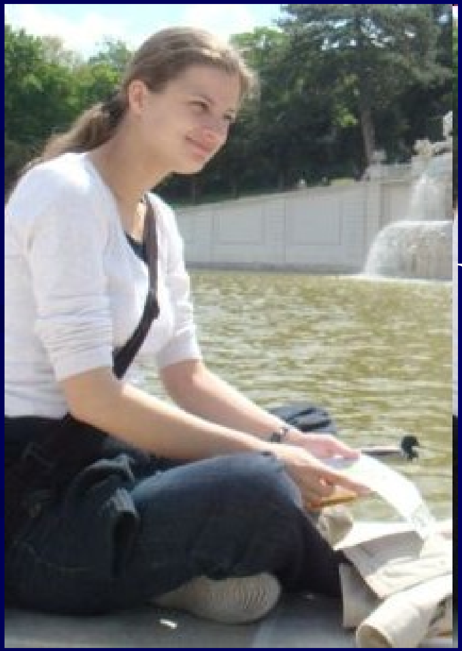
Evolvability of Autocatalytic Reaction Networks

Mauro Santos

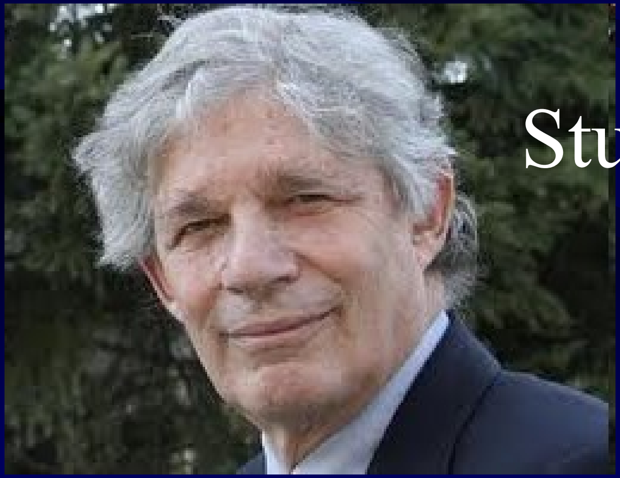
Universitat Autònoma de Barcelona

Collegium Budapest

Chrisantha Fernando



Vera Vasas



Stuart Kauffman

Eörs Szathmáry

Darwinian Principles of Evolution by Natural Selection

- Different individuals in a **population** have different morphologies, physiologies, and behaviours (phenotypic variation).
- Different phenotypes have different rates of survival and reproduction in different environments (differential fitness).
- There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable).

Darwinian Principles of Evolution by Natural Selection

- Important generality in the principles: No particular mechanism of inheritance is specified, but only a correlation in fitness between parent and offspring.
- Natural selection requires reproduction and heredity: there must be some sort of autocatalysis.

Basic Model of Population Genetics

Genotype:	A	a
Frequency:	p_t	$q_t = 1 - p_t$
Fitness:	w_{A}	w_{a}

Change in frequency: $$p_{t+1} = \frac{p_t w_A}{p_t w_A + q_t w_a}$$

Basic Model of Population Genetics

This simply haploid model already assumes a lot:

- The entities (genes) A, a are autocatalytic.
- There is a genotype – phenotype mapping.
- Fitness acts on the ‘phenotypic’ output (reproductive success), and changes in frequencies are mediated by the correlation between genotype and phenotype.

Questions

In Oparin's view coacervates were the first units upon which natural selection could be said to operate. Even though individual molecules in solution may not have been autocatalytic, there may have been selection among variants of a given molecular species when incorporated within a coacervate, so that the coacervate itself would evolve.

- Was Oparin right? Can autocatalytic sets be considered units of evolution in the Darwinian sense?
- Is some sort of ‘genotype – phenotype’ mapping needed?

Autocatalytic Sets: Some Essentials

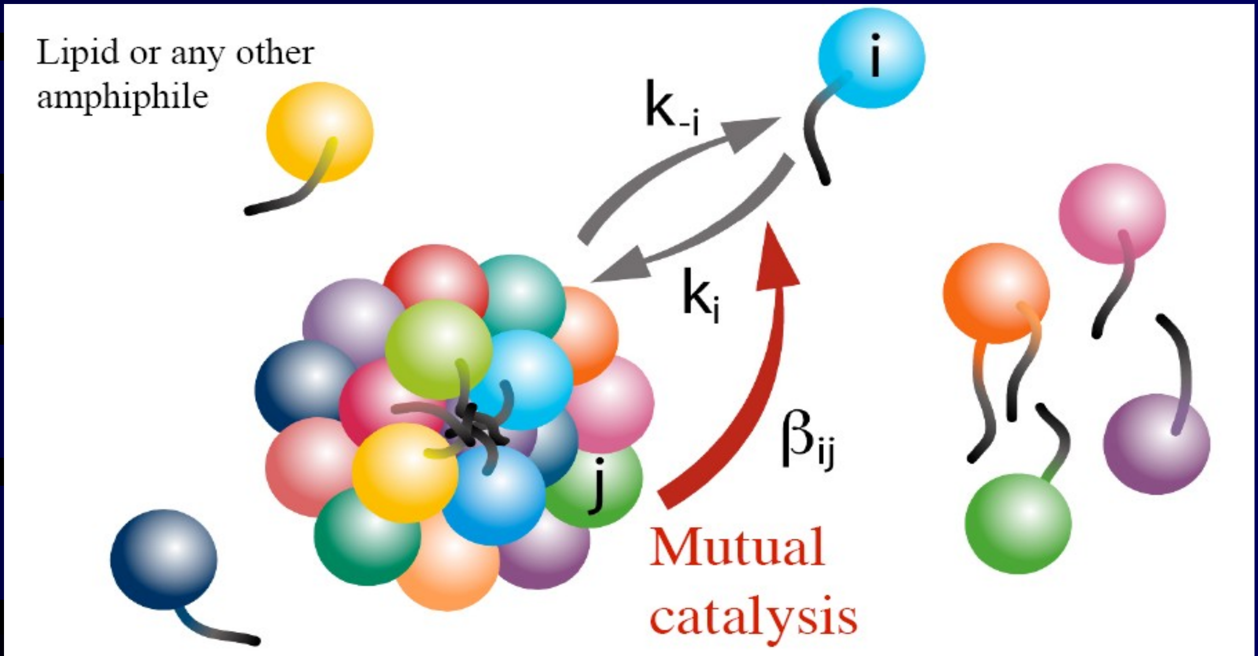
- The essential feature of so-called autocatalytic sets is that they can collectively self-replicate (grow exponentially), even if none of the molecular species can individually self-replicate.

