# Evolution towards the Hypercycle:

## A spacial Model of Molecular Evolution

Camille Stephan-Otto Attolini

Bled, February 2004

Institute for Theoretical Biochemistry

University of Vienna

# The RNA World Hypothesis

The RNA World hypothesis for the beginning of life resides in the idea that RNA molecules were able to form a close and self contained system of replication and catalysis.
It is necessary then, that some simple molecules were able to replicate and present catalytic functions.

## RNA capable of catalysis

According to Scheuring, RNA present catalytic activities themselves. This is why we assume that at some point a system of organized replicators could have existed with auto and catalyzed replication. But...

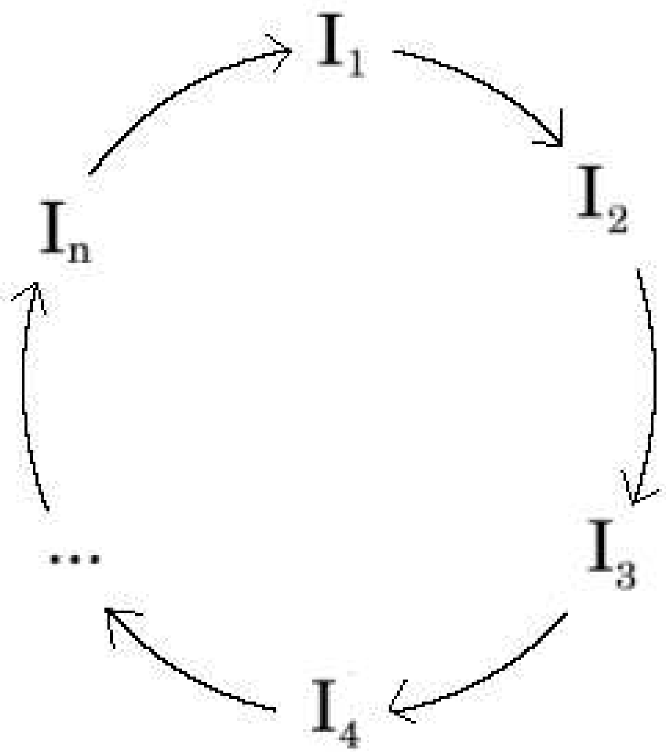
## There is a maximum size for the sequences

Eigen presented in 1971 a crucial result: given the replication fidelity that could exist in such a system, there is a maximum length for the replicators in order to maintain a minimum amount of information in the next generation. Estimating the copying fidelity, sequences could be of no more than a few dozens nucleotides.

# Eigen's Paradox or the Catch 22

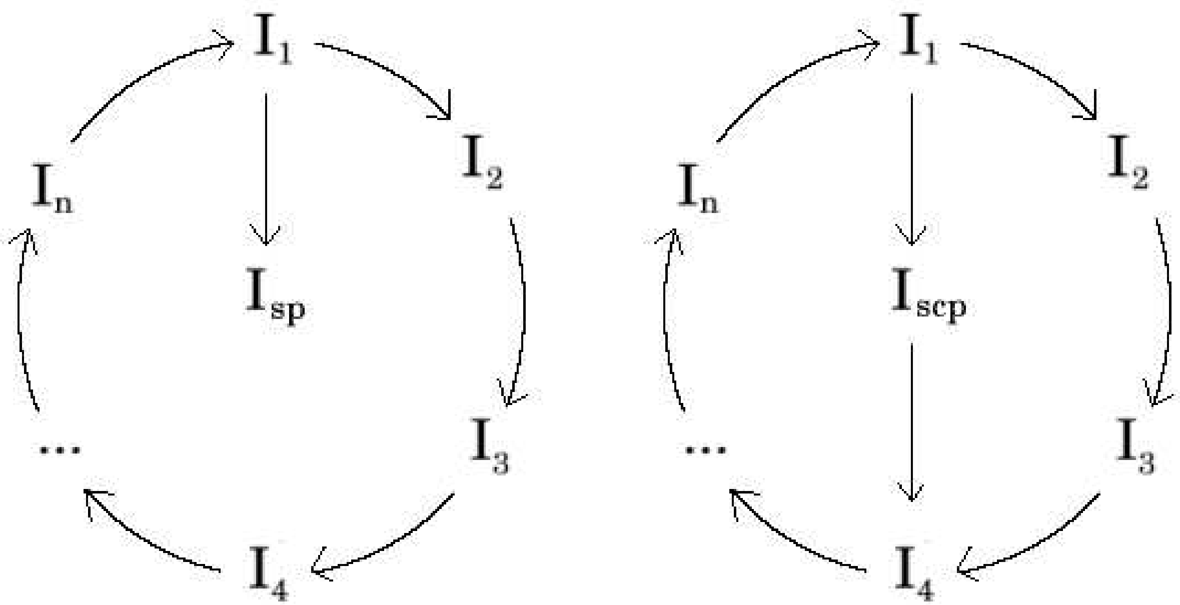
What happens then, if the fidelity of replication depends on the size of the replicator (the shorter the best) and the larger the more information can be stored, and a better copying fidelity attained?...

