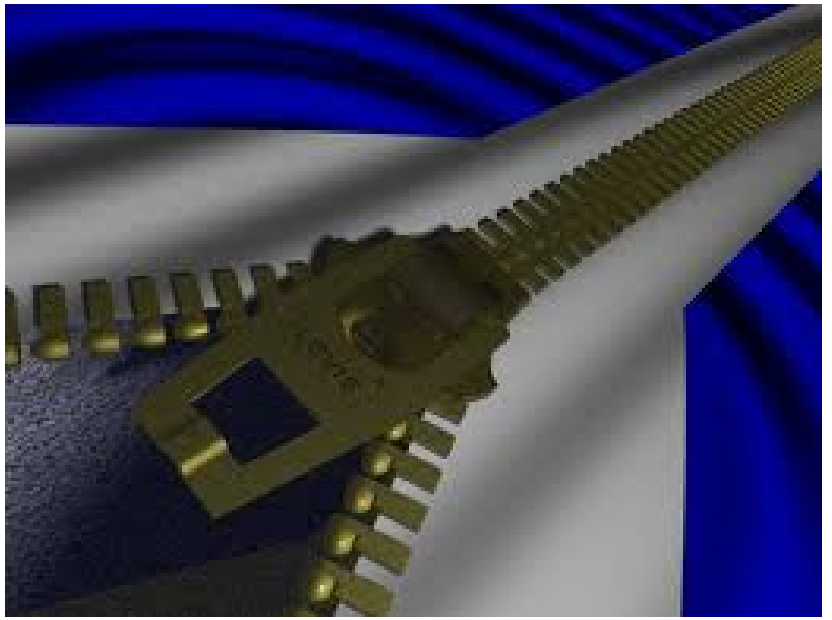
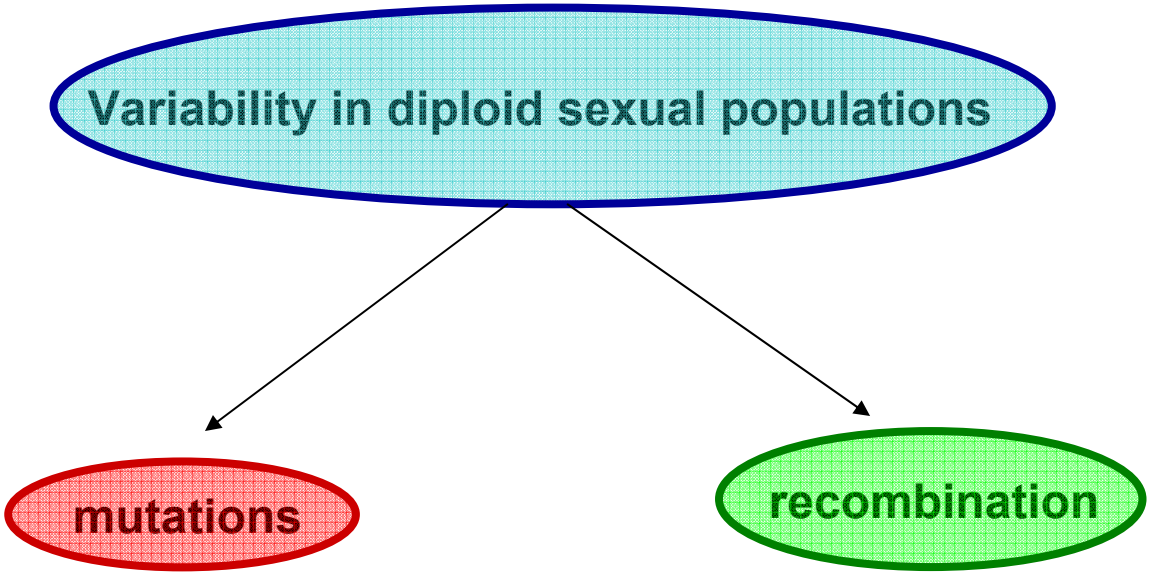
The Zipper effect: Why different positions along the chromosome suffer different selection pressures