Autocatalytic Sets without Genotype – Phenotype Mapping

- Graded Autocatalysis Replication Domain: GARD model (Segré D, Ben-Eli D, Lancet D 2000. Compositional genomes: Prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Natl Acad Sci USA 97:4112–4117.)

- Autocatalytic sets of polymers (Kauffman SA 1986. Autocatalytic sets of proteins. *J Theor Biol* 119:1–24.)

GARD Model



N_G = number of different kinds

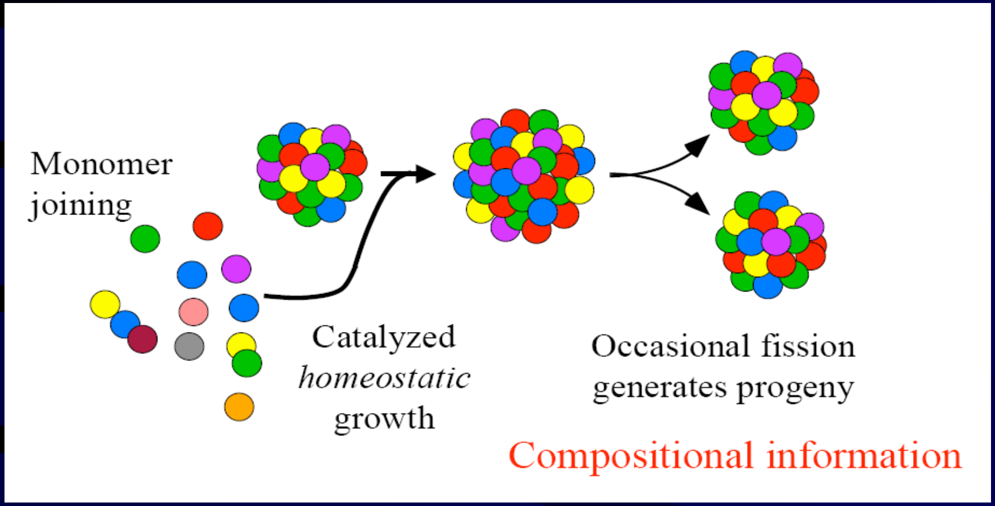
n_i = count of molecules of kind i in the assembly

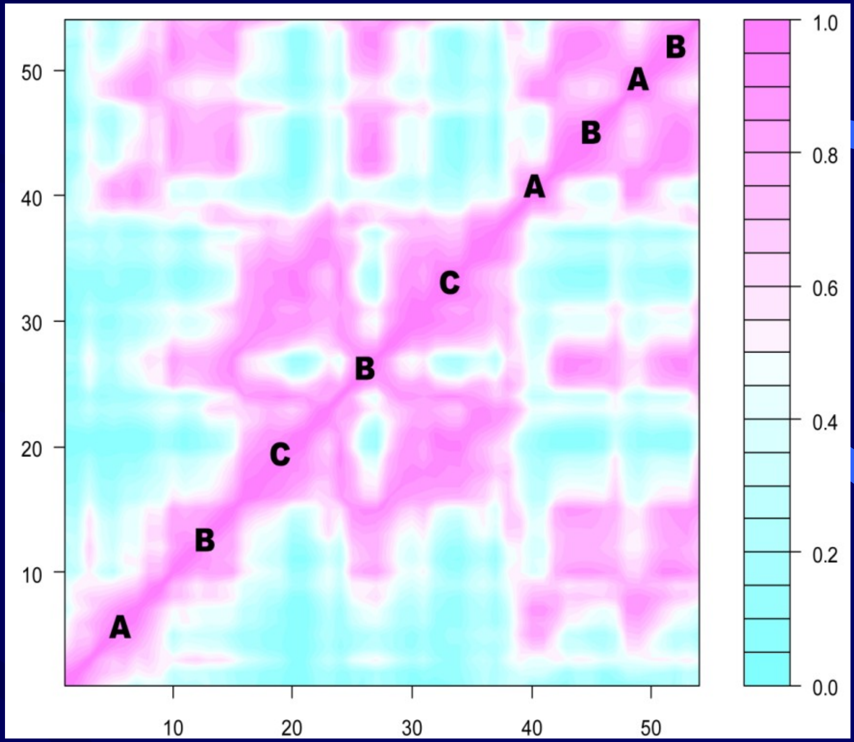
$N = \sum_{i=1}^{i=N_G} n_i$ (assembly size)

$$\frac{dn_i}{dt} = \left(\rho_i k_i N - k_{-i} n_i\right)\left(1 + \frac{1}{N}\sum_{j=1}^{j=N_G} \beta_{ij} n_j\right)$$

