

Evolution of RNA Genomes from a Quasispecies Perspective

UNIVERSITÄT LEIPZIG



Nabor Lozada-Chávez¹, Irma Lozada-Chávez², Marco V. José¹

¹ Institute of Biomedical Researches, UNAM, MEXICO

² Interdisciplinary Center for Bioinformatics (IZBI),
UNIVERSITY OF LEIPZIG, GERMANY
nlozada@biomedicas.unam.mx, ilozada@bioinf.uni-leipzig.de
marcojose@biomedicas.unam.mx



References:

- [1] Forterre, 2006. *Vir. Res.* 117:5-16.
- [2] Eigen and Schuster, 1977. *Naturwissenschaften.* 64:541-565.
- [3] Domingo et al., 2005. *Vir. Res.* 107:129-139.
- [4] Bull et al., 2005. *PLoS Computational Biology* 1(e61):0450-0460.

OVERVIEW

Viruses are biological systems that have played an important role in the early evolution of life, in the generation of living genetic diversity and the evolution of contemporary cellular systems [1]. However, the evolution of viruses is itself a complex issue of study due to its controversial nature: their dependency of the cells to replicate, a “simple” genomic composition and a huge diversity of mutation rates and mechanisms developed to infect almost any living cell on Earth.

In 1977, Manfred Eigen and Peter Schuster proposed the **quasispecies** model (see Equation 1) [2], which has been applied to understand the evolution of RNA viruses, among other genetic phenomena such as the early steps in the origin of life. In general terms, *quasispecies* are clouds of genotypes (i.e. master sequences) that appear in a population at mutation-selection balance. Thus, viruses when modeled as *quasispecies* can be conceived as a distribution of replicating, small and mutant genomic sequences which reach an equilibrium state of infinite size by processes of high mutation rate and fitness-selection from a finite initial population [3].

In this work, we have created a genetic algorithm (Figure 2) that incorporate additional biological “realism” into the basic models in order **to assess the evolution of Human immunodeficiency virus 1 (HIV-1) as quasispecies**. Therefore, we are evaluating the impact of the genome size, effective population size and differential mutational fitness on our evolutionary dynamics. From our modeling, we are able to measure **the most novel properties of quasispecies** [4]:

- a) **Eigen limit or error threshold**, a mutation rate below which populations equilibrate in a traditional mutation-selection balance and whereby the favored genotype can maintain the replication of information despite high mutation rates (Figure 3C).
- b) **Error catastrophe**, the phase transition above the mutation rate whereby the loss of the favored genotype or viral identity is carry out through frequent deleterious mutations (Figure 3B and 3D).
- c) **Extinction catastrophe**, the complete loss of the genotypes of the viral population through lethal mutations (Figure 3A).

