

Evolution of mutation rate in RNA-like replicator systems

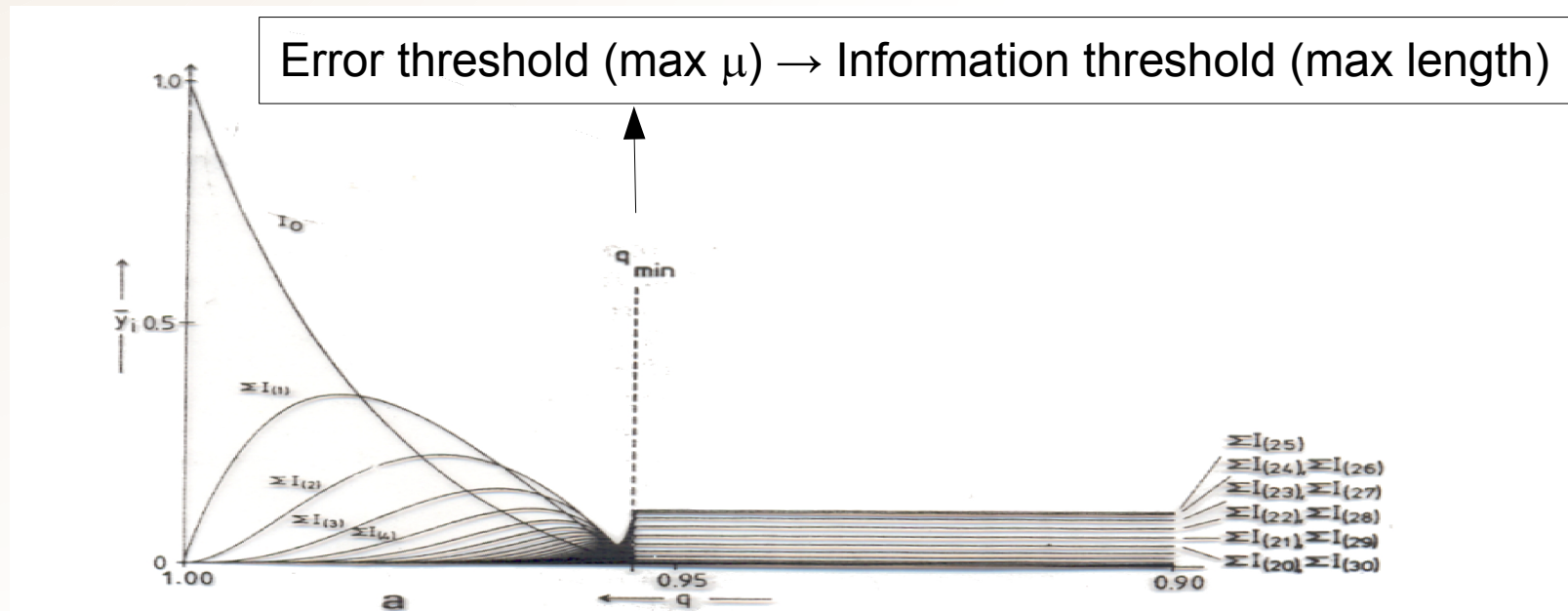
Enrico Sandro Colizzi and Paulien Hogeweg. Utrecht University.



Claude Monet, 1886. Travelling waves in a rugged landscape.

Introduction

- A quasi-species is defined as a stationary distribution of macromolecular species with closely interrelated sequences
- It appears in a population of replicators at mutation-selection equilibrium



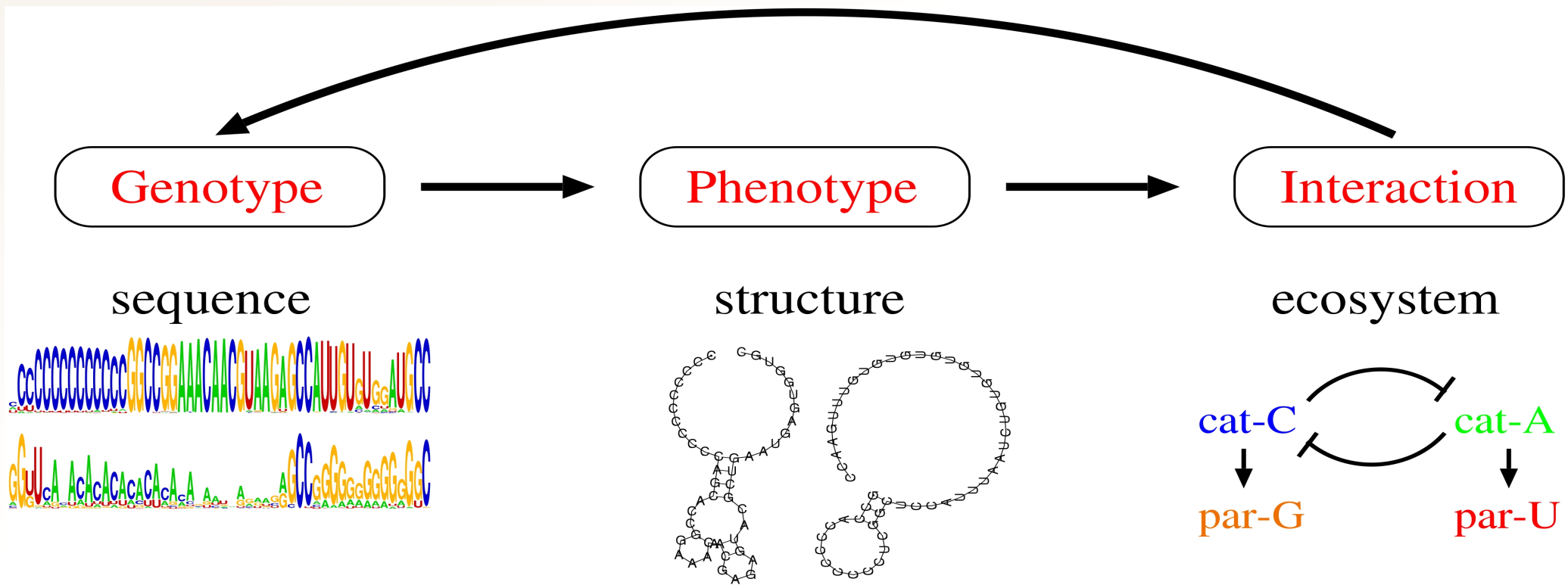
- Eigen's paradox states that it's impossible for a sequence to evolve longer than its Information threshold, for it would need some more accurate replication machinery, which could be encoded only in a longer genome

Introduction

- Addressing Eigen's paradox means answering to the question:
How is it possible to evolve / integrate new information,
past the Information threshold?
- Two possibilities:
 - Ecosystem-based solution
 - Single molecule-based solution (= genome)

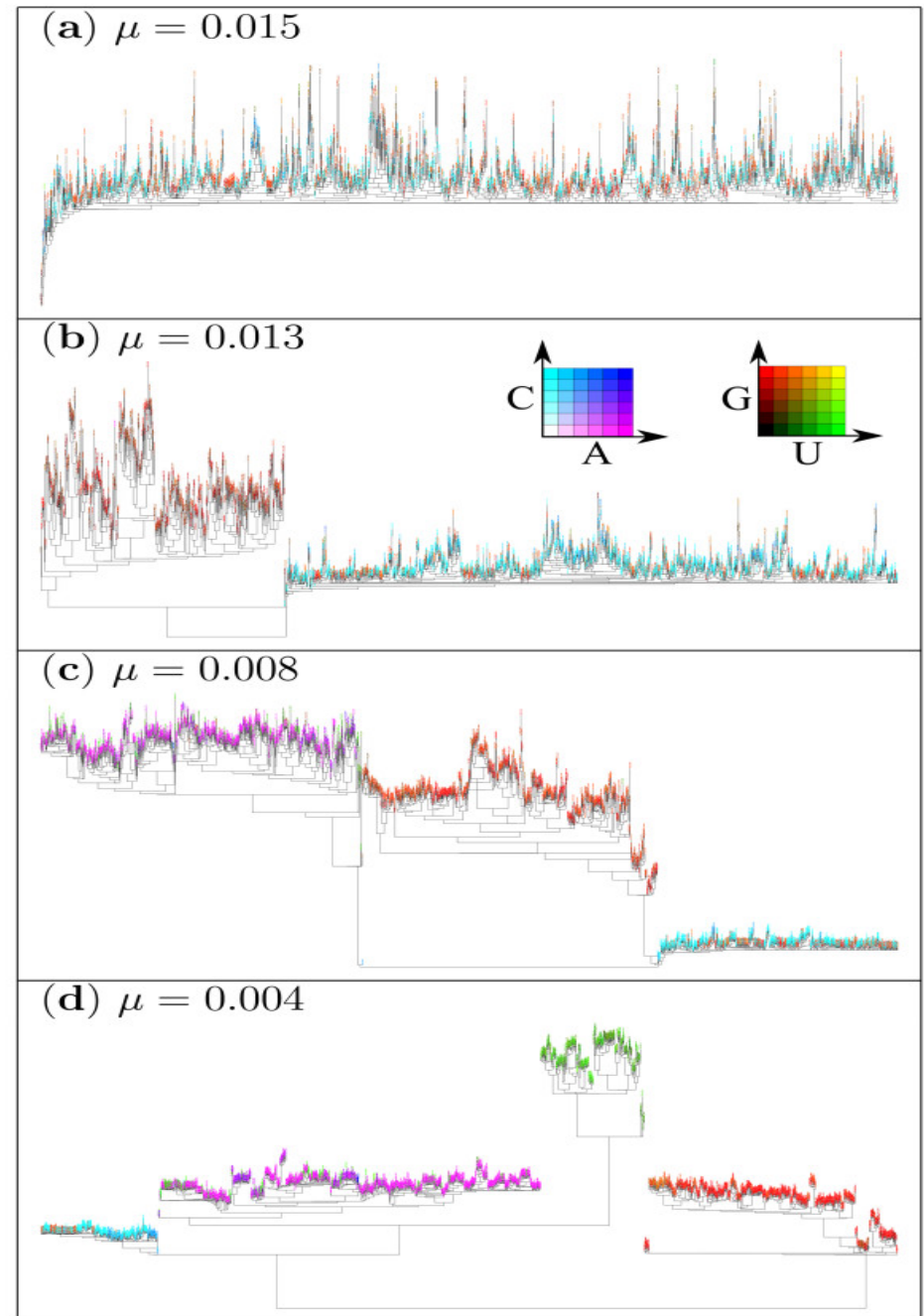
Introduction

Multilevel, eco-evolutionary models show that ecological complexity may evolve as a result of the feedback among information (genotype), function (phenotype), organization (interactions)



