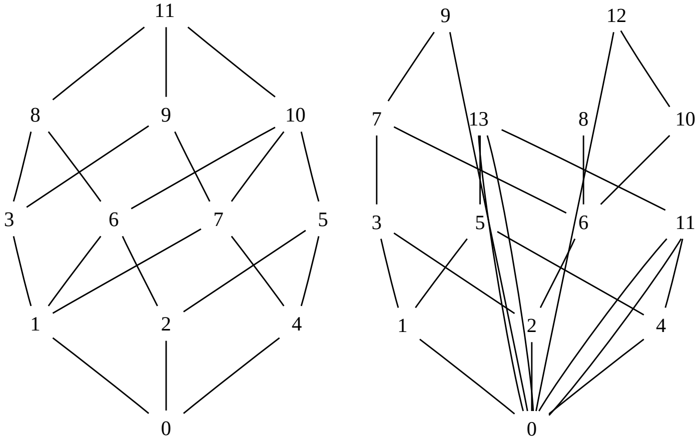
Self-maintaining and closed sets of molecules and reactions

$\downarrow$

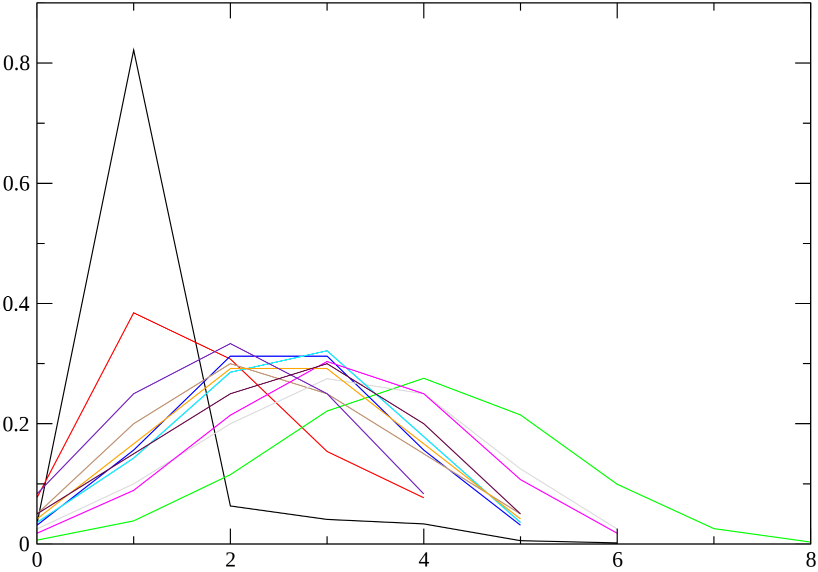
chemical organizations

- Hierarchies of organizations
- Shape of Hierarchies $\rightarrow$ robustness
- Size distribution of organizations $\rightarrow$ robustness, modularity