Forward reaction — Backward reaction — Catalytic enhancement

GARD Model





'Compositional correlation carpet'

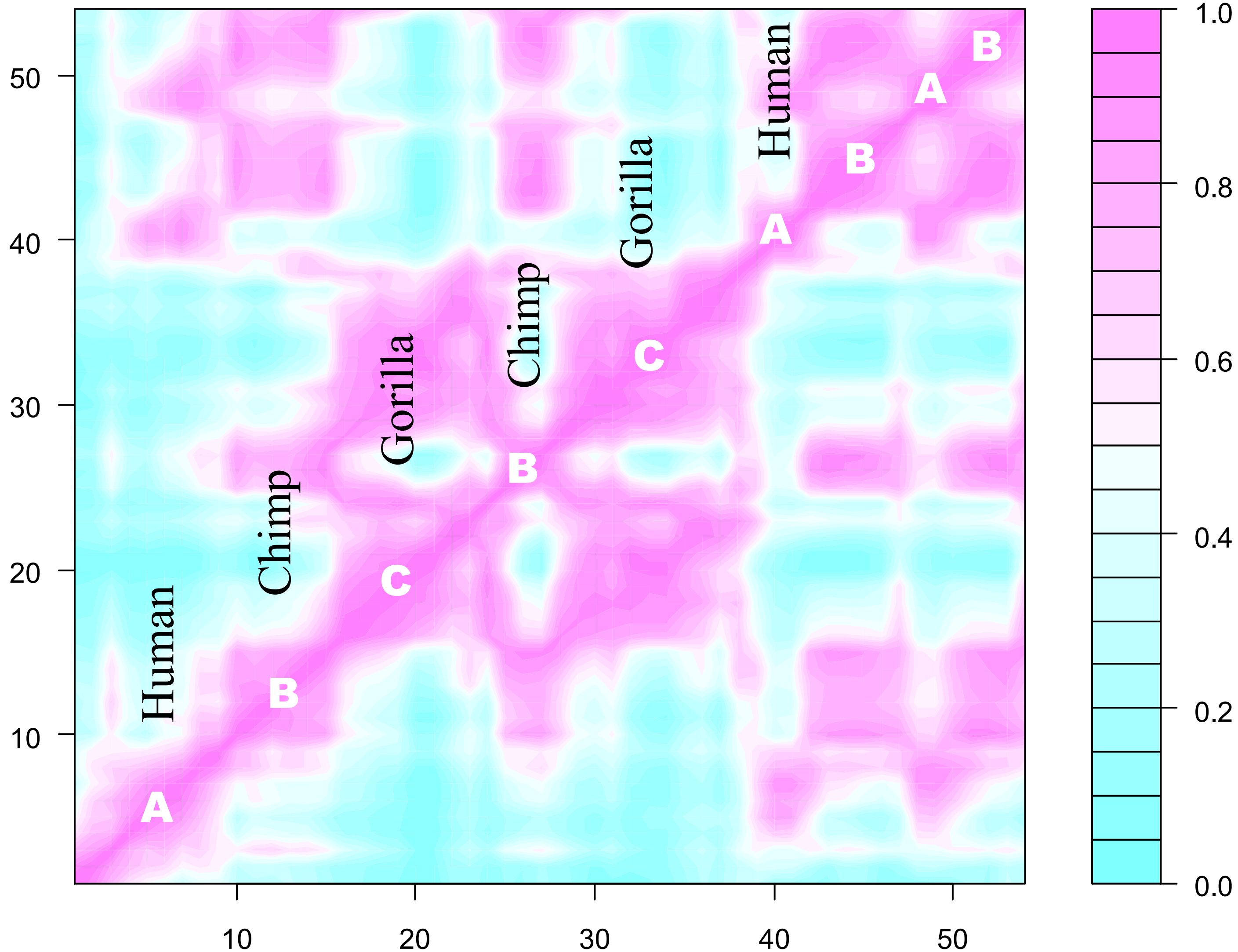
Claims (GARD Model)

- Computed behavior may constitute a demonstration of natural selection in populations of molecules without genetic apparatus (Segré et al. PNAS 2000).

Darwinian Principles of Evolution by Natural Selection

- Different individuals in a population have different morphologies, physiologies, and behaviours (phenotypic variation).
- Different phenotypes have different rates of survival and reproduction in different environments (differential fitness).
- There is a correlation between parents and offspring in the contribution of each to future generations (fitness is heritable).

Time correlation matrix of H for 54 generations.
A, B and C mark the three different composomes.