Introduction

- One function only is explicitly defined: being able to replicate other molecules.
- The evolution of parasites creates new niches and triggers an increase of complexity
- The eco-evolutionary dynamics consists of a chain process of niche creation and speciation
- By pre-setting the mutation rate to low values, up to four species can coexist



Mutation rate as phenotypic parameter

- Variable, sequence-independent per-base mutation rate (keeping sequence length constant)
 - μ is defined for every molecule, included those that don't have catalytic activity
- Upon replication:
 - The number of mutations in the new sequence is (stochastically) determined by the replicase
 - If mutations happened: $\mu_{\text{new molecule}} = \mu_{\text{parent}} \pm \delta$
 - (δ is a small random number)
 - If no mutations happened: $\mu_{\text{new molecule}} = \mu_{\text{parent}}$

Mutation rate as phenotypic parameter

- In order to initialize the field at the Error threshold, sequences have to be pre-evolved to tolerate high mutation rate
- The initial population consists entirely of the master (most abundant) sequence at the end of the pre-evolution step
- μ as high as possible (Error threshold):
 $\mu = 0.013 - 0.015$ chance of substitutions per base, per replication (~ 50% chance per molecule, per replication)

Mutation rate as phenotypic parameter

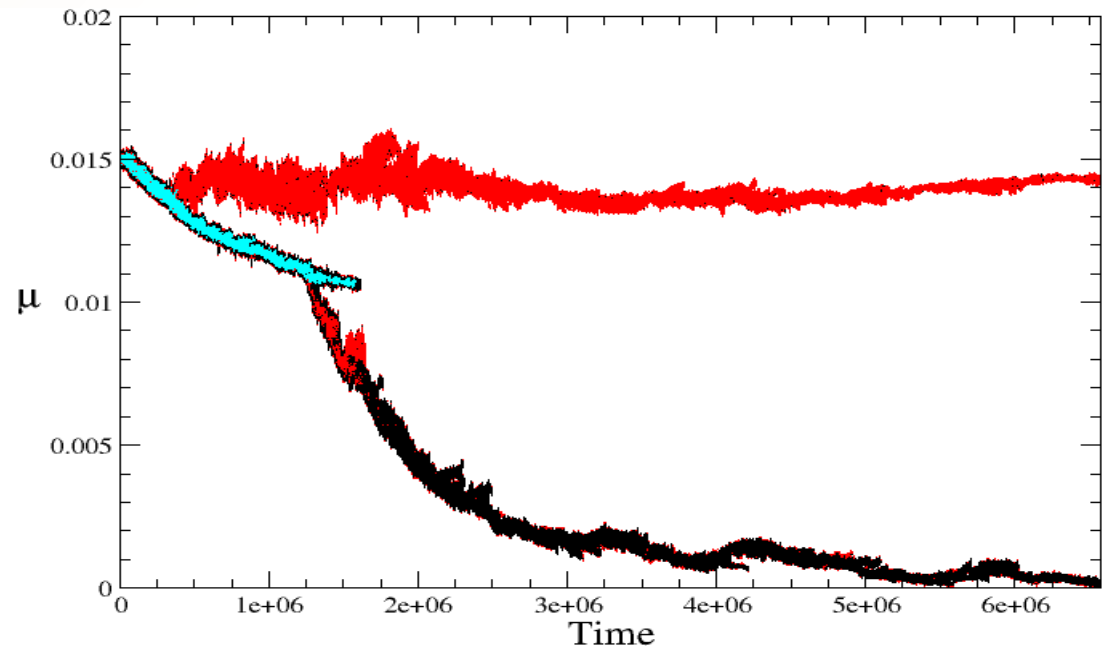
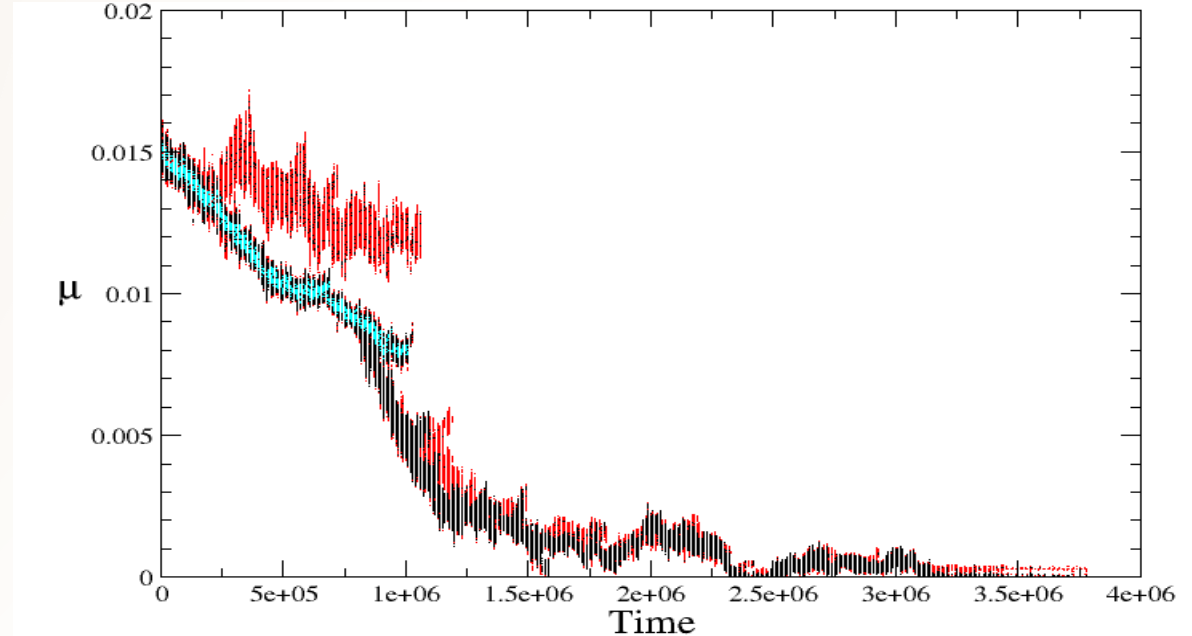
- The average μ for the catalytic lineages decreases because it reduces the mutational load

- Parasites speciate

The average μ_{parasite} drifts

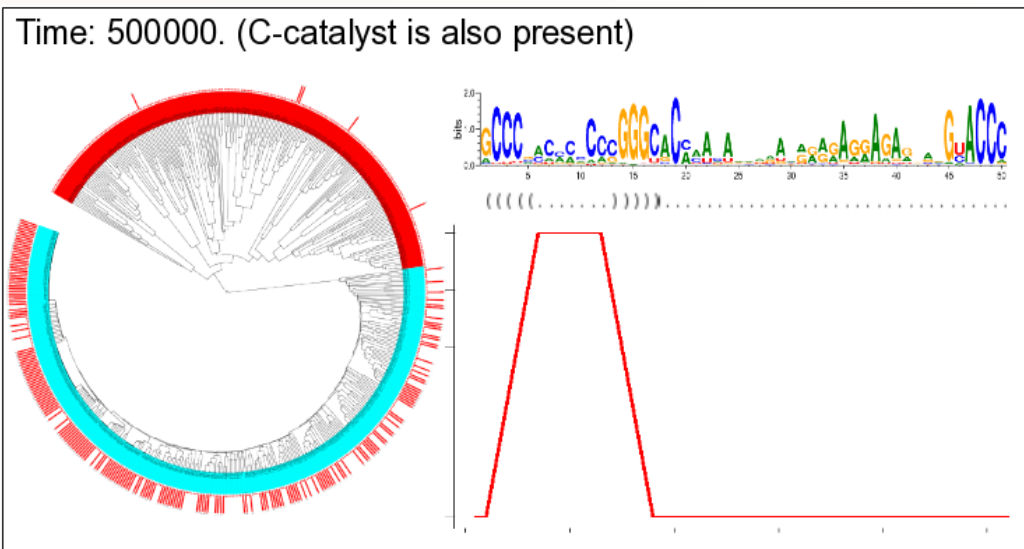
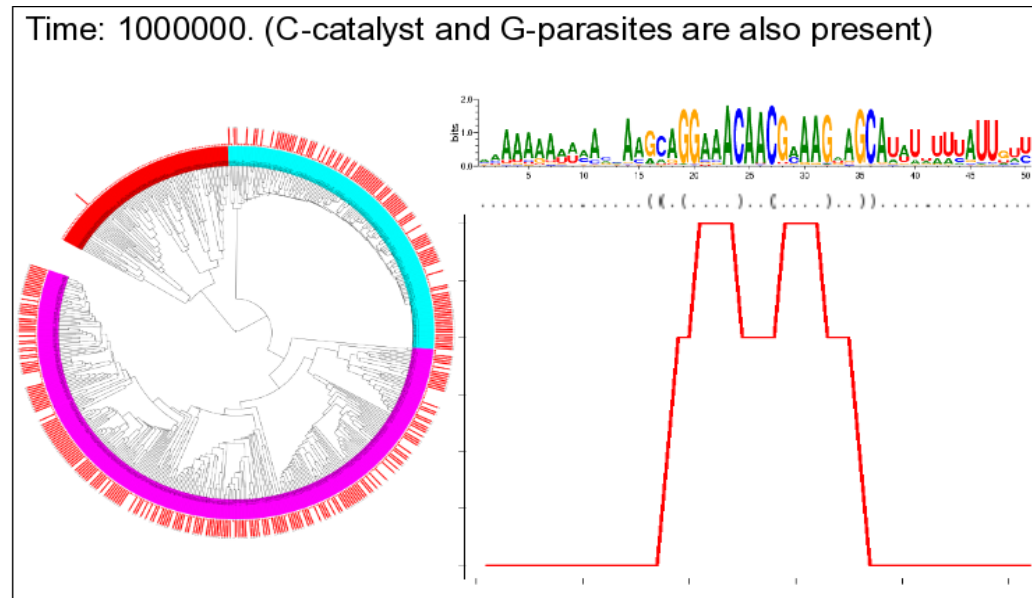
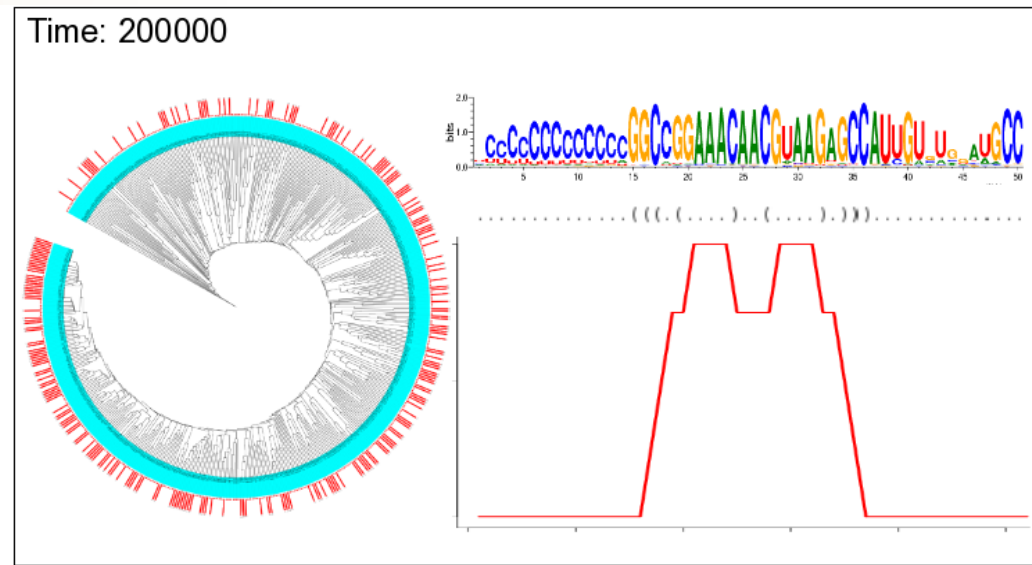
- Three species coexistence on a transient