# The Hypercycle comes into scene...




Catalytic activity organized in a cyclic way. The information is carried by many small molecules instead of only a large one.

# But...




The hypercycle in the stirred reactor is unstable against parasites.

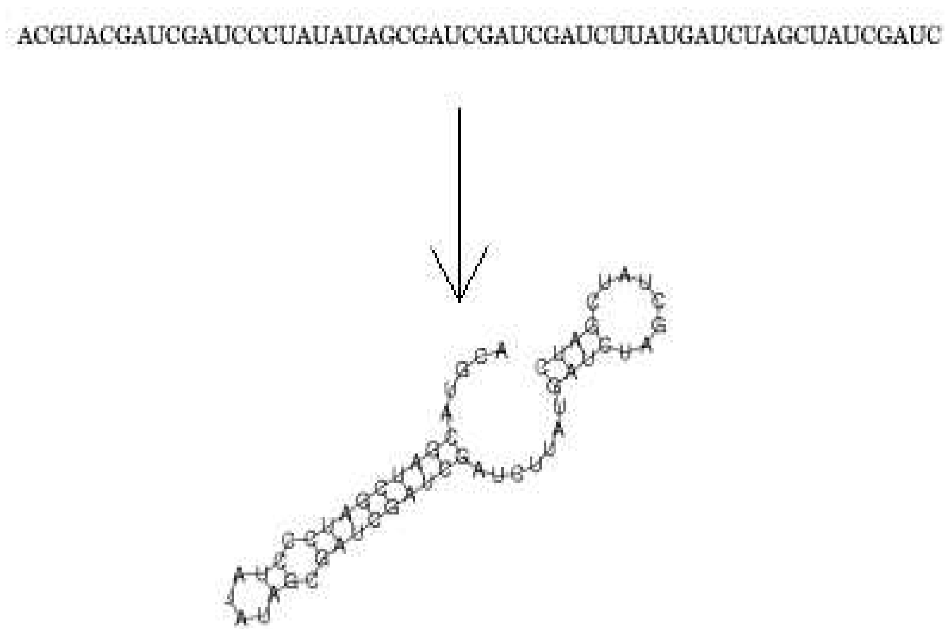
## One possible solution presented by Boerlijts and Hogeweg

When reactions take place in a lattice, instead of a stirred reactor, spacial patterns emerge and the system results stable against selfish parasites. The idea of life beginning in surfaces was first proposed by Wächtershäuser in a sea vents scenario.