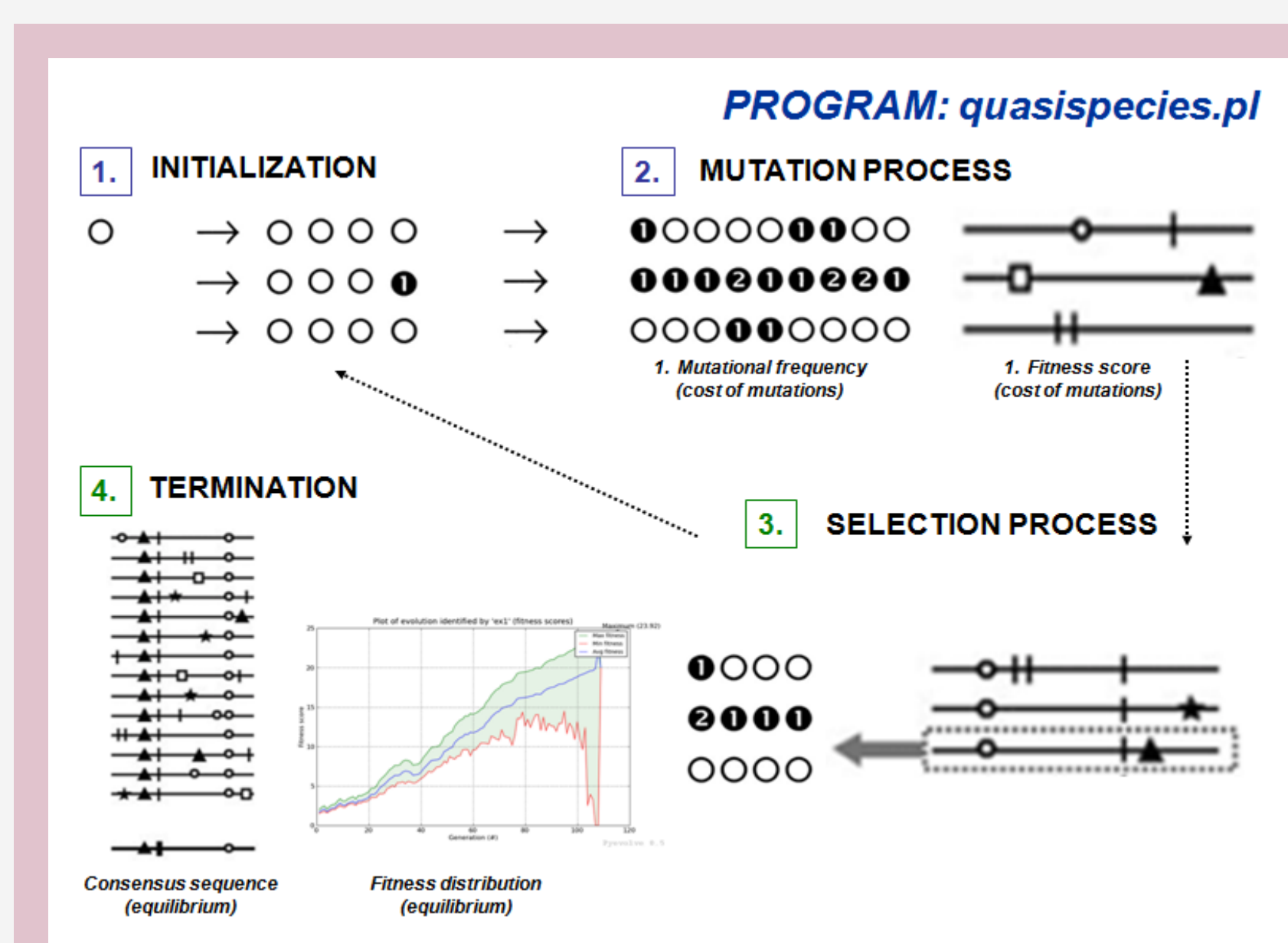
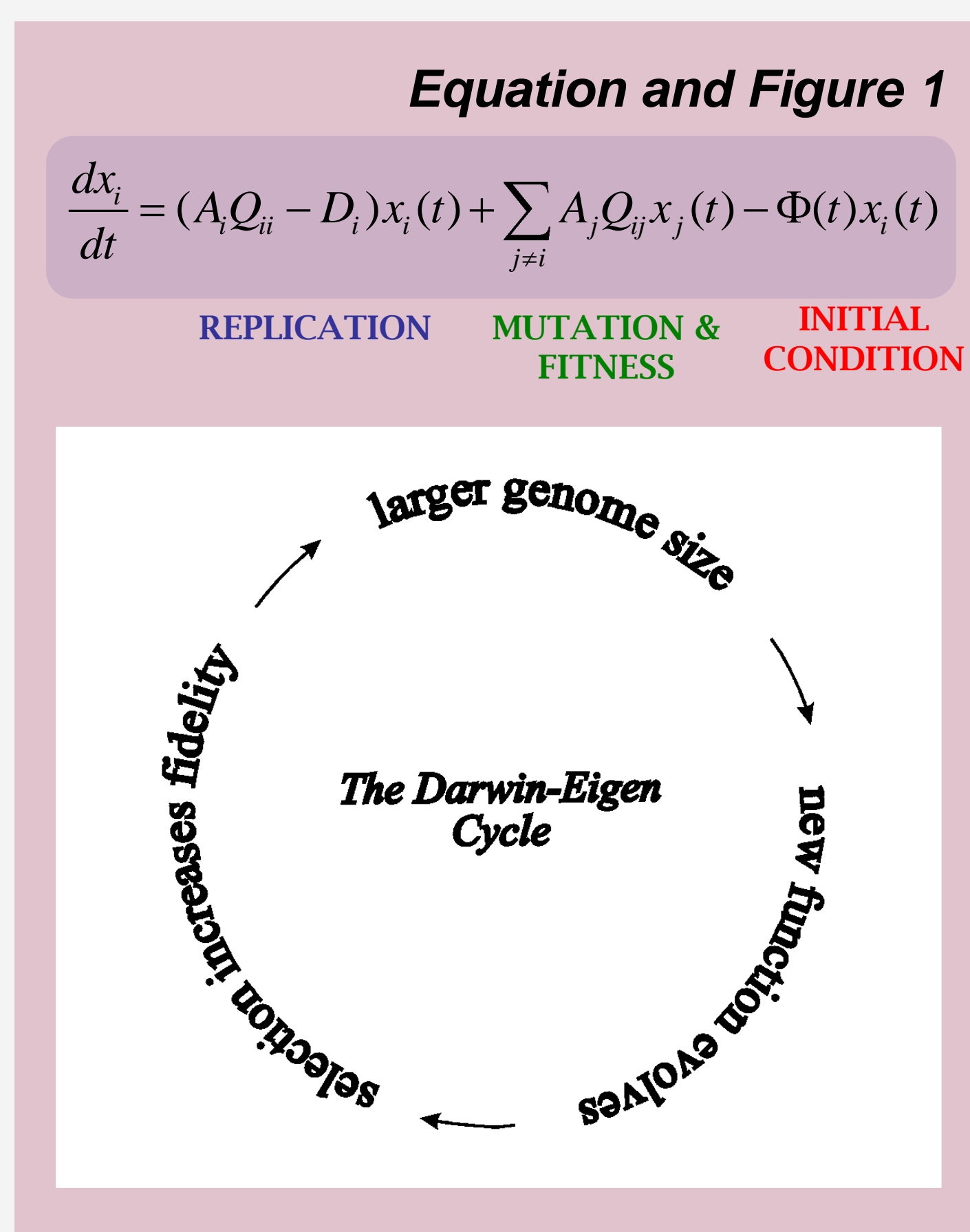