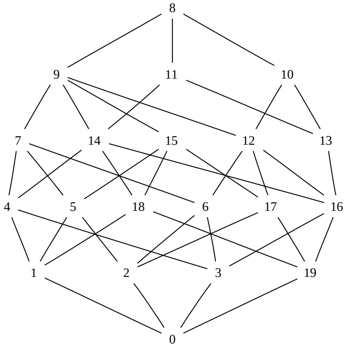
Chemical Organizations

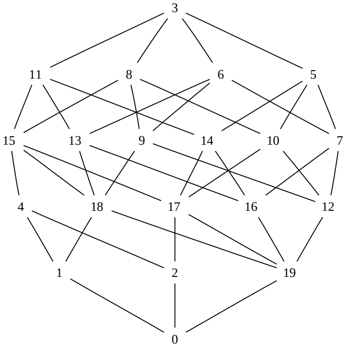


Level Size Distribution - after 500 generations

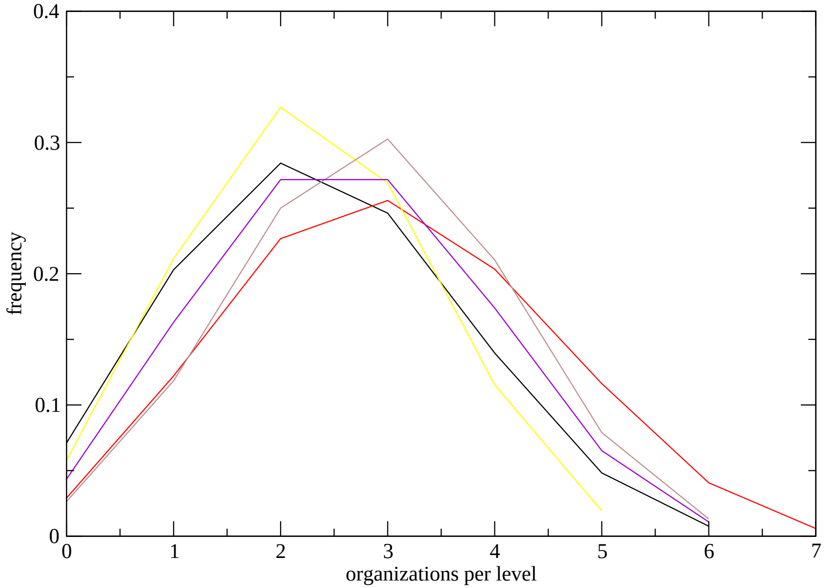


Chemical Organizations

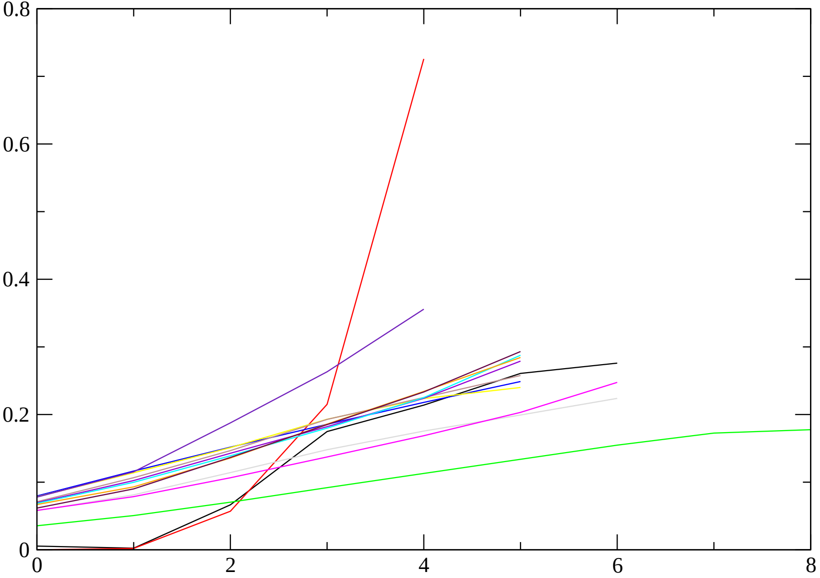




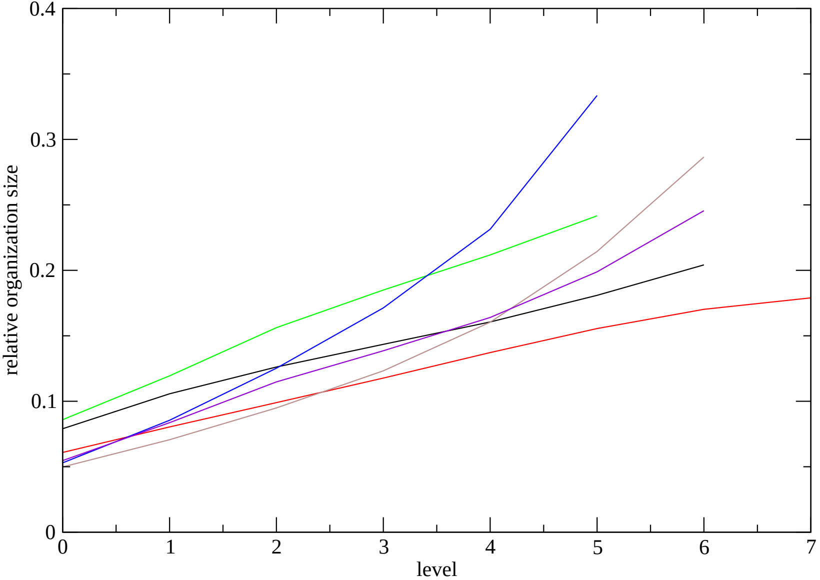
Changing Environment - after 500 generations



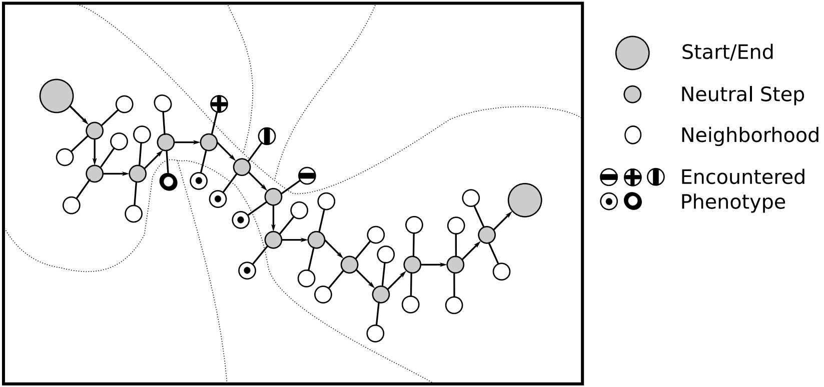
Organization Size Distribution - after 500 generations