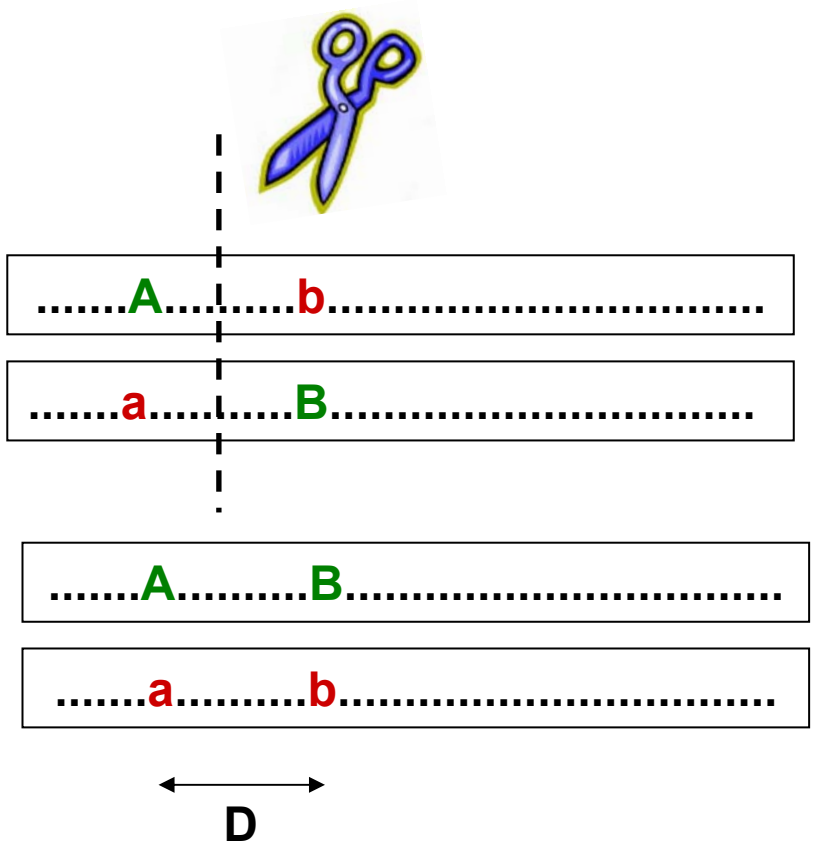
P.M.C. de Oliveira, S. Moss de Oliveira, Physica A 390 (2011) 492

INSTITUTO DE FÍSICA
Universidade Federal Fluminense





We propose that loci near the cromosome tips are more effectively cleaned by recombination than those near the chomosome centre → clusters of neighbouring, orchestrated functioning genes, supposed to be more robust against genetic deffects, are more likely to be found at the chromosome centre.



A and B – wild type

a and b - mutated

The larger the distance D is, the larger is the probability of cleaning both loci on the corresponding gamete. Therefore, a spatial correlation along the genome emerges as a consequence of the crossing mechanism.

❖ Some genes are not a single piece of adjacent bases but are formed by a certain number of such pieces which we refer to as sub-pieces.

❖Although separated, the sub-pieces forming a gene (or a cluster of independent genes) are located at the same region along the chromosome.