(Four species is possible with different catalytic structure)



Mutation rate as phenotypic parameter

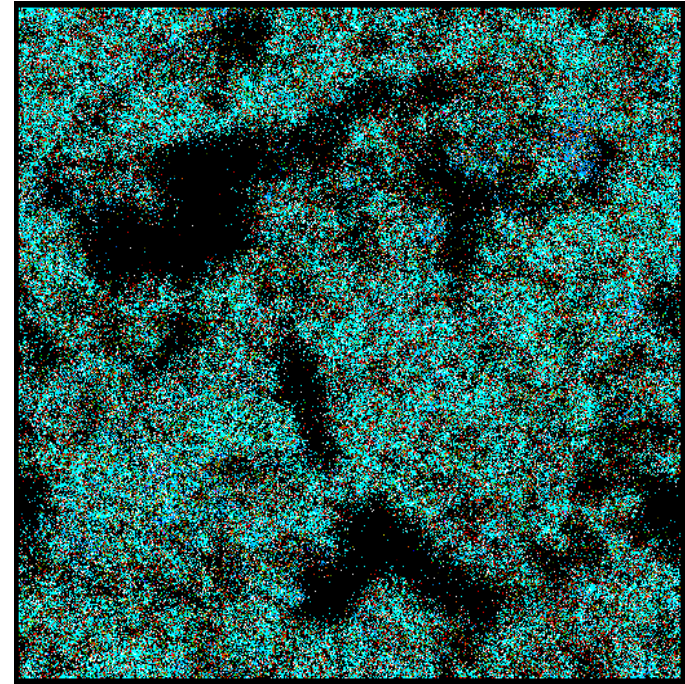
- $\mu = 0.015$: One quasi-species
- $\mu = 0.013$: Speciation of a parasite
- $\mu = 0.009$: New catalytic quasi-species
- $\mu < 0.009$: Red queen dynamics
- Different quasi-species speciate by evolving a different nucleotide usage for their interactions



Mutation rate as phenotypic parameter

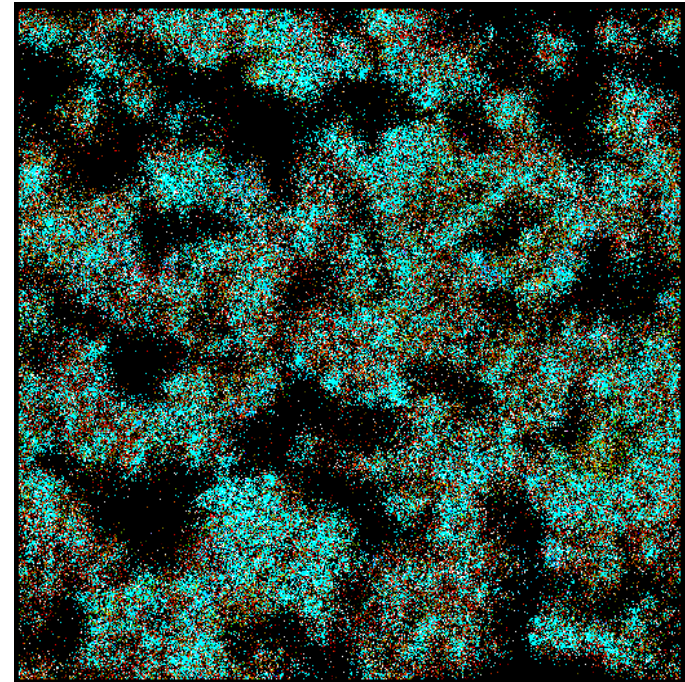
First catalytic quasi-species:

- Only ~10% H=1 mutants of the master sequence are functional (*steep* quasi-species)
 - *Junk* molecules may interfere with parasites (if present)
- μ decreases slowly through back mutations that restore the master sequence



Parasitic quasi-species:

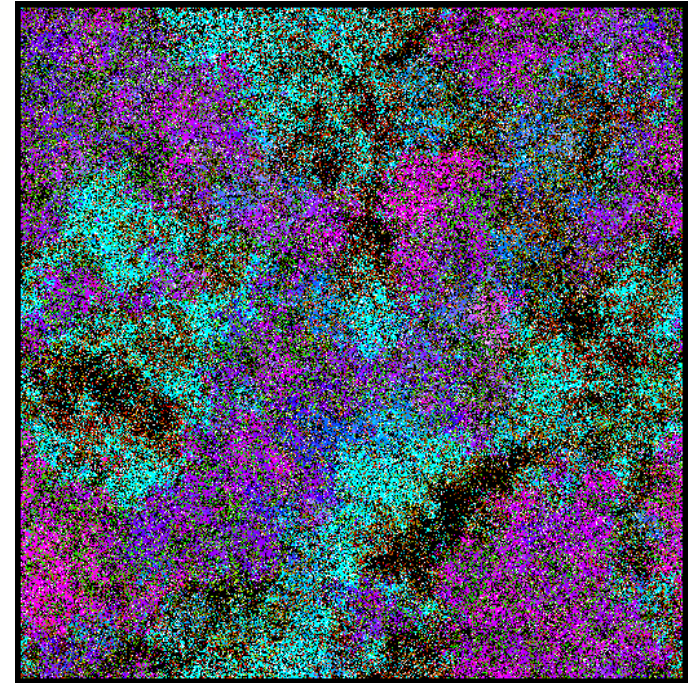
- Parasites are not predefined functionalities, their secondary structure is free to change (albeit the portion used for interactions)
- Its foldings are the most abundant in structure space



Mutation rate as phenotypic parameter

Second catalytic quasi-species:

- Lower μ , triggers the evolution of a parasite, which creates a niche for a new catalytic species
- New, *spread* catalytic quasi-species:
 - up to 50% H=1 substitutions of some sequences make a functional catalyst
 - evolution pushes the quasi-species to the most connected parts of the neutral network
- It can evolve only when μ is low enough:
- High neutrality \rightarrow high variability \rightarrow
 - \rightarrow lower replication rate
 - \rightarrow (but also) faster decrease of $\mu \rightarrow$ starts Red queen dynamic (no long standing quasi-species)