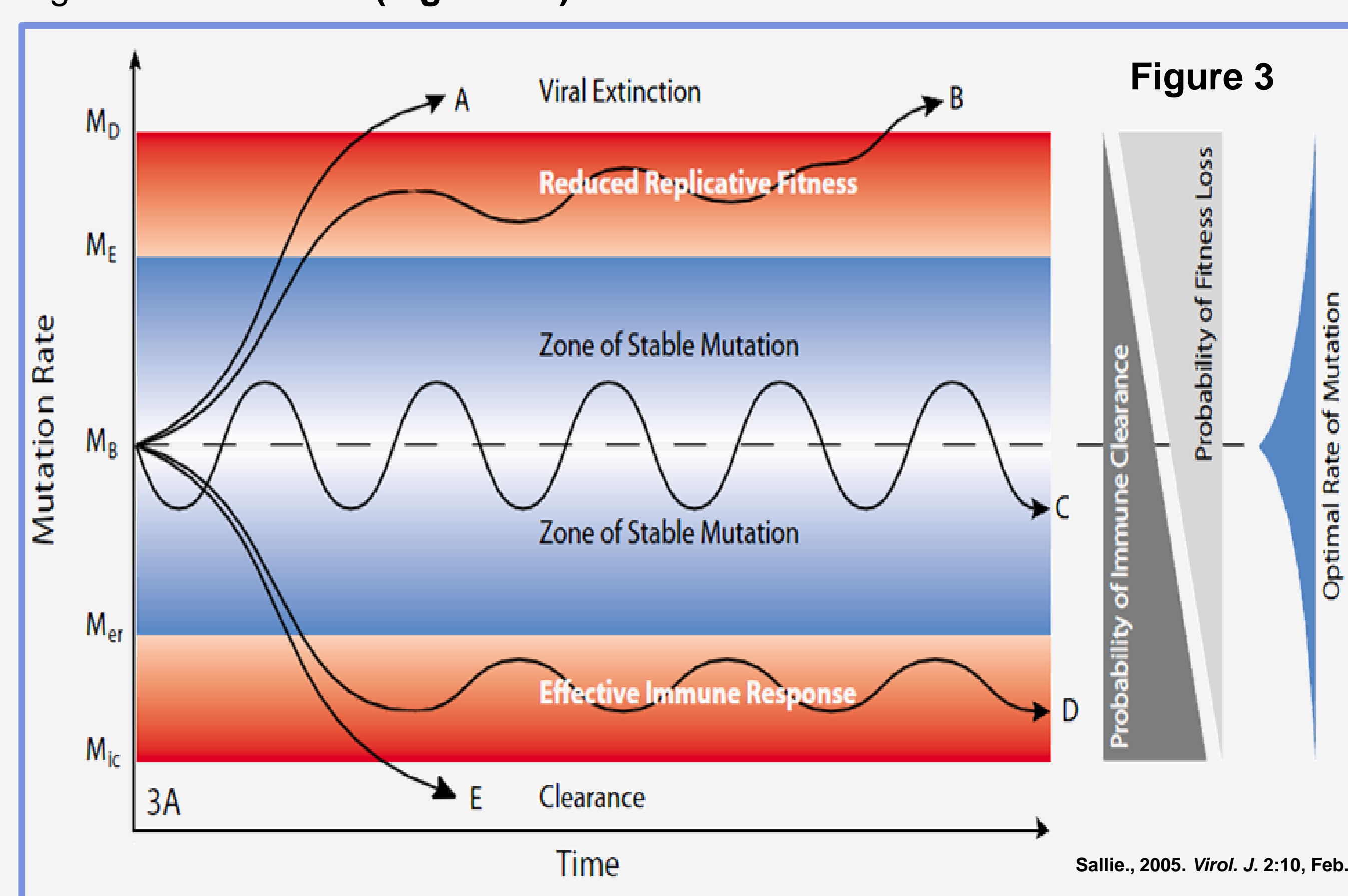