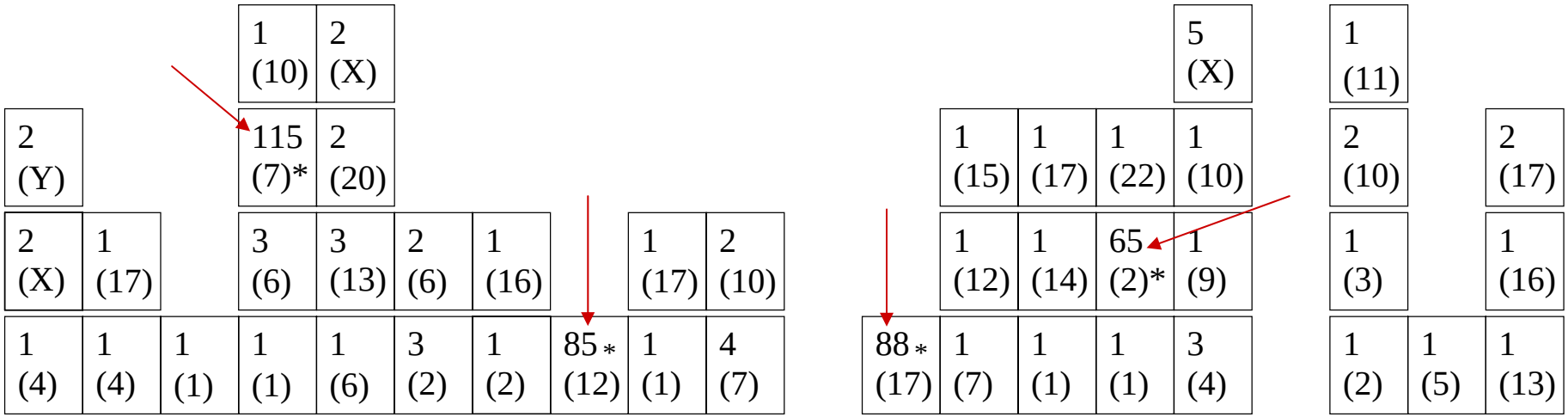
❖ In general the expression of these multi-loci genes are extremely related to the regulatory mechanism, as the HOX genes, responsible for the most primitive and fundamental features of embryogenesis.

❖We took from http://genome.ucsc.edu/ the initial and final positions of 417 HOX genes along the 24 chromosomes.

❖ Since chromosomes differ in size, we divide each one into 20 adjacent segments, each one containing 5% of the chromosome`s lenght. Any HOX-gene is entirely located inside such a segment.

❖We counted how many HOX-genes were found in the first segment of each chromosome, then in the second segment, etc.

Distribution of HOX genes



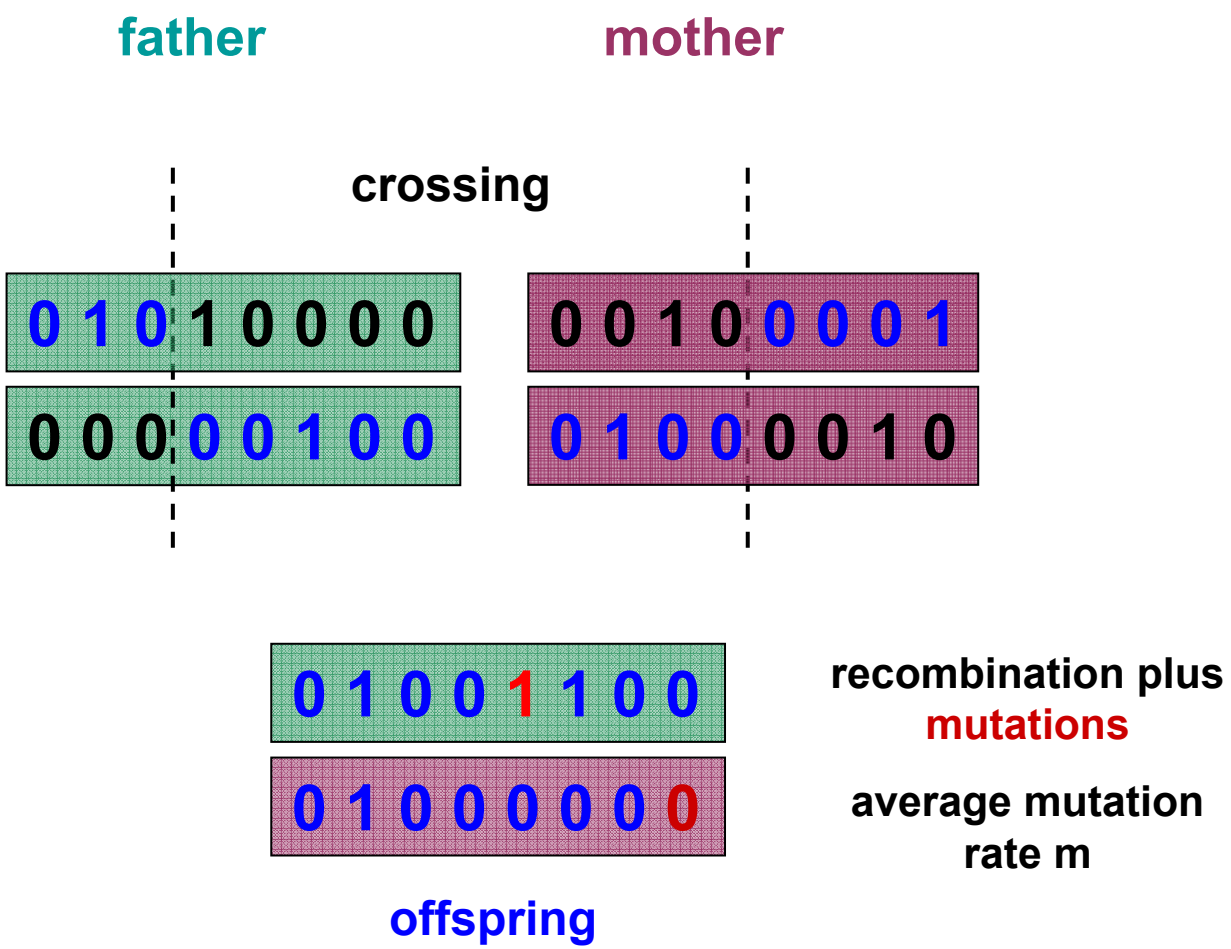
❖ **Along the horizontal direction, each of the 24 human chromosomes were divided into 20 adjacent segments. Each retangular box displays the number of HOX genes inside those segments and the number or letter of the corresponding chromosome inside parentheses.**