Scala Naturae according to GARD

An Eigen equation for compositional assemblies

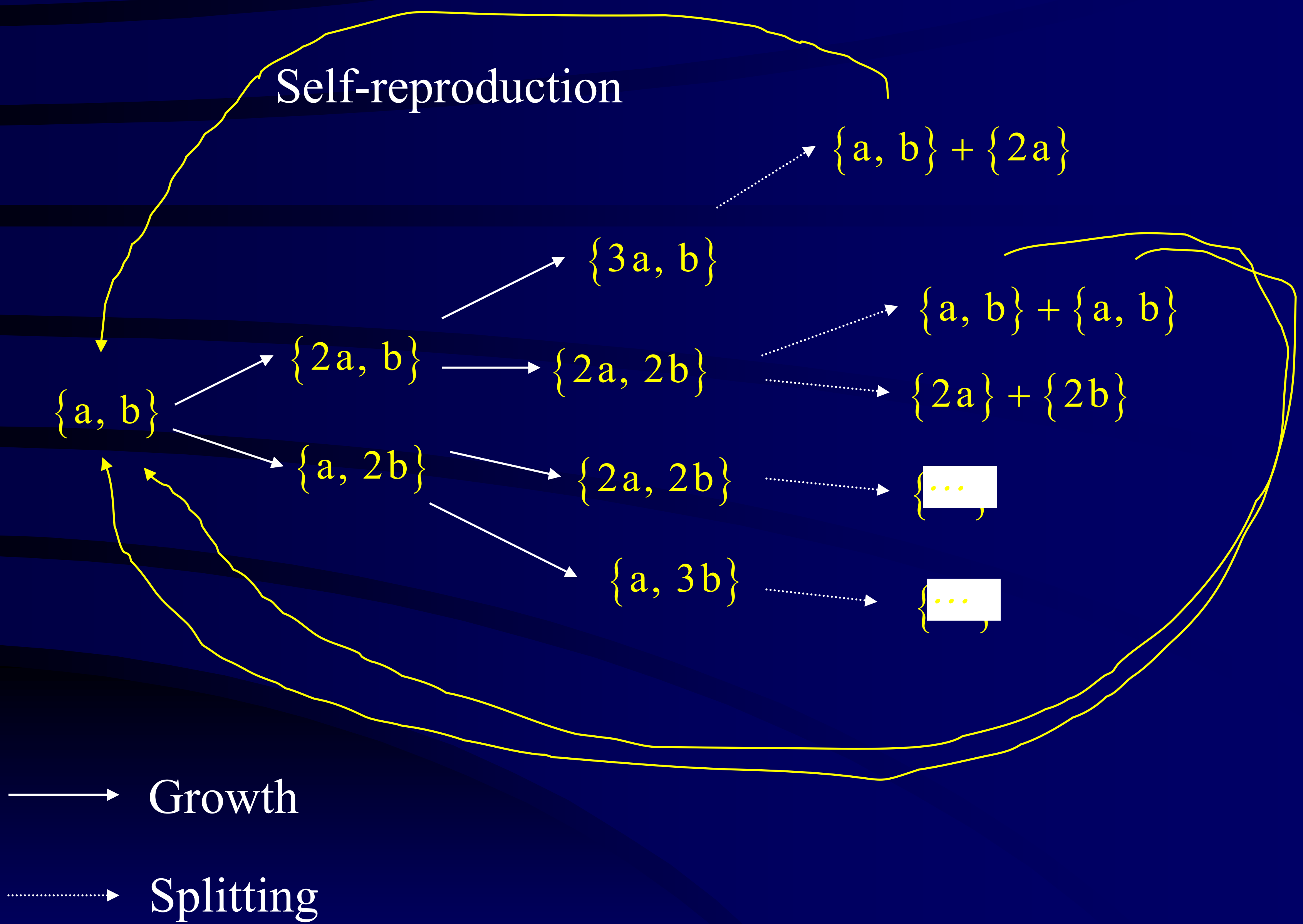
Density of assembly *k*

outflow

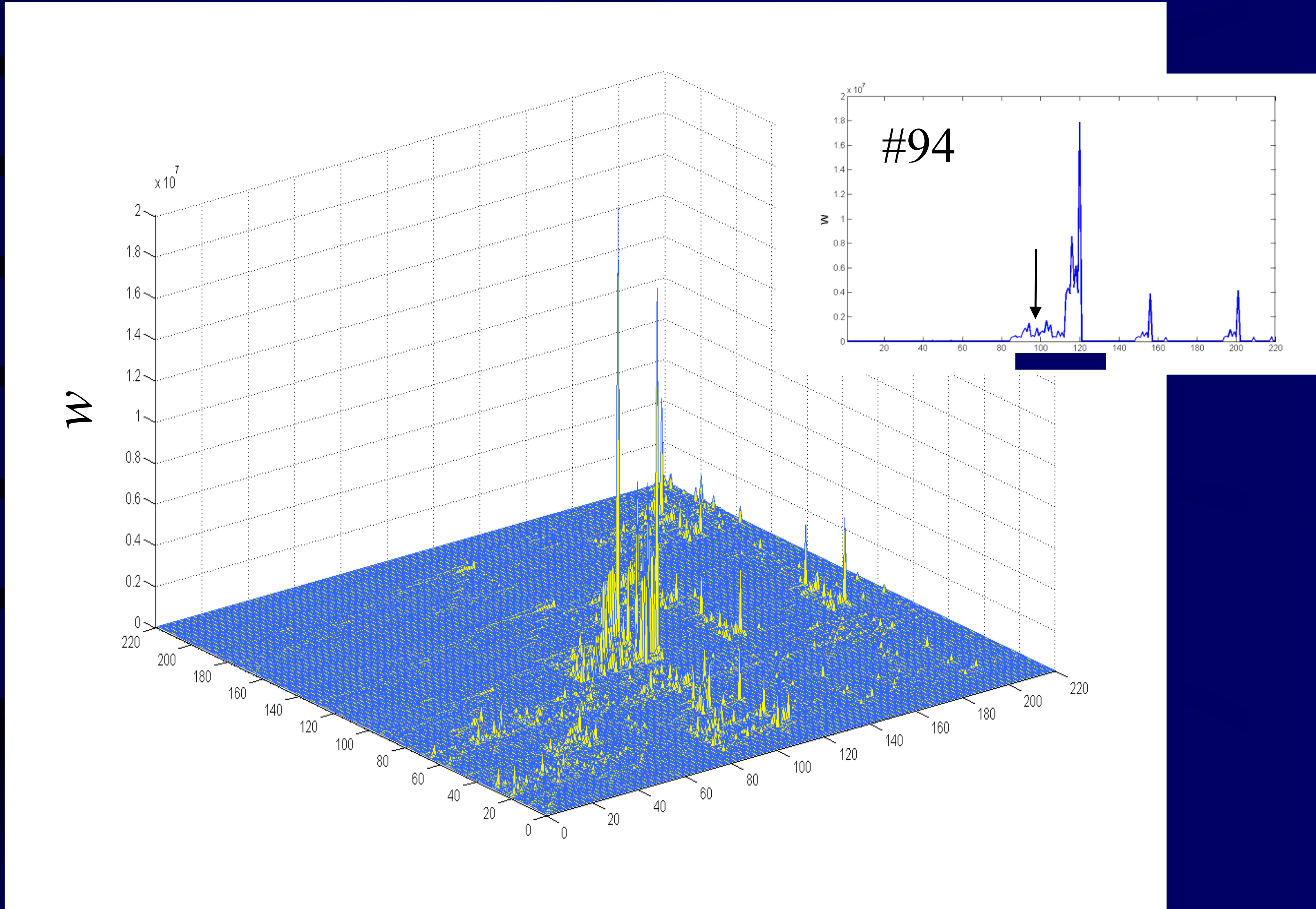
$$X_k' = (r_k - E)X_k + \sum_{l=1}^{\Omega} \mu_{kl} X_l$$

Self-reproduction

Mutation rate from *l* to *k*



Self-reproduction
{a, b}
{2a, b}
{a, 2b}
{3a, b}
{2a, 2b}
{2a, 2b}
{a, 3b}
{a, b} + {2a}
{a, b} + {a, b}
{2a} + {2b}
Growth
Splitting



An exact solution for the replication-mutation equilibrium distribution