Mutation rate as phenotypic parameter

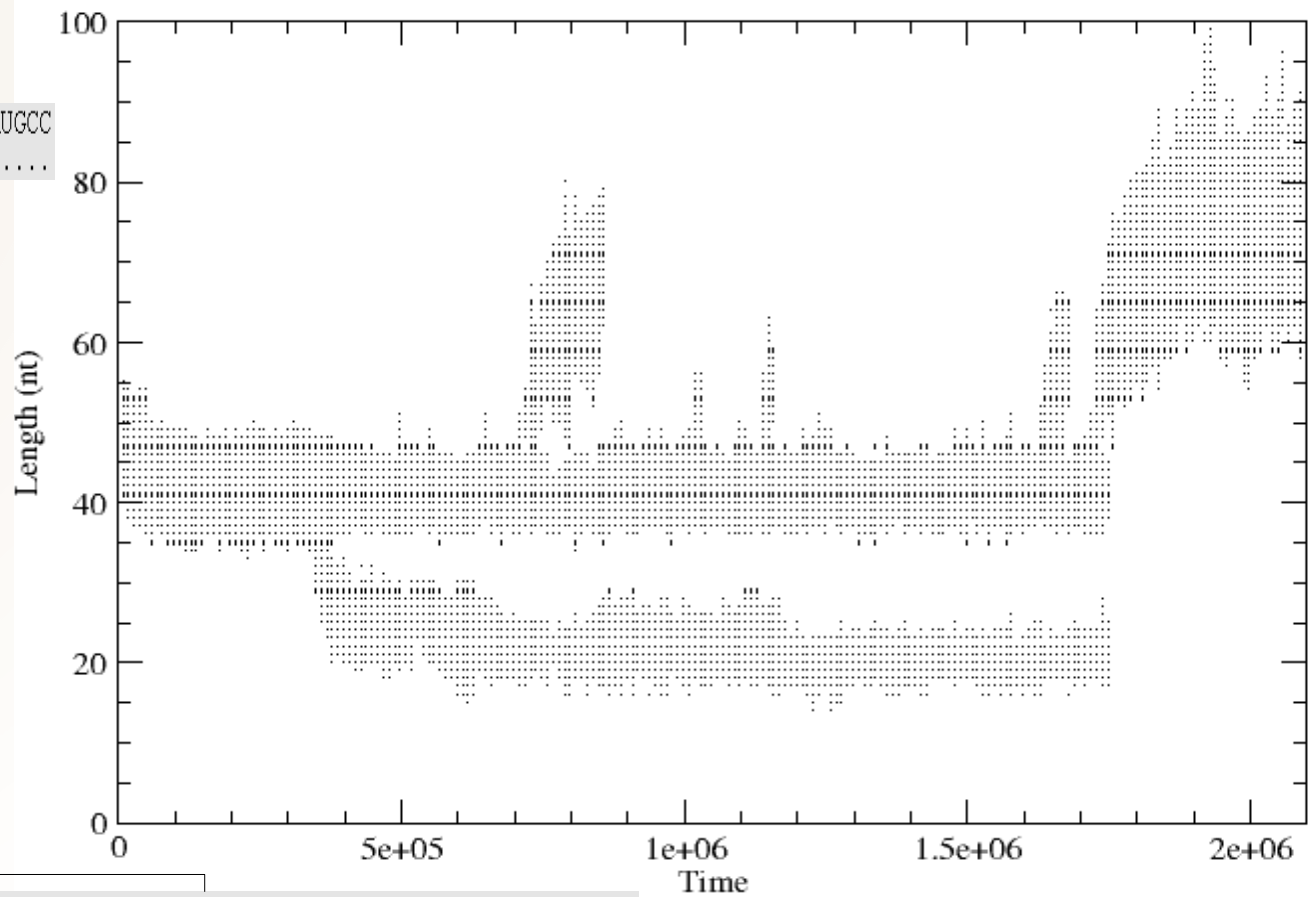
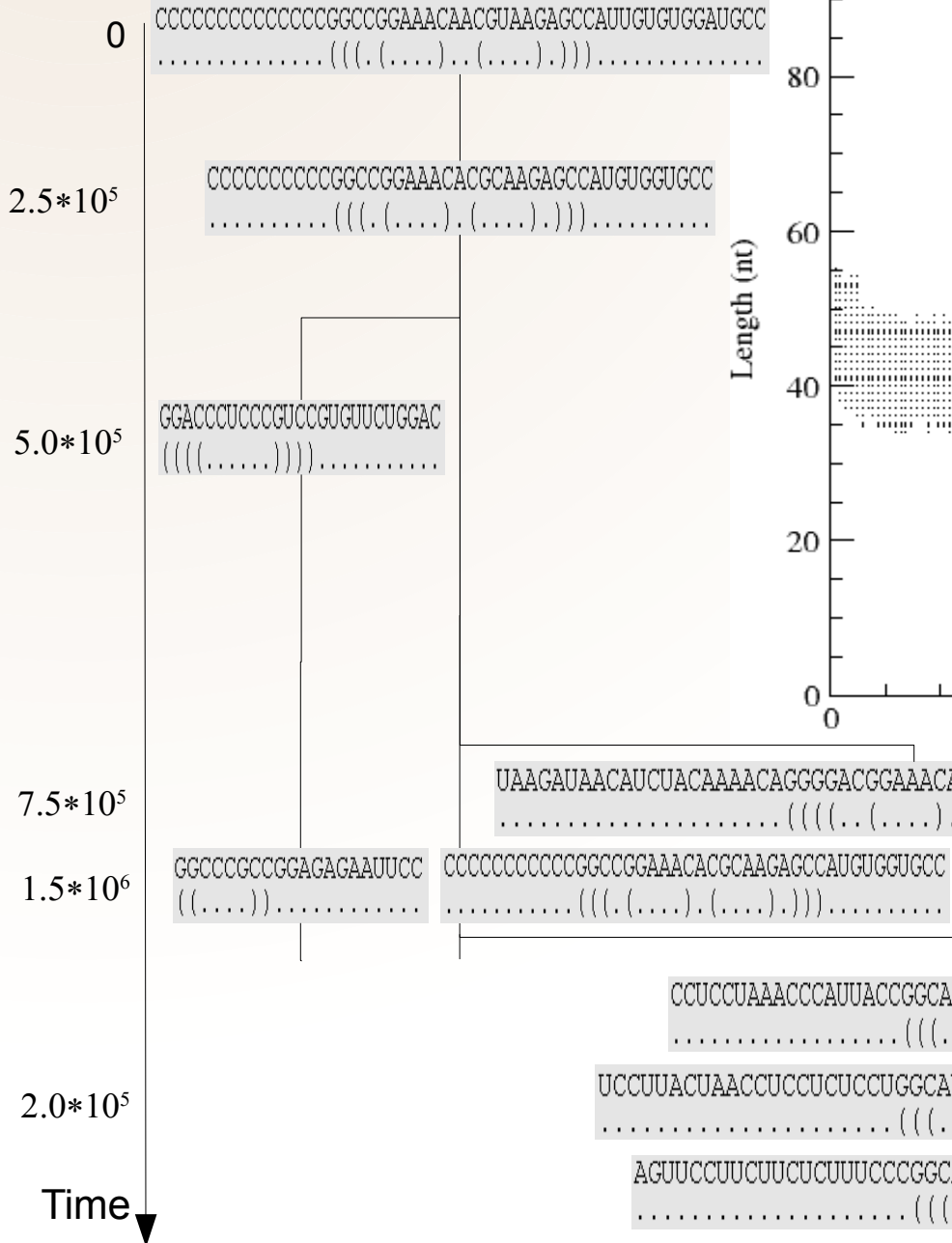
- So far:
 - even if formulated independently, sequence and mutation rate feedback on each other,
 - thereby influencing the eco-evolutionary dynamics, and being influenced by it.
 - The *steepness* of the first quasi-species makes the system avoid Red queen dynamics, making the mutation rate decrease slowly to the point where the ecological dynamics allows a *spread* quasi-species to arise.

Sequences with variable length

- Now, let the length of the sequences change and the per-base mutation rate be constant:
 - $\mu_{\text{subst}} = 0.013$
 - $\mu_{\text{in/del}} = \mu_{\text{subst}} / 10$
- The implicit per-molecule mutation rate changes as sequences grow longer or shorter
- With previous model, in this parameter regime, only the first catalytic quasi-species and its parasite could evolve

Sequences with variable length

Ancestry:



Sequences with variable length

- The average sequence length for the *steep* quasi-species shrinks (50 nt → 42 nt):
 - The original catalyst was at its error/information threshold, “stretched” to 50 nt.
- A short parasite evolves (~21 nt)
 - Get rid of several “useless” nucleotides
- After a long time (and a few unsuccessful tries), the neutral catalyst evolves and out-competes the previous quasi-species:
 - the eco-evolutionary mechanism of niche creation and speciation is the same
 - the new catalyst is long (60-90 nt)