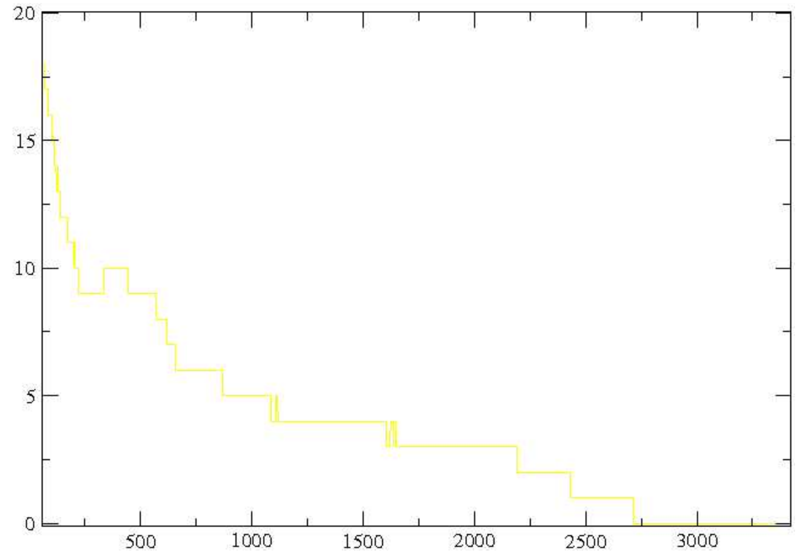
## The genotype-phenotype map

The problem of finding the relation between genotype and phenotype is too complex to be solved nowadays. One approach often used is the relation between sequence and structure in RNA molecules.

- A *secondary structure* $\Psi$ is a special type of contact structure, represented by a list of base pairs $(i, j)$ with $i < j$ on a sequence $x$, such that for any two base pairs $[i, j]$ and $[k, l]$ with $i \leq k$ holds: $(i)$ $i = k$ if and only if $j = l$, and $(ii)$ $k < j$ implies $i < k < l < j$.



# Quasispecies and evolution toward a predefine target.




Evolution of molecules with selection and mutation is known to present a stair wise behavior. It also shows a quasispecies distribution for the individuals per sequence.

# **O**ur Model

# The basic properties of the model are:

- We use a square lattice with continuous border conditions (which means a torus).
- Each square can be empty or occupied by only one sequence.
- Sequences are 56 bases long and are supposed to be self replicators.
- Toffoli-Margolus method is used for diffusion.

# The basic properties of the model are:

- We use a square lattice with continuous border conditions (which means a torus).
- Each square can be empty or occupied by only one sequence.
- Sequences are 56 bases long and are supposed to be self replicators.
- Toffoli-Margolus method is used for diffusion.

## The basic properties of the model are:

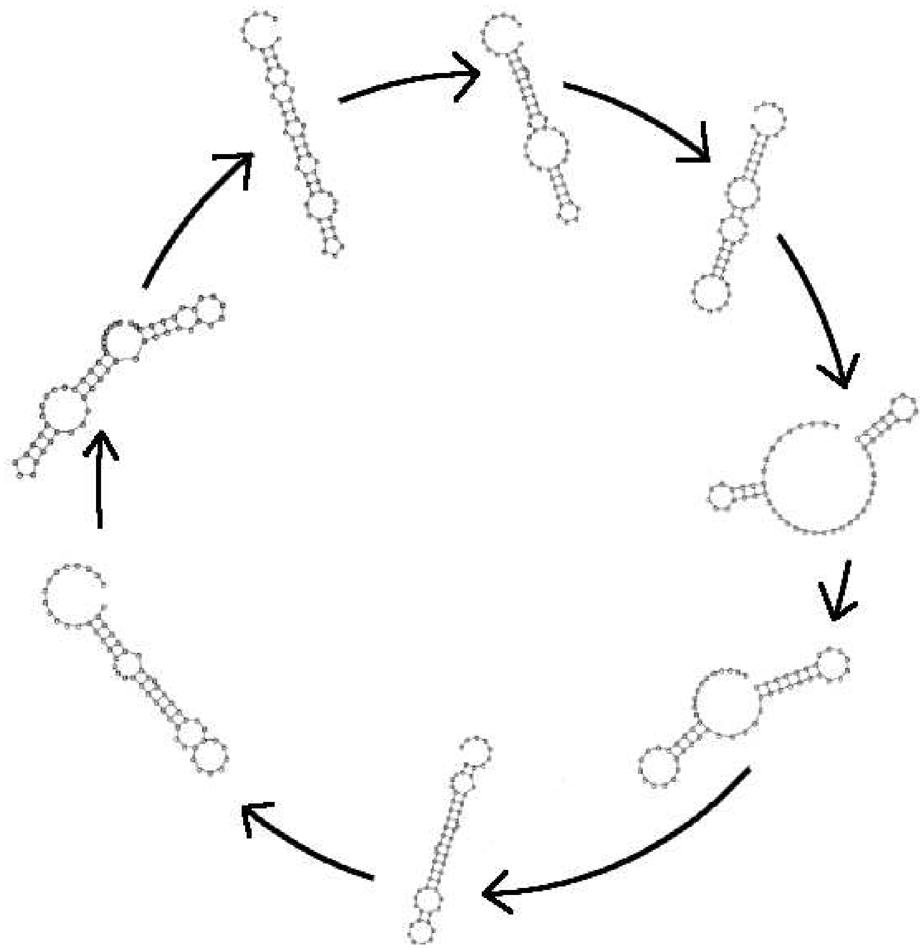
- We use a square lattice with continuous border conditions (which means a torus).
- Each square can be empty or occupied by only one sequence.
- Sequences are 56 bases long and are supposed to be self replicators.
- Toffoli-Margolus method is used for diffusion.