❖ **Clusters of many orchestrated-functioning HOX genes can be recognised by the large numbers of genes inside the same box.**

Conclusion: clusters of multi-loci, multi-function genes, for which mutations were shown to be[1] more likely to help in adaptation to the current environment tend to be located at the centre, while single-locus genes for which mutations rarely improve their genetic functionality tend to be at the tips. That is, since mutations are randomly distributed along the chromosome, the crossing mechanism turns out to be the evolutionary ingredient locating some genes preferentially near the centre and others near the tips.

1 D.J. Earl, M.W. Deem, Proc. Natl. Acad. Sci. USA 101 (2004) 11531.

The Model (L=8)


father
mother
crossing
0 1 0 1 0 0 0 0
0 0 0 0 0 1 0 0
0 0 1 0 0 0 0 1
0 1 0 0 0 0 1 0
0 1 0 0 1 1 0 0
0 1 0 0 0 0 0 0
recombination plus
mutations
average mutation
rate m
offspring

- The wild allele 0 is considered dominant;
- The mutated deleterious allele 1 decreases individual survival only if it is present at the same locus in both bit-strings;
- The survival probability is given by: $S = X^{N+1}$ where N is the number of 1-1 loci along the individual genome;
- X is a parameter slightly smaller than 1 representing selection pressure; the larger the X is, the weaker is the selection mechanism;
- At each time-step each individual is killed with probability $P = 1-X^{N+1}$;
- After the death mechanism, still at the same time step, the survivors breed in order to restore the population size of 1000 individuals;
- The parents of each offspring are selected at random;
- The initial population has only wild alleles.