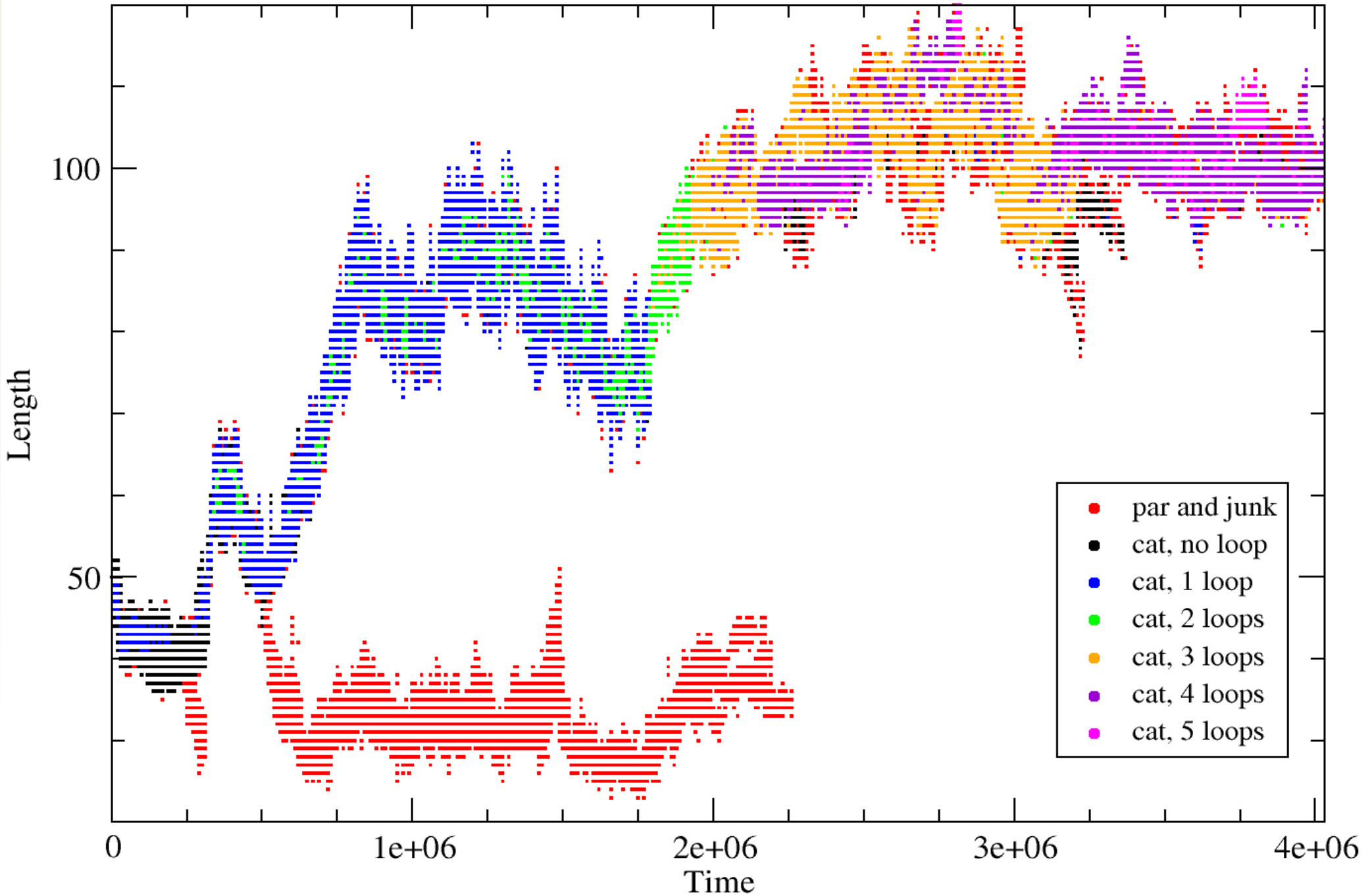
Sequences with variable length

- Surprising because with constant length set to 50 nt, $\mu=0.013$ is too high for the *spread* quasi-species to evolve
- Here it does evolve and it does so to be the longest sequence
 - Very high neutrality:
 - up to 70% H=1 substitutions yields functional catalytic structures
 - Very high variability:
 - compensated for by long dangling ends
 - low exploitability by parasites (at least, so far...)
- However, high “intrinsic” variability + high mutation rate can lead to “delocalization” (Muller's ratchet) if the population is small:
 - In the small field, the system may go extinct

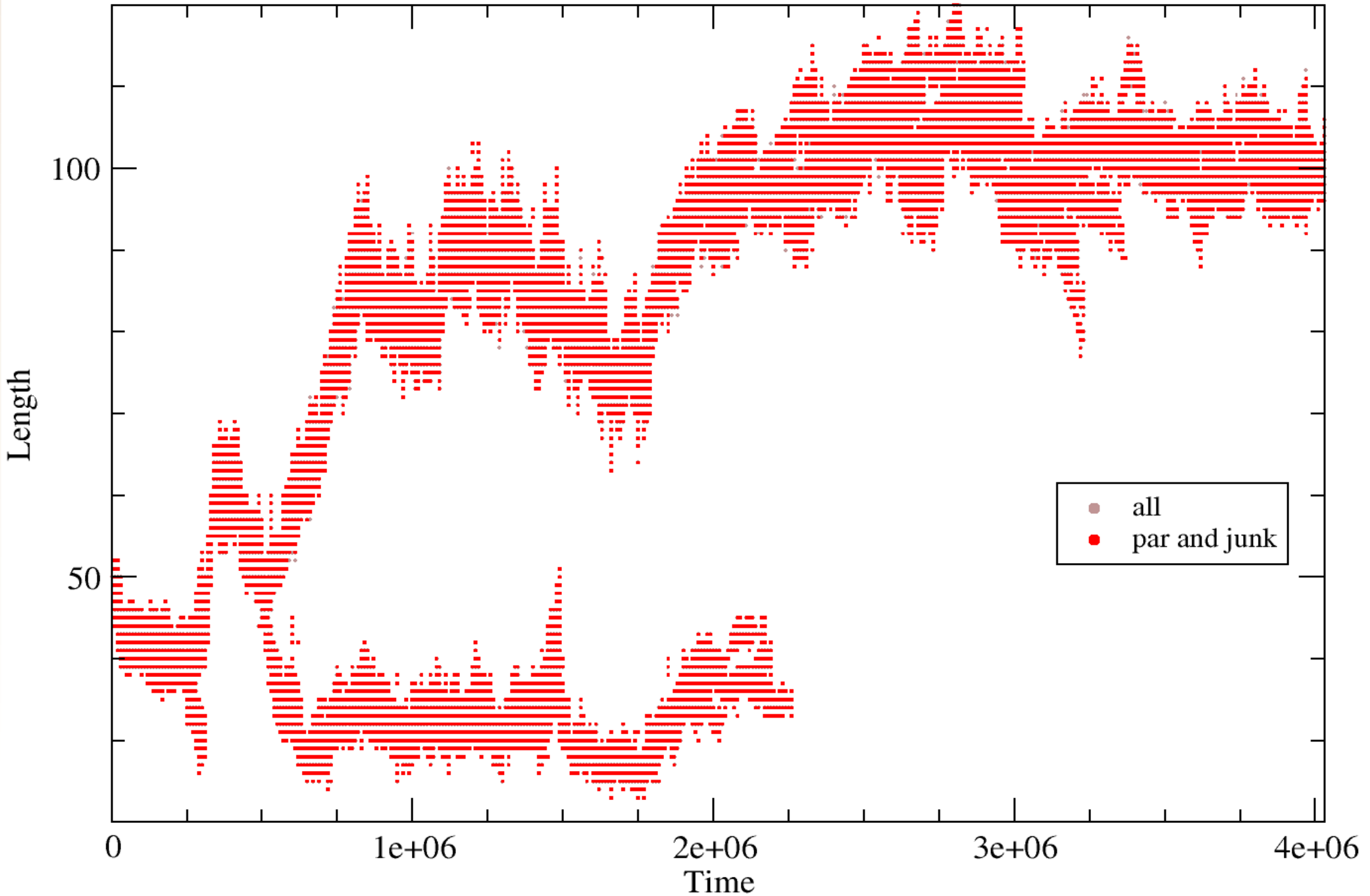
Evolving the proof reader

- The evolution of the long A catalyst suggests a mechanism to escape Eigen's paradox:
 - If a sequence containing the information for a replicase becomes longer, it may be able to host the extra information to become a *better catalyst*
- We define *better catalyst* a sequence that folds as an old catalyst but possesses at least one hairpin loop on its dangling ends
 - Each new hairpin loop contributes to decrease mutation rate
 - This implies a trade-off between dangling ends (replication rate) and mutation rate