## The basic properties of the model are:

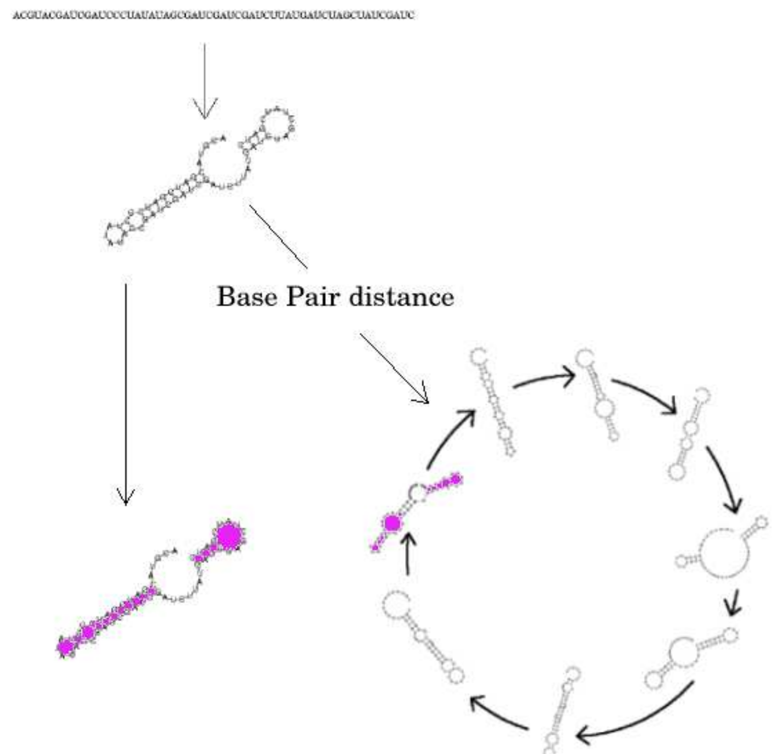
- We use a square lattice with continuous border conditions (which means a torus).
- Each square can be empty or occupied by only one sequence.
- Sequences are 56 bases long and are supposed to be self replicators.
- Toffoli-Margolus method is used for diffusion.