In the classical, sequence-based model the mutation matrix has a definite structure imposed by sequence space in the sense that for any mutation rate $u < 1$ the one-error mutants are more frequent than the two-error mutants, and so forth.

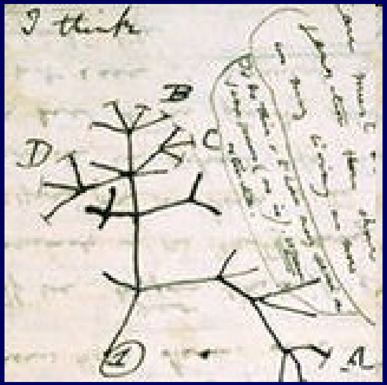
This structure (which is missed in GARD) is fundamental for evolution (phylogeny).

Remember

Darwinian evolution has two ingredients:

- Descent with modification

- Natural selection



Darwin's vision of descent with modification

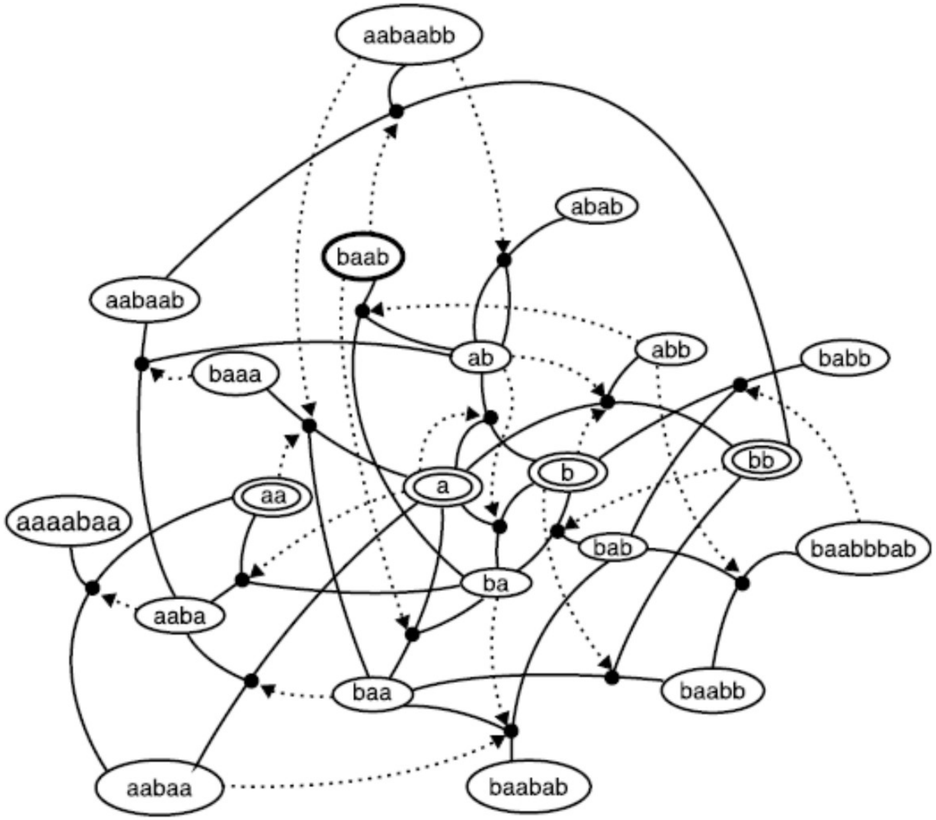
$$\frac{dn_i}{dt} = \left(\rho_i k_i N - k_{-i} n_i\right)\left(1 + \frac{1}{N}\sum_{j=1}^{j=N_G} \beta_{ij} n_j\right)$$