✓ **Extinction may occur if there is less than 2 survivors;**

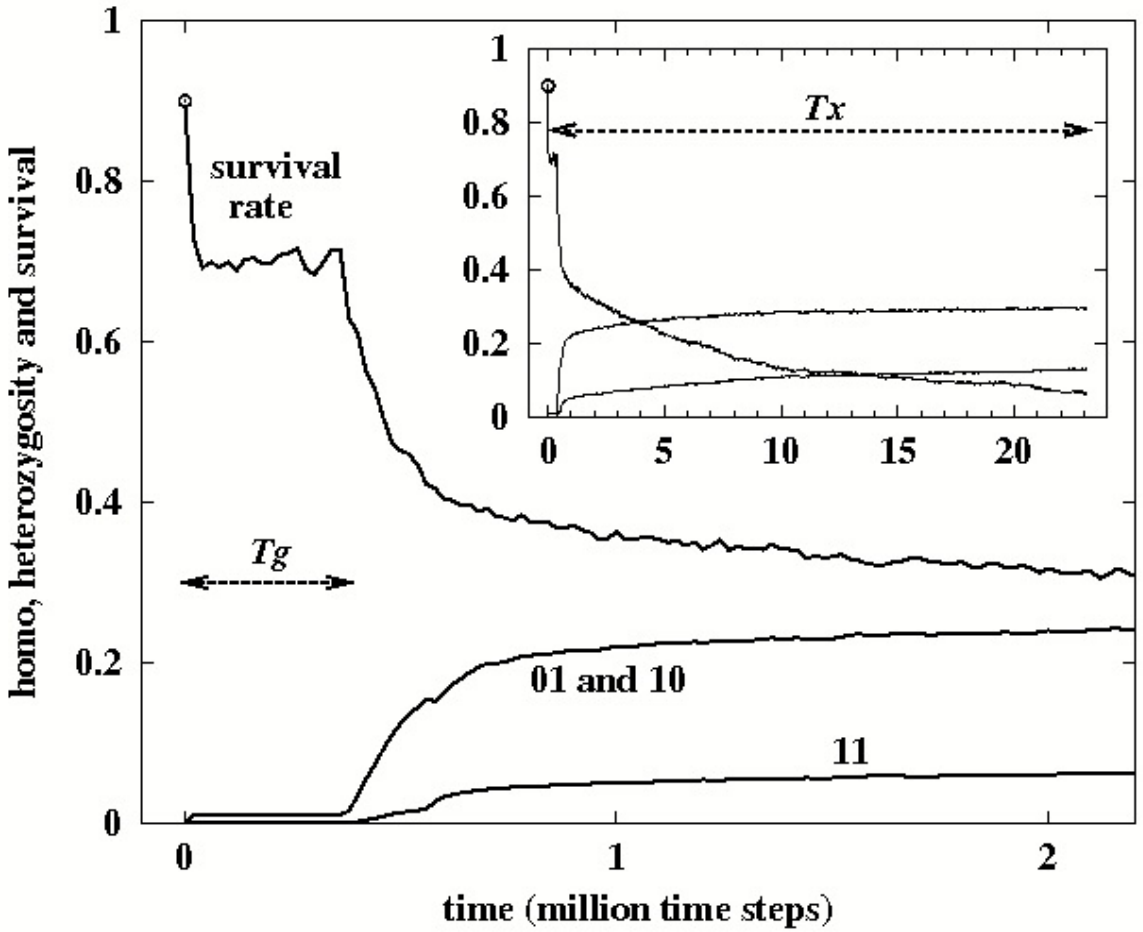
✓ **Population always survives[2] for m < 1; but it can also survive for an extremely long time for m ≈ 1.2, when extinction may appear or not for different random seeds;**

❖ **We fixed m = 1.2 and followed the evolution of several populations differing only by the initial random seed.**