# The target set.

The basic configuration of the target set is an 8 members hypercycle

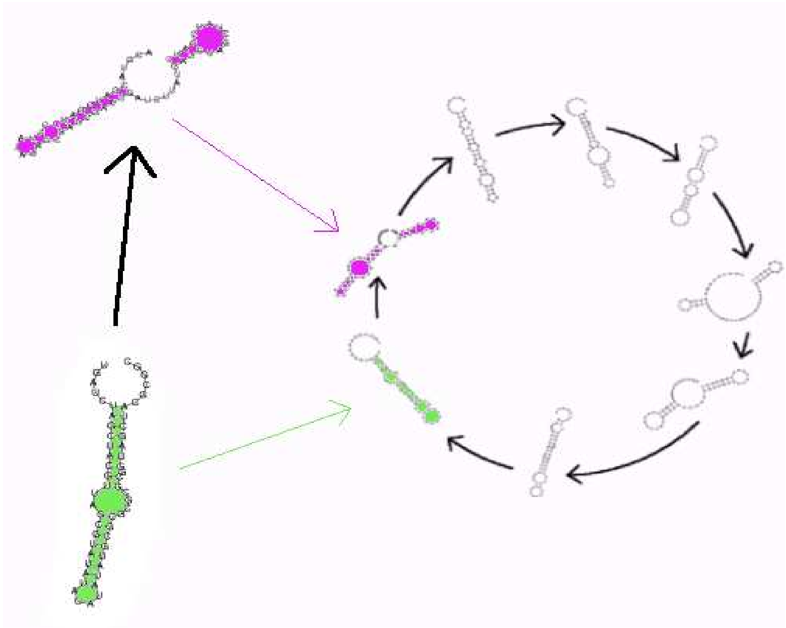


# How we define the way a sequence behaves...



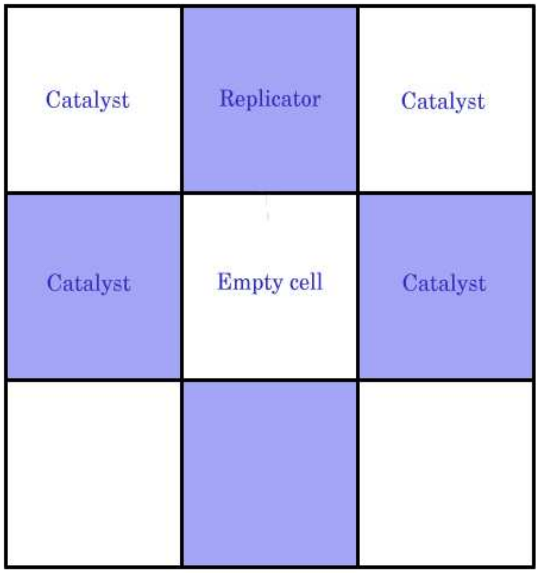

For each target, the base pair distance is computed. We define the sequence to belong to the group of the closest target.

# Replication rates



Once the sequence's group is defined, its self replication rate and decay rate are computed depending on the distance to the corresponding target. Catalytic rates depend also on this distance, and on the topology of the target set.

# Replication rules




Colored squares are possible invaders to the empty cell. Probability of replication depends on the total replication rate for each replicator and a probability for staying empty.

# Initial conditions

- Some parameters of the system are the ratio between self replication and catalyzed replication and decay rate.
- The initial number of sequences in the lattice can vary, giving different results ranging from the survival of only one group to the formation of spatial patterns like those found by B. and H. This will also be very important for the diversity of the system.
- Diversity in our system is calculated by computing the average distance between all sequences of each group for each time step.

# Initial conditions

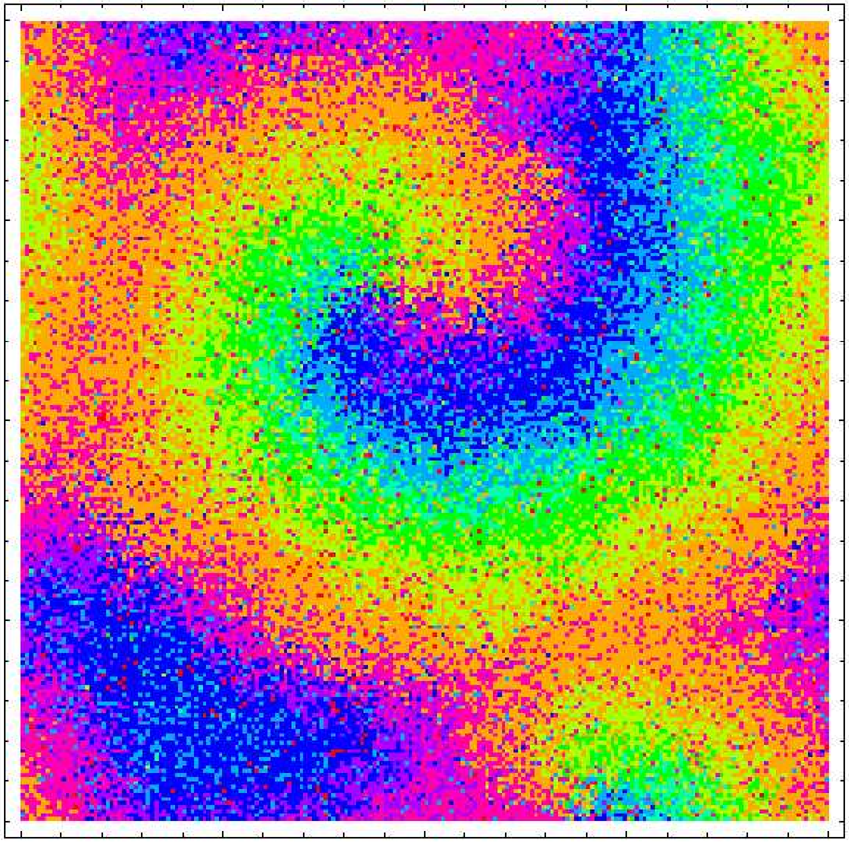
- Some parameters of the system are the ratio between self replication and catalyzed replication and decay rate.
- The initial number of sequences in the lattice can vary, giving different results ranging from the survival of only one group to the formation of spatial patterns like those found by B. and H. This will also be very important for the diversity of the system.
- Diversity in our system is calculated by computing the average distance between all sequences of each group for each time step.