Forward reaction — Backward reaction — Catalytic enhancement

Dynamics is dominated by the beta matrix.

Large-scale outcomes that arise in the simulations result from hidden, small-scale processes.

Kauffman’ model: polymer chemistry



- Catalysed ligation/cleavage reactions of peptides
- *P*: probability that a given molecule catalyses a given reaction
- Food set

Kauffman 1986. *JTB* 119:1-24.

Kauffman’s original paper describes **autocatalytic sets** in relation to the food set: reflexively autocatalytic and food generated (F-generated) **RAF**

Kauffman's Polymer Chemistry

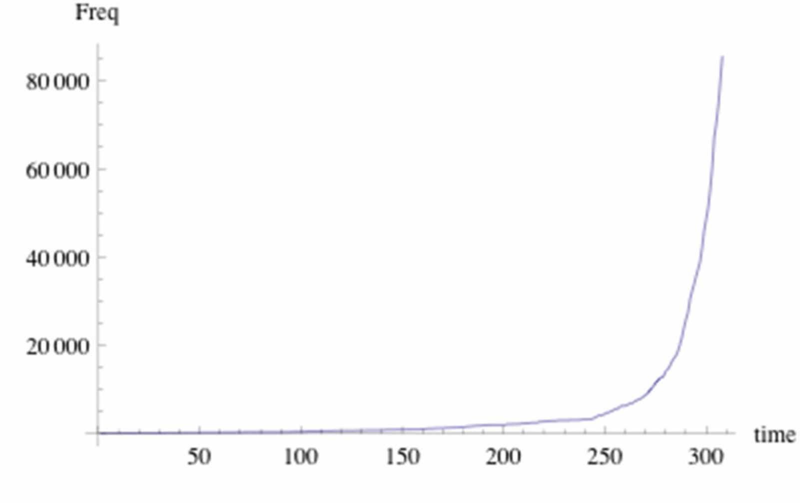
- There exists a large 'food set' of abundant polymers naturally formed in the environment, up to some level of complexity (up to length *L*).
- Each molecule has certain probability *P* of catalyzing each ligation-cleavage reaction.
- Kauffman suggested that 'some autocatalytic sets will reproduce more rapidly than others and hence will have higher Darwinian fitness …. we have evolution (*sensu* Darwin) without a genome'.

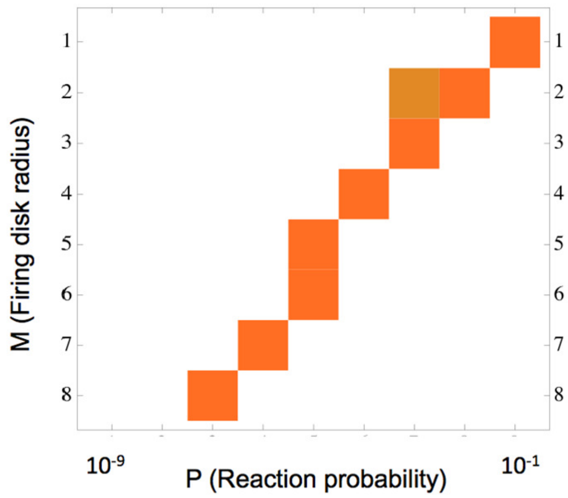
Kauffman's Polymer Chemistry

- However, no rigorous analysis of the putative evolvability of Kauffman's RAF has been carried out so far.
- Previous analysis just 'forgot' that Darwinian evolution is a population concept …what is needed is to simulate a population of compartments enclosing autocatalytic sets.

Reimplementation of Kauffman's model

Mathematical model: supracritical growth





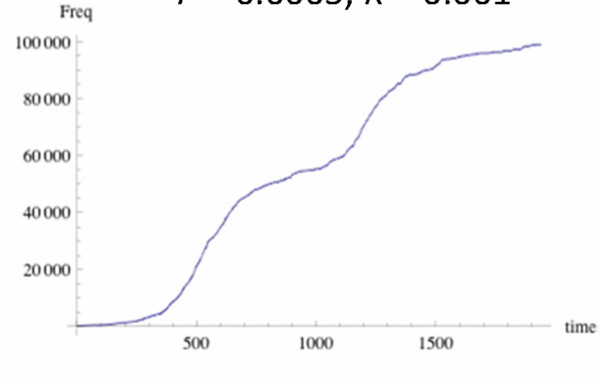
Chemical implementation in a flow reactor: logistic growth

Farmer et al. 1986. *Phisica D* 22:50-67.