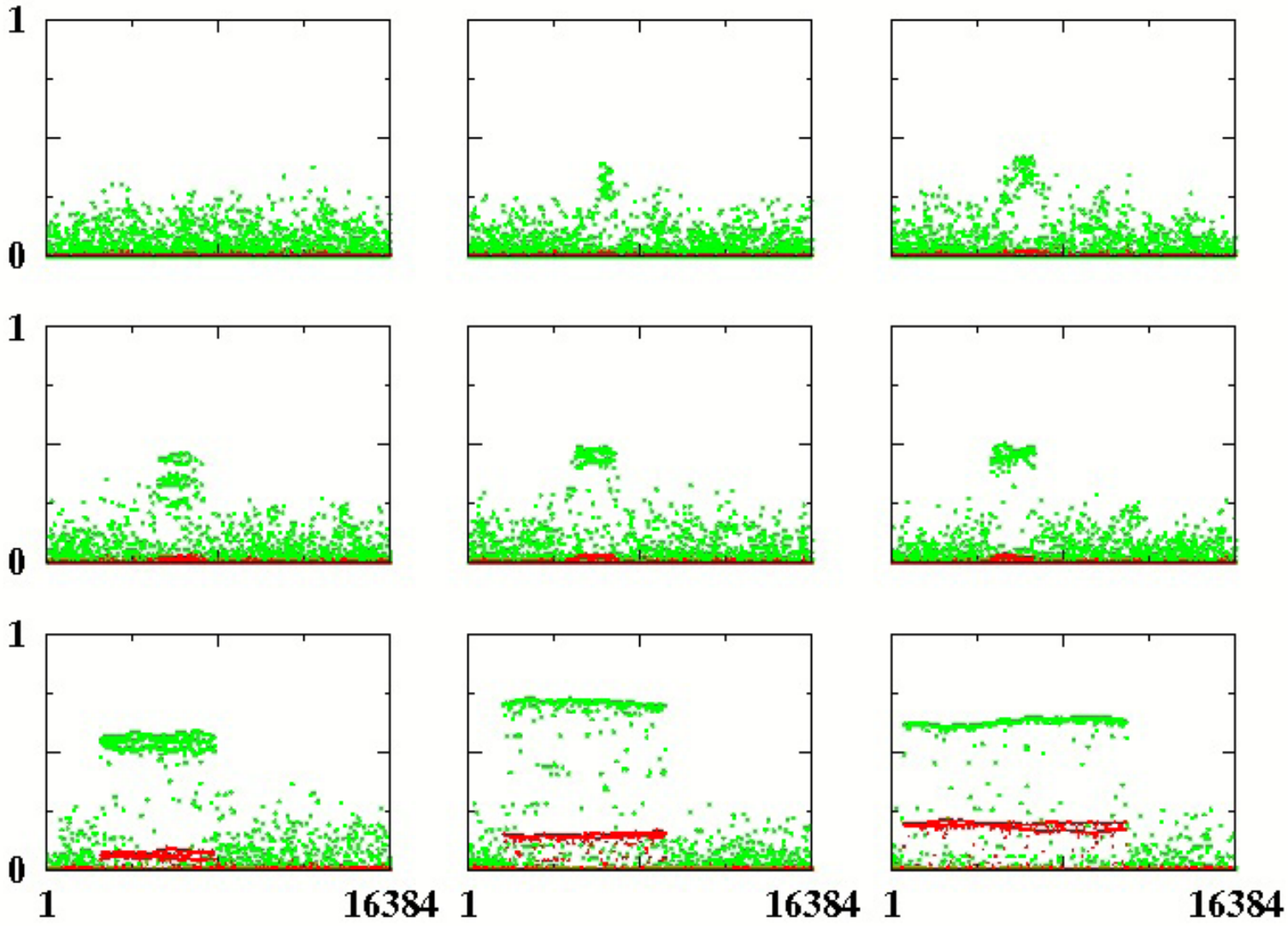
2 PMC de Oliveira, SMO, D Stauffer, S Cebrat and A Pekalski Eur. Phys. J. B 63 (2008) 245.

N=1.000; M=1.2; L=16.384; X=0.9;

T_X = extinction time = 23.128.835; T_g = time when genetic degeneration suddenly starts and continue until T_X



Vertical axis: fraction of loci; Horizontal axis: loci positions
green → heterozygous loci (01 or 10); red → homozygous 11



❖ **Tg depends on the genome lenght:**

for L = 16.384 T_g is generally between 1 million and 10 million time-steps;