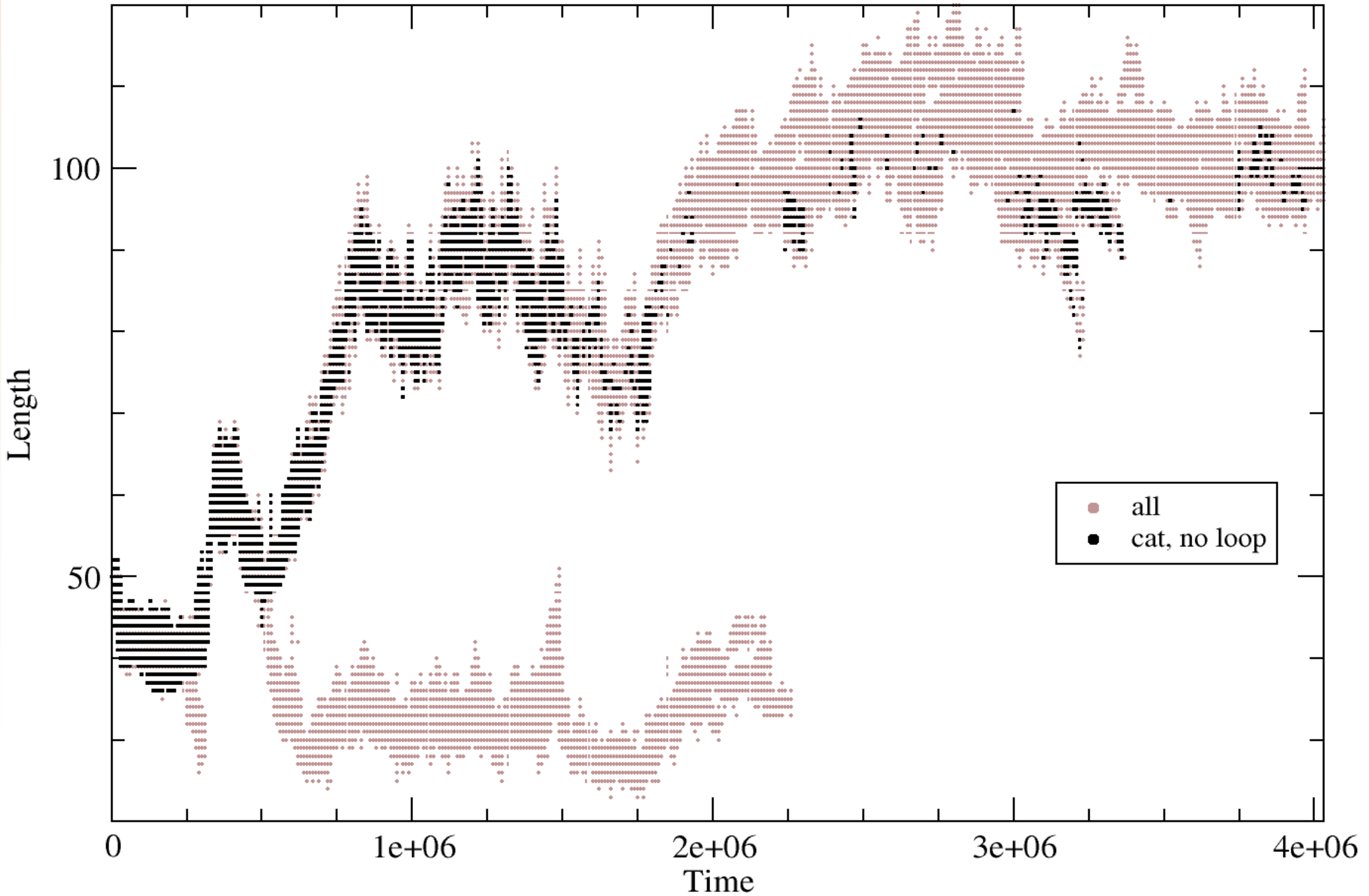
Evolving the proof reader



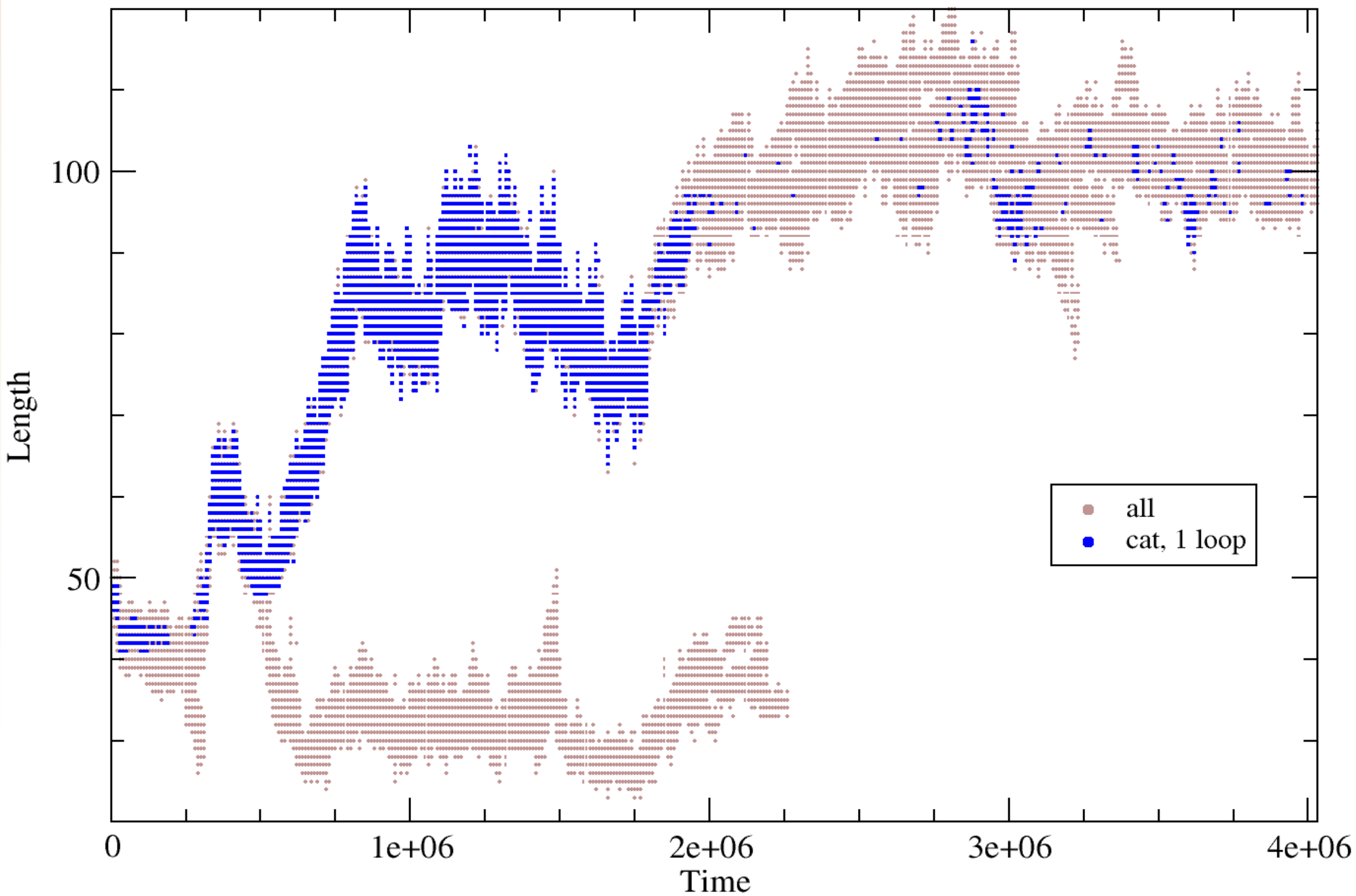
Evolving the proof reader



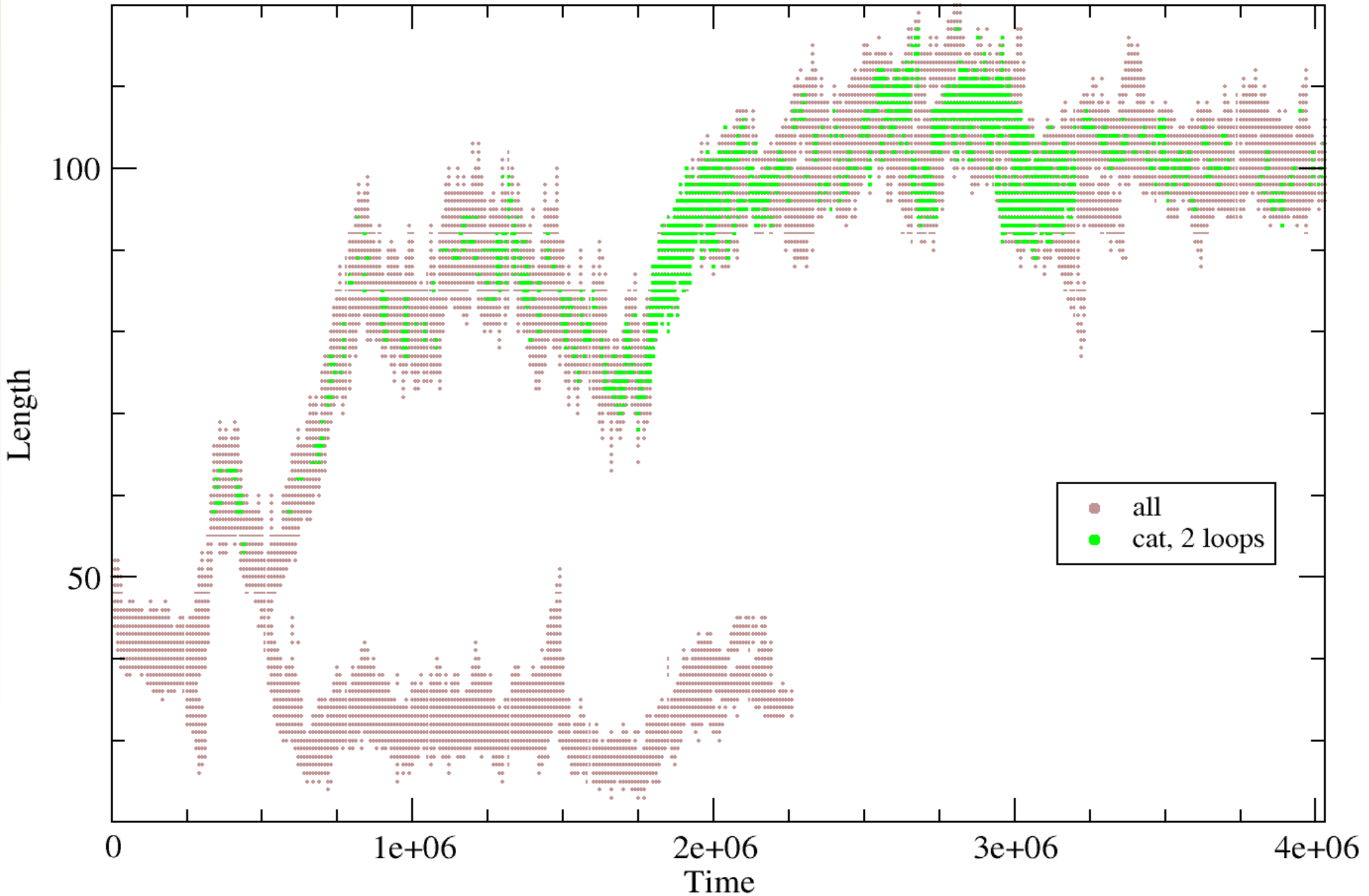
Evolving the proof reader



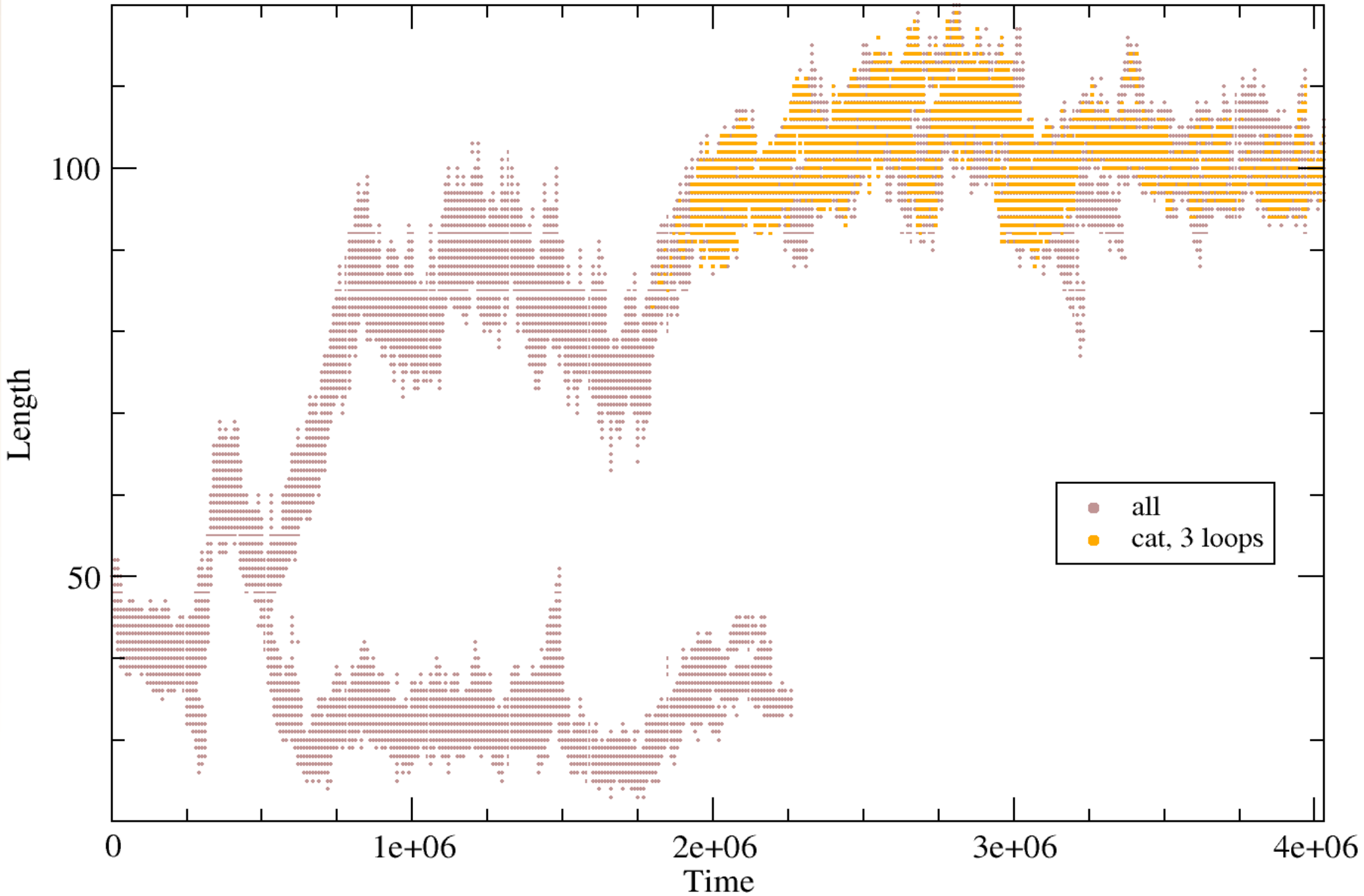
Evolving the proof reader



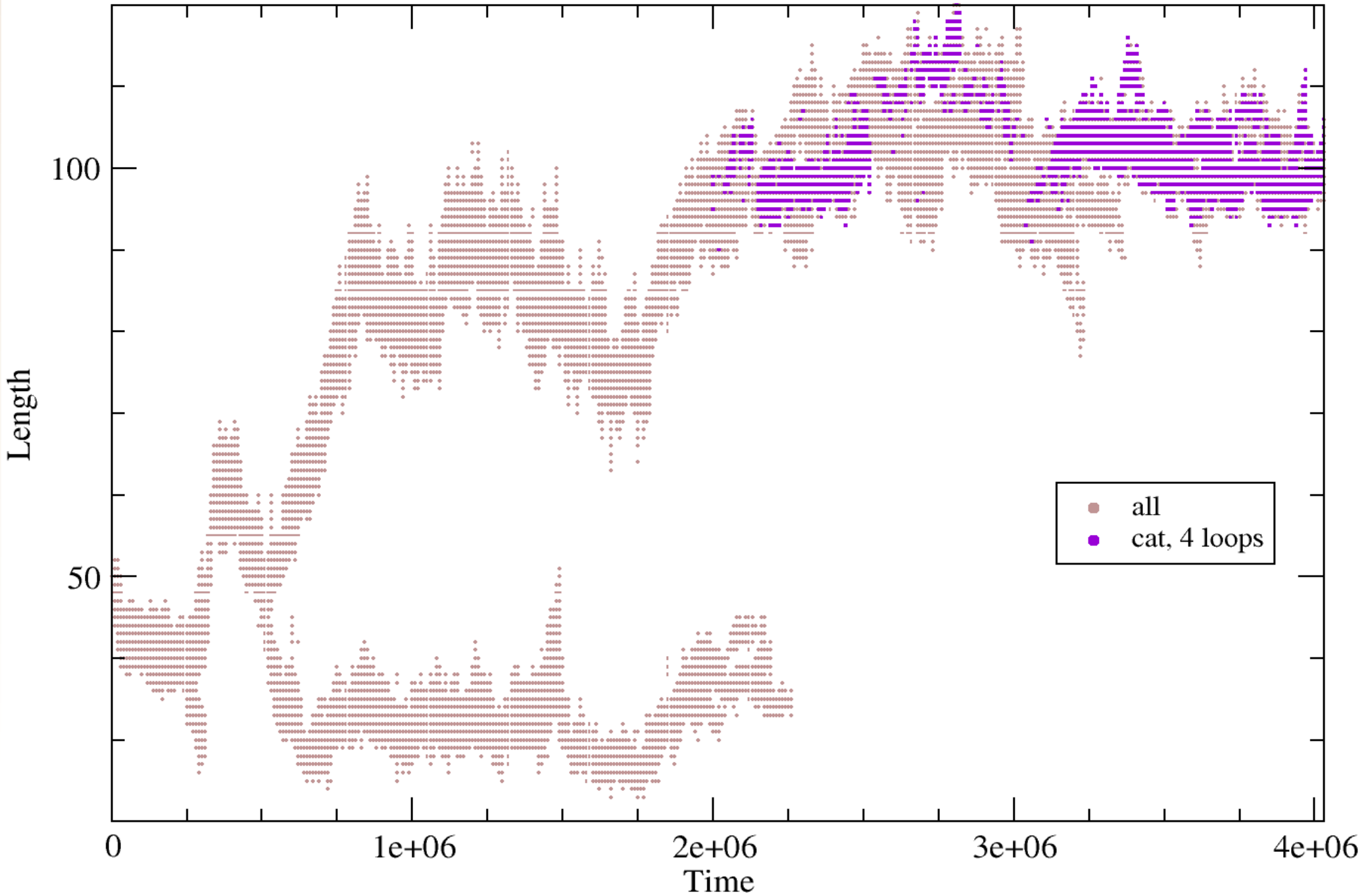
Evolving the proof reader



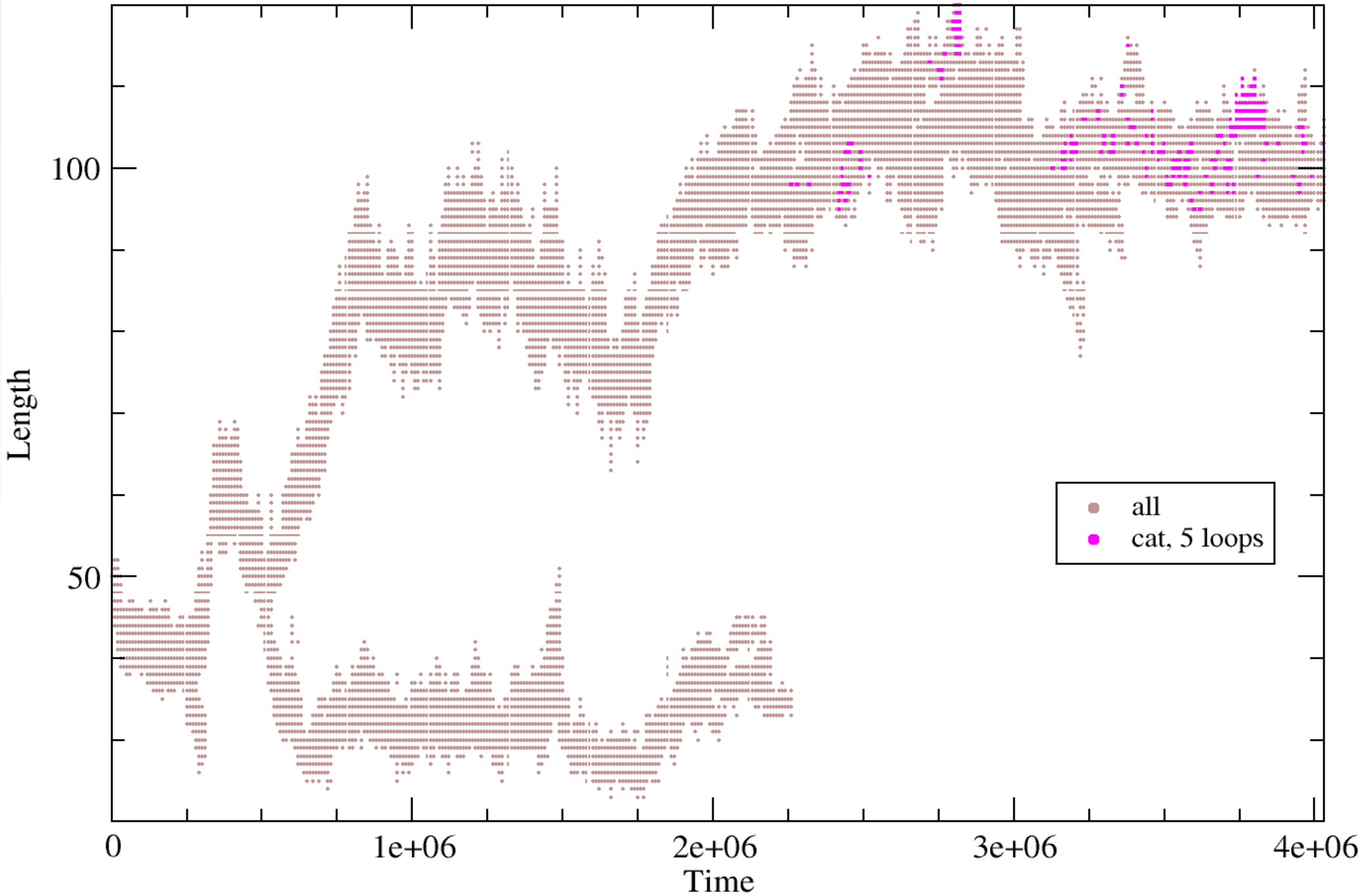
Evolving the proof reader



Evolving the proof reader

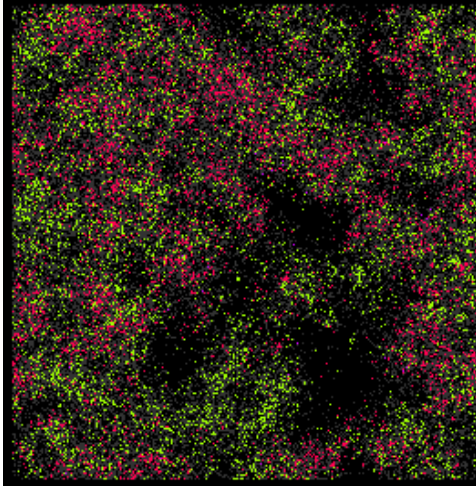
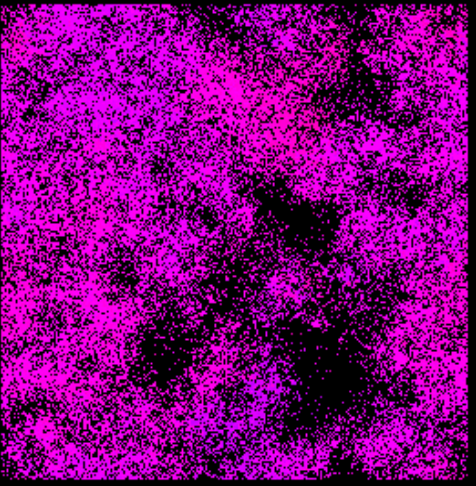
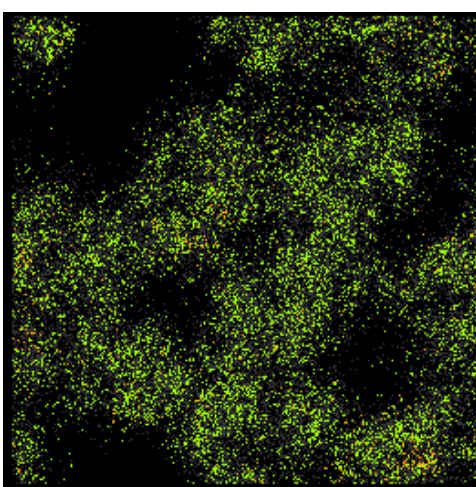
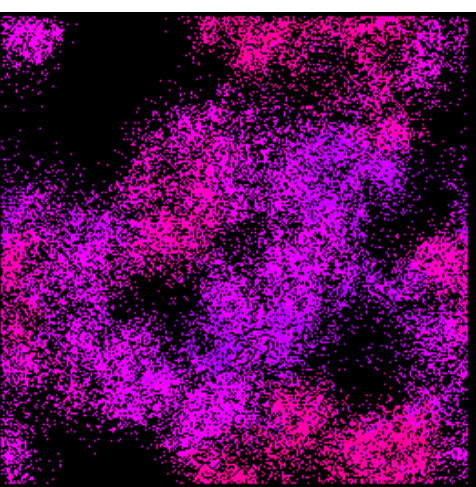


Evolving the proof reader



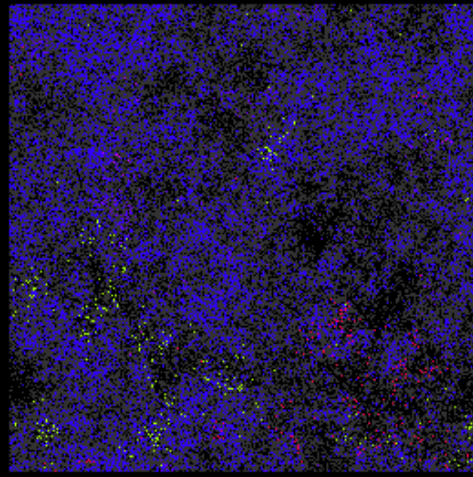
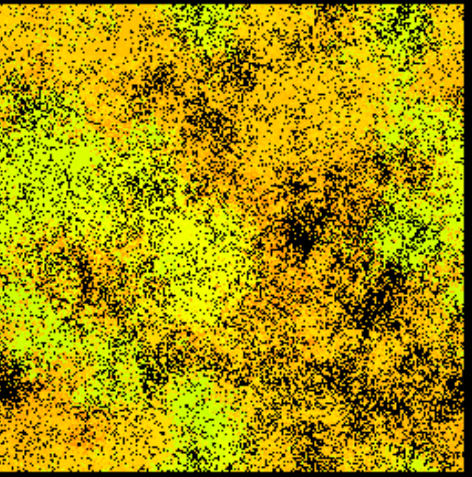
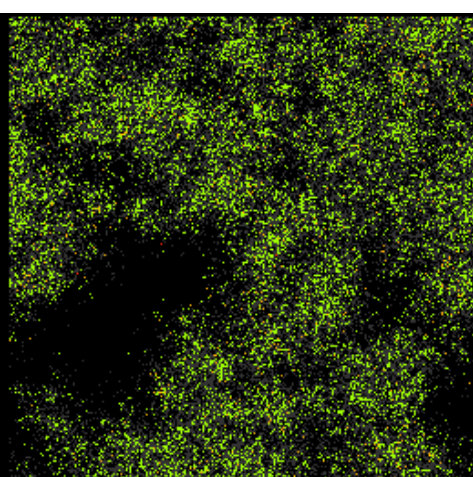
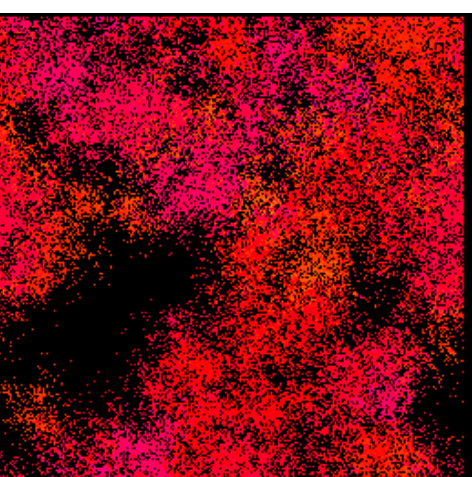
Evolving the proof reader