Earlier criticisms of Kauffman's model

Lifson: *P* is a composite probability
P'= probability of being a catalyst
P''= probability of catalysing a particular reaction
$\Rightarrow$ with high P' and P'' there *is* supracritical growth

P = 0.0003, *K* = 0.001



Szathmáry: 'paradox of specifity'
K: probability that a given molecule inhibits a given reaction

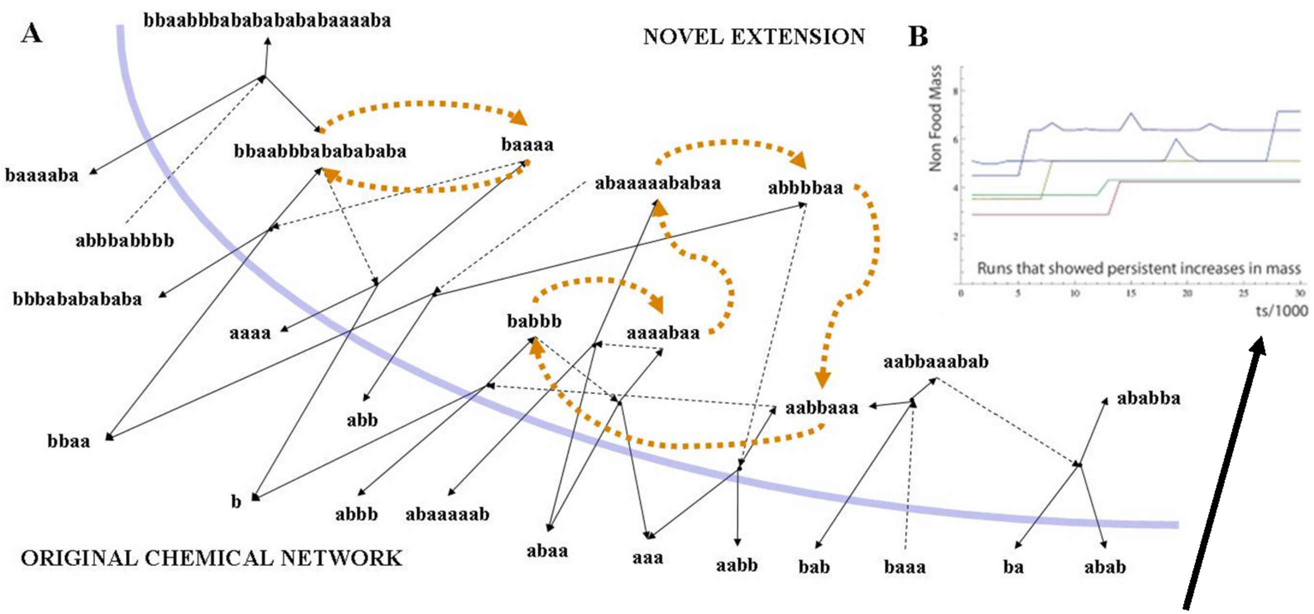
∑: only quantitative issues

Can Kauffman's autocatalytic sets be units of evolution?

- No! The original Farmer-type networks always have one attractor. Kauffman's original polymer chemistry when enclosed in a finite space will eventually crystallize into the same attracting network and can never ever be a Darwinian unit (similar to the GARD model).
- However, the addition of rare novel species can result in the ignition of a novel self-reproducing viable loop. In this case selection could work.