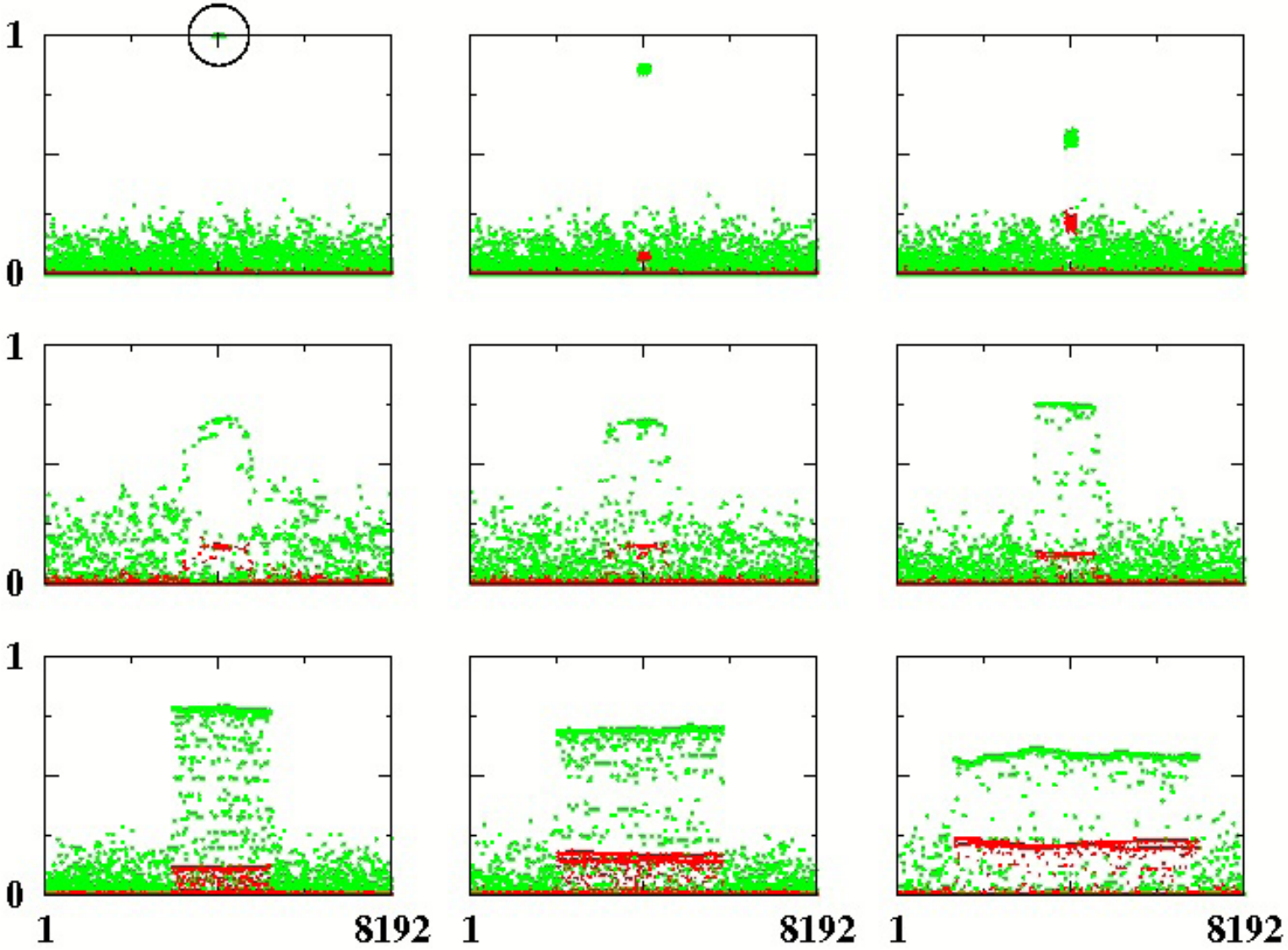
For L = L_{max} = 32.768 T_g < 1.000.000;

For L = 8192, 4096 and 2048 the expected degeneration didn`t appear after 300.000.000 time-steps while all our runs for L = 16.384 showed mutation meltdown;

➢ **Near the extremities of the linear genomes the degeneration is inhibited: the front propagation speed towards the tips slows down.**

Instead of waiting a long time until genetic degeneration spontaneously starts, we decide to induce it artificially for L = 8.192. We allowed the population to evolve normally for 1 million time-steps and then we toss 128 among the 256 loci located around the genome centre, and randomly set these loci to 01 or 10 for all individuals.

T=0; 1; 10; 100; 1.000; 10.000; 100.000; 1.000.000; 10.000.000.



Questions to think about:

1. **Why do we have 23 chromosome pairs instead of a single long one?**

2. **Why diploid genomes are linear instead of circular like bacteria?**

3. **Why genetic material seems to be more robust against degeneration near chromosome tips, compared to the centres?**