- However, it doesn't work (at least in the small field) if the gain in replication accuracy doesn't adequately compensates for the loss in replication rate

After invasion of the <i>spread</i> quasi-species	GREEN original μ RED lower μ (1 loop)	Average length ~ 60 nt
<ul style="list-style-type: none">• For each new loop: $\mu_{\text{new}} = 0.5 * \mu_{\text{old}}$• New catalyst manages to invade		
<ul style="list-style-type: none">• For each new loop: $\mu_{\text{new}} = 0.8 * \mu_{\text{old}}$• New catalyst cannot invade		

Evolving the proof reader

- If the better catalyst invades the first time, more loops will be discovered. Longer and longer catalysts will invade.
- Else, the length of the first *spread* quasi-species stabilizes

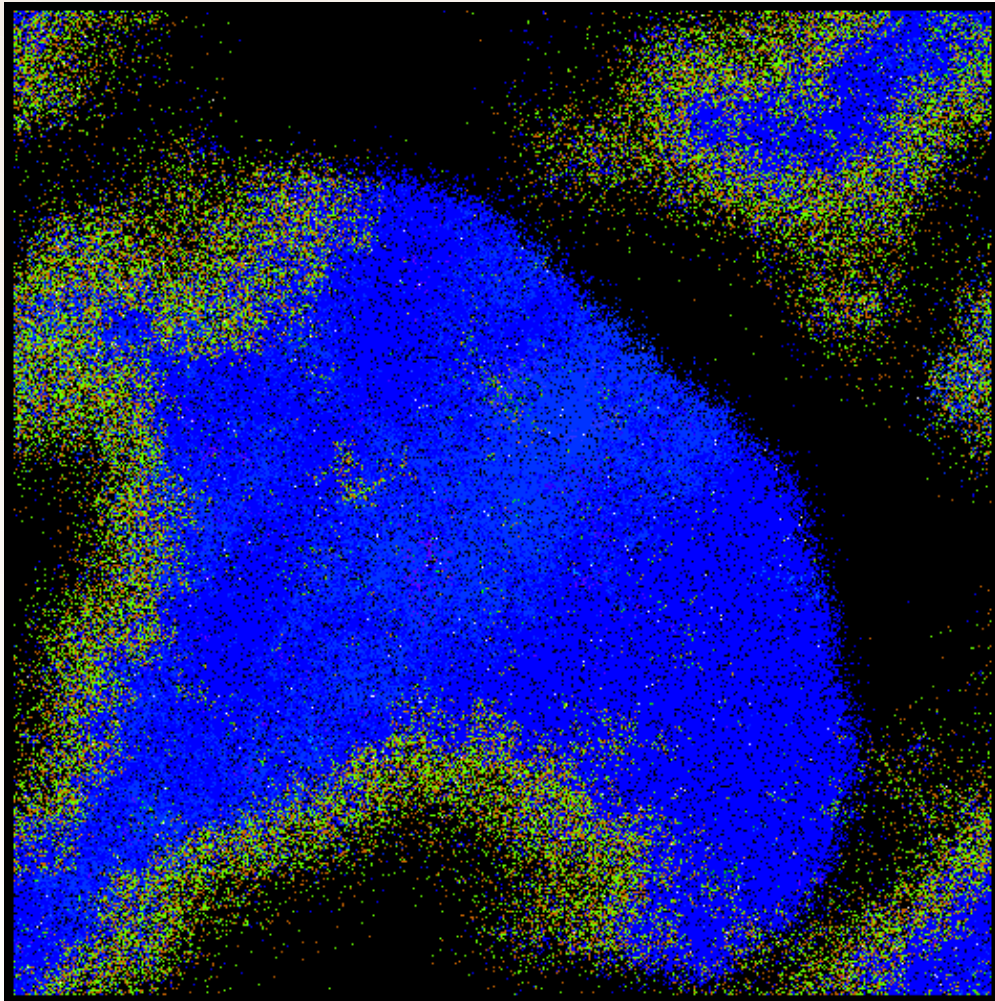
Time = 3.6×10^6	GREEN original μ BLUE lower μ (4 loops)	Length
<ul style="list-style-type: none">• For each new loop: $\mu_{\text{new}} = 0.5 * \mu_{\text{old}}$• Length: ~ 100 nt (increasing)		
<ul style="list-style-type: none">• For each new loop: $\mu_{\text{new}} = 0.8 * \mu_{\text{old}}$• Length: ~ 75 nt (stable)		

Evolving the proof reader: conclusions

- The process for the evolution of the “proof reader” consists of a cycle of:
 - Niche creation due to steep quasi-species exploitation by parasites
 - Speciation of a neutral, longer sequence
 - Loop discovery (lower μ)
 - Steepening of the quasi-species
- In conclusion:
 - A storage-based (single molecule) solution to Eigen's paradox can evolve from ecological complexity
 - The mechanism for the decrease of the mutation rate depends on the feedback between the evolved ecological interactions and the structure of the quasi-species

Aknowledgements: Paulien Hogeweg, Nobuto Takeuchi,
Folkert K. de Boer, Thomas Cuypers,
my fellow master students for the *moral* support

Thanks for your attention



Gauguin, The wave.