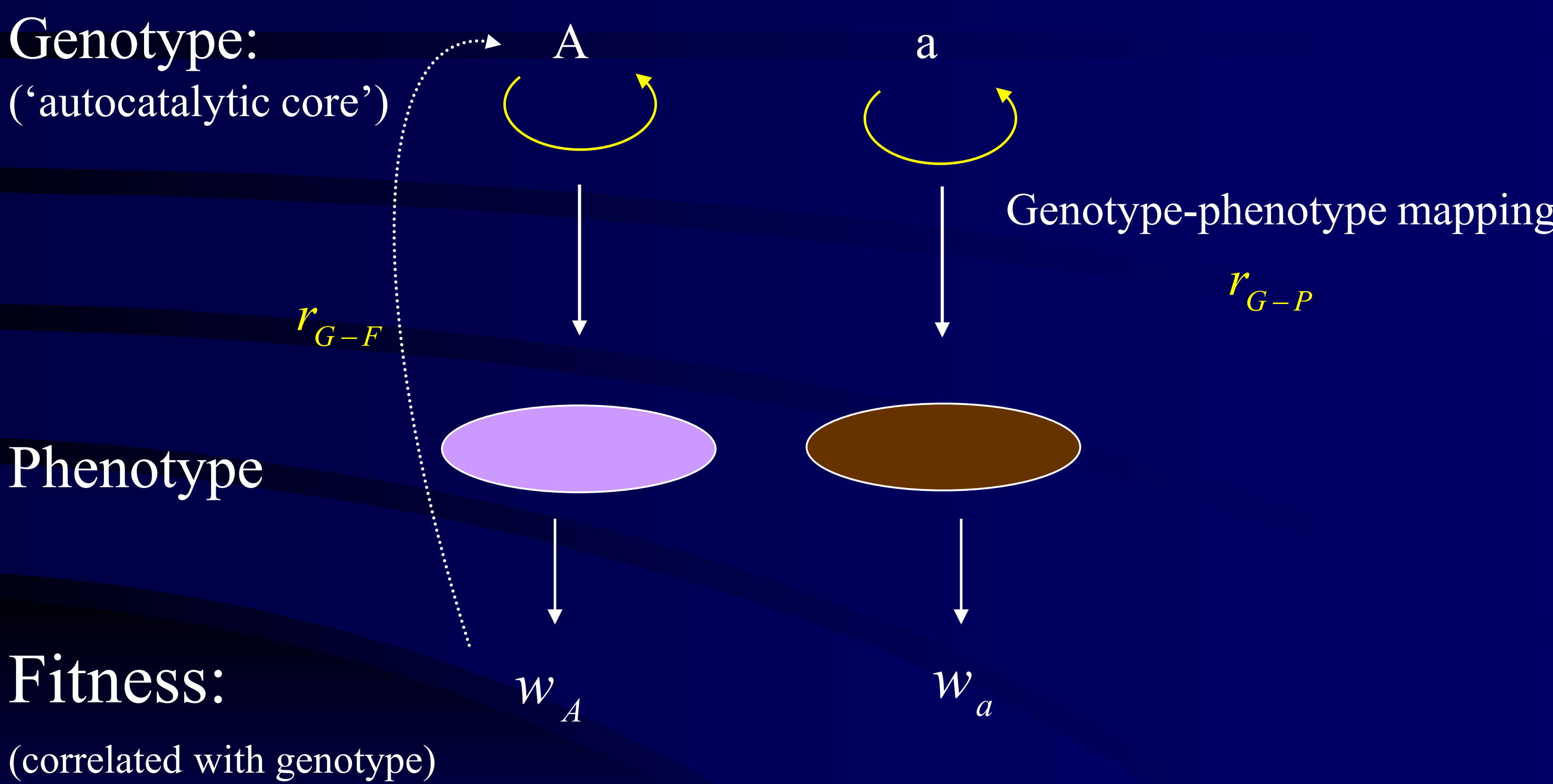
Introducing chemical avalanches

Extra: rare uncatalysed reactions result in novel species
Rarely: non-food generated autocatalytic loops



tendency to increase in mass and complexity

Analogy with the Basic PopG Model

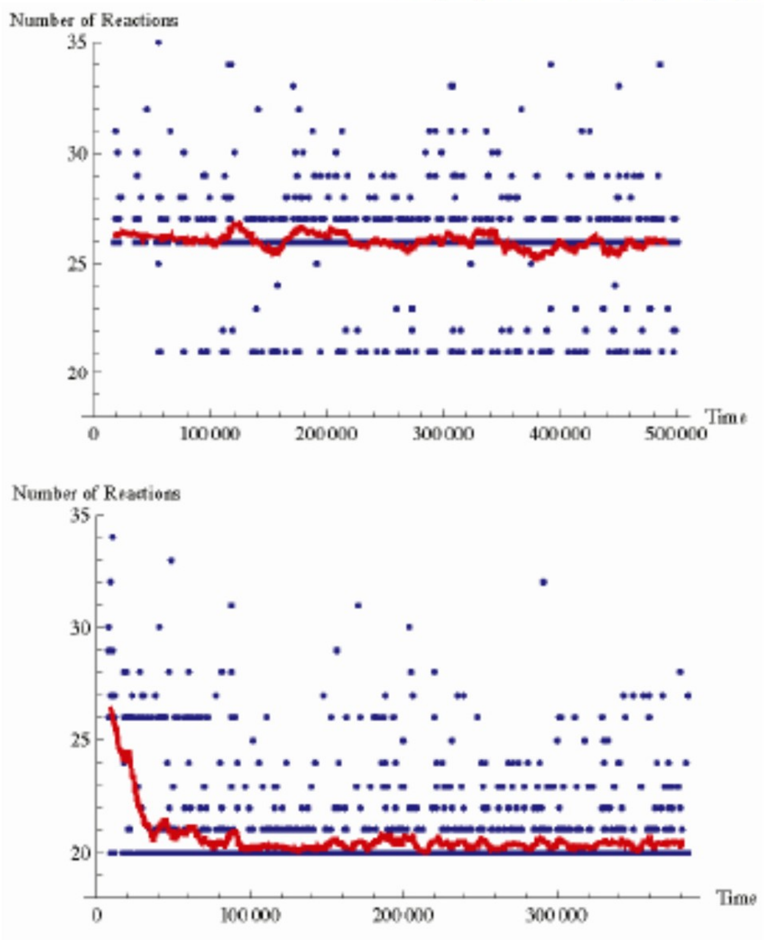


Summary of the evolvability analyses

Selection is possible in the novel species model

	Standard model	Inhibition model	Novel species model
Compartments	✓	✓	✓
Multiple attractors	✗	✓	✓
Selectability	✗	✗	✓

Natural selection experiments



WITHOUT SELECTION
viable loop enables a higher growth rate and therefore the network with the large viable loop (characterized by 26 reactions) constitutes the most frequent network type

SELECTION AGAINST THE LOOP
1 percent fitness advantage (S = 1.01) attributed to the networks without the loop, it is possible to reduce its frequency

∑: Selection *is* possible

Summary

- Rare novel species generate reaction avalanches
- Novel extensions to the network (non-food generated autocatalytic loops) ⇨ bits of heritable information

- Properties:
  - Attractor-based evolution
  - Weakly correlated variation
  - Limited heredity

- Small catalytic networks are selectable in a population

OTKA

COST

gencat
www.gencat.net

EUROPEAN
SCIENCE
FOUNDATION
SETTING SCIENCE AGENDAS FOR EUROPE

MINISTRY OF
SCIENCE AND
TECHNOLOGY

SEVENTH FRAMEWORK
PROGRAMME