# Initial conditions

- Some parameters of the system are the ratio between self replication and catalyzed replication and decay rate.
- The initial number of sequences in the lattice can vary, giving different results ranging from the survival of only one group to the formation of spatial patterns like those found by B. and H. This will also be very important for the diversity of the system.
- Diversity in our system is calculated by computing the average distance between all sequences of each group for each time step.

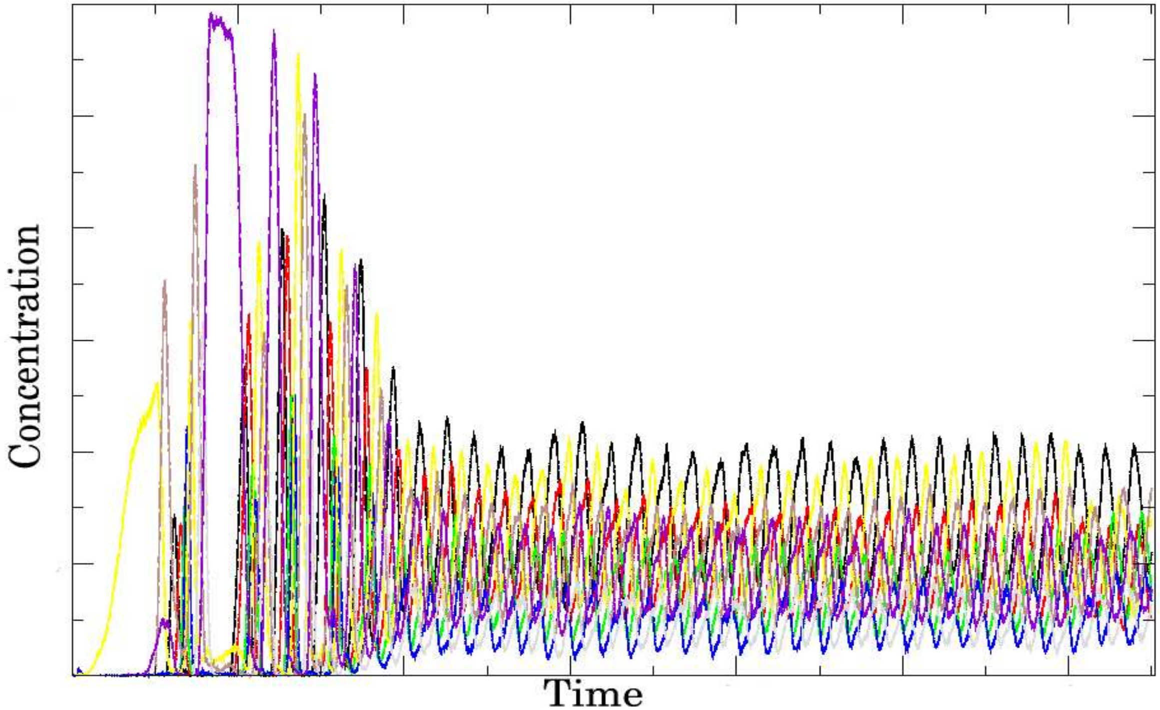
# Results

# Spirals!!



Spacial patterns are formed after some period.