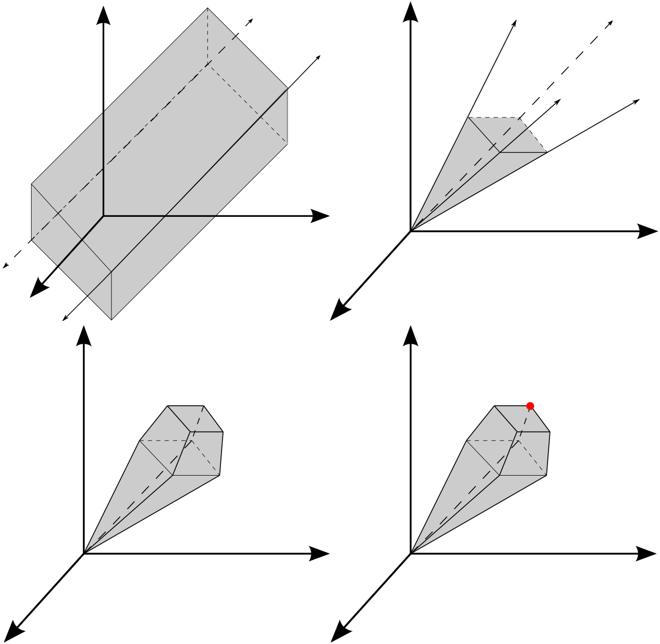
Changing Environment - after 1000 generations



Work in Progress

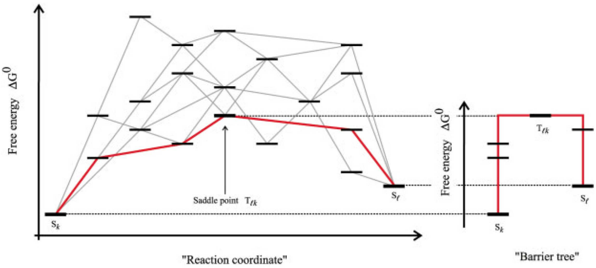


FBA



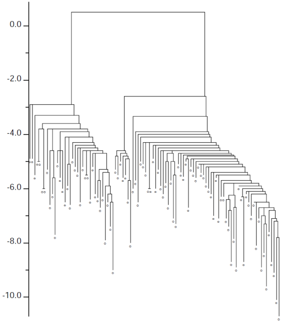
Flux barrier analysis

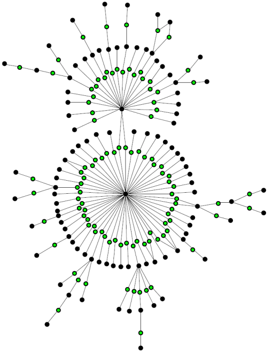
- linear optimization: EMs modeled as system of linear equations
- constraints: limits on reactions, exclusion of combinations of EMs
- barrier tree



Reaction barrier analysis

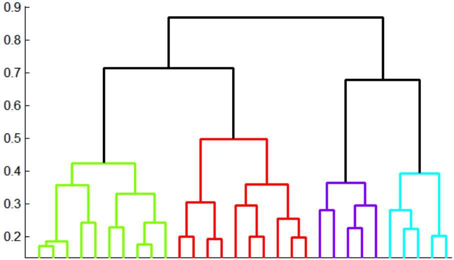
- linear optimization: stoichiometrix matrix
- constraints: limits on reactions, exclusion of combinations of reactions
- barrier tree





Flux similarity

- Compute pairwise similarity of elementary modes
- similarity between metabolites (in+out / all) through topological indices
- similarity between enzymes/reactions by comparing transition state structure



Conclusion

- Summary
  - Structural and Functional measures for Robustness and Modularity
  - Follow the Law (Connectivity Distribution)
  - Size Matters (Knockout set Size)
  - Shape too (Chemical organization Hierarchy)
- Outlook
  - Investigate single networks (flux barriers, flux similarity)
  - Different scenarios (Horizontal Gene Transfer)
  - Structural modularity (Clustering Coefficient)

Acknowledgements

Christoph Flamm

Peter Stadler
Konstantin Klemm
Martin Mann
Markus Rohrschneider
Peter Dittrich
Dennis Goerlich