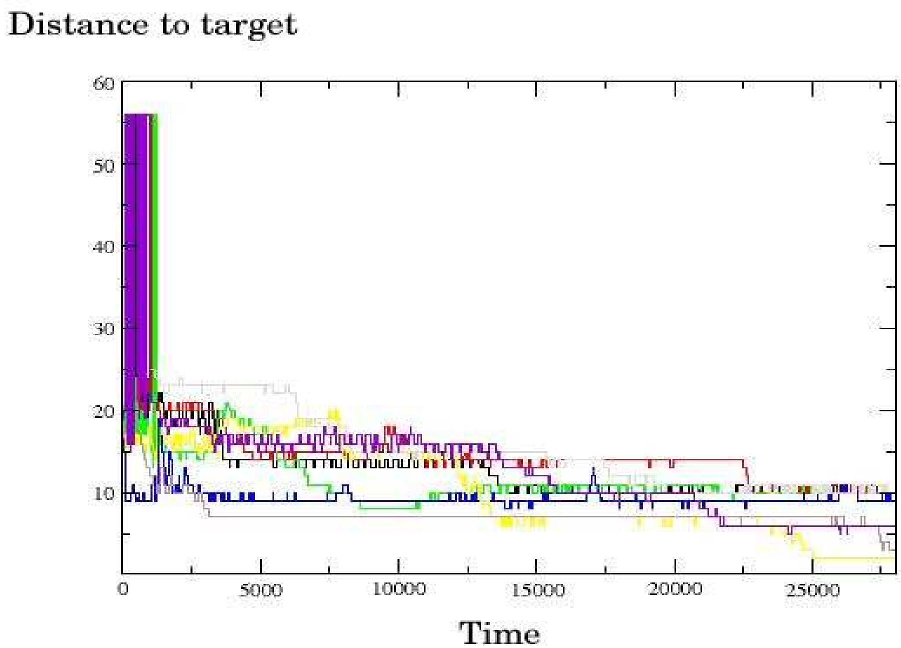
# Concentration




Each colored line corresponds to the concentration of a different group, after a period of disorder a change in the system's behavior occurs and an oscillating pattern can be seen.

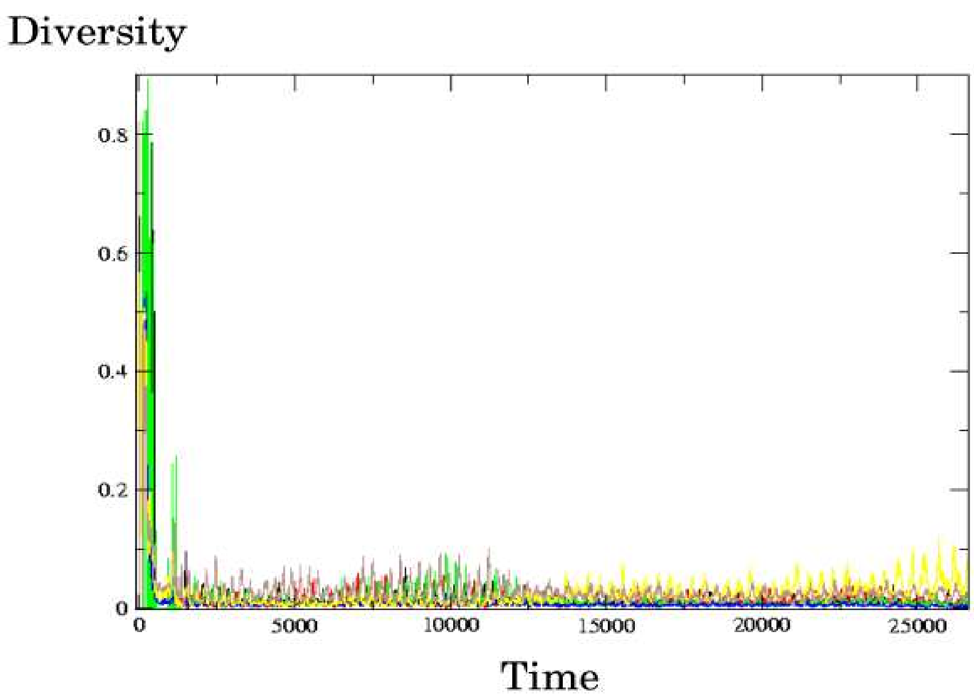
# Distance to the target

Fitness is defined as the inverse of the distance to the target. It is possible that some groups find the target while other's evolution is lowered or even stopped by the system organization.



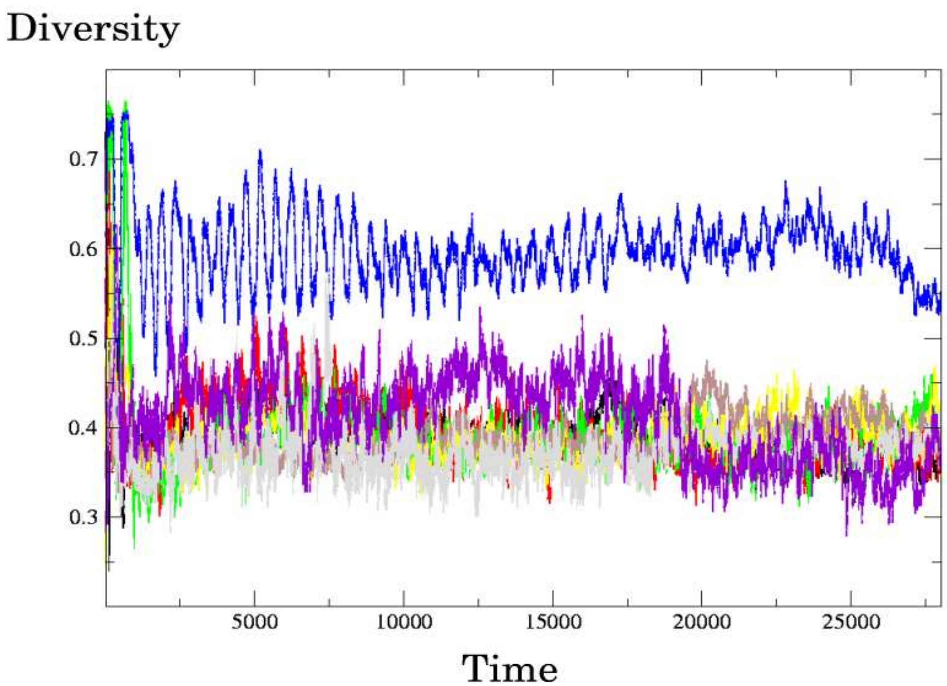
# Diversity

If number of sequence at the beginning is low, the system's diversity will fall to almost zero and stay like that.



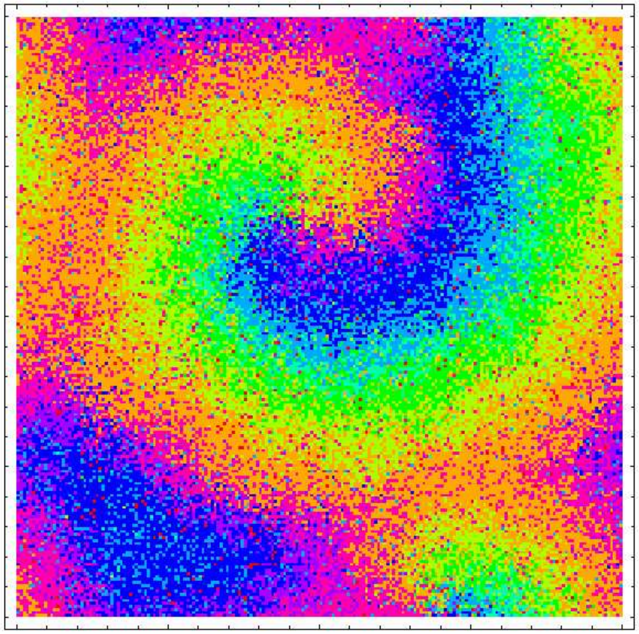
# Diversity

If, in change, the number of sequences is high, then diversity will stay high oscillating the same way as the number of individuals per group.



# Resistance to parasites: Selfish parasite

As it was found before, the spacial organization of the system makes it stable against parasites. In the case of the selfish parasite, the invader is moved to the outer layer of the spiral leaving it with no chance of killing the system.



# Resistance to parasites: Short-cut parasite

In the B-H model, the short-cut parasite was able to kill the system by repeated shortening of the hypercycle. In our system, this kind of parasite was expelled from the system, making it lose its fitness and so stopping it to return to the system in a harmful way.

# Resistance to parasites: Short-cut parasite