Figure 2. Biological Parameters for the algorithm:
Mutation rates: 1 mut/site/rep.
Freq. nt/mutation: A=13%, T=23% C=24% G=40%
Effective population size: 10,000 indiv.
Production of virions: (10%) 1,000 particles
Type of mutations: in/del; Ds and Ns; frameshift.

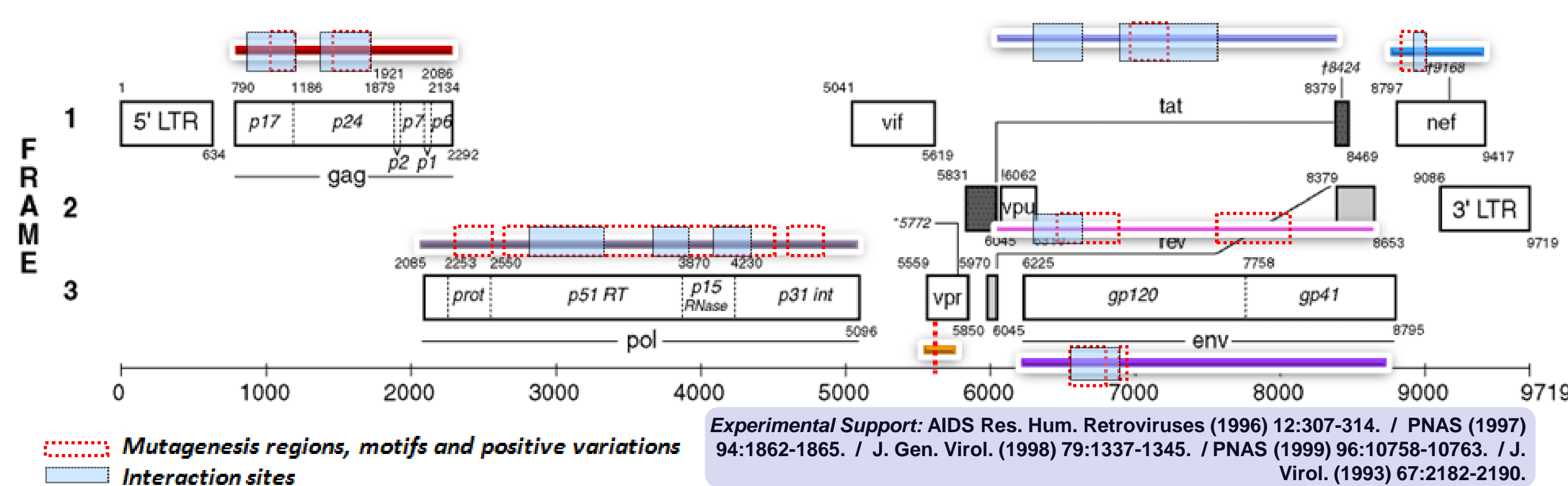


METHODOLOGY

Estimation of a MUTATIONAL FITNESS for HIV-1 (strain HXB2 ID:K03455)

In order to calculate a **mutational fitness** (positive, negative or neutral selection) for the **functional regions** and codons in HIV-1 genome, we perform the following steps:

- 1 **Collection of 2364 complete genomes of non-recombinants HIV-1** from 1982 to 2009 in 35 countries (NCBI and Los Alamos databases).
- 2 **Redundancy analysis:** Final data set of 124 viral genomes.
- 3 **Alignments:** ClustalW, Muscle and manual edition (BioEdit).
- 4 **Characterization of functional domain & motifs:** lethal and advantageous mutations in: PFAM, Prosite, Superfamily and Los Alamos databases.



Estimate the Selection in viral genome by:

- 5 **Codons** HyPhy (dS-dN) and DnaSP (Tajima's D)
- 6 **Functional regions** (Tajima's D)

Figure 7 Overview of selective pressure in HIV-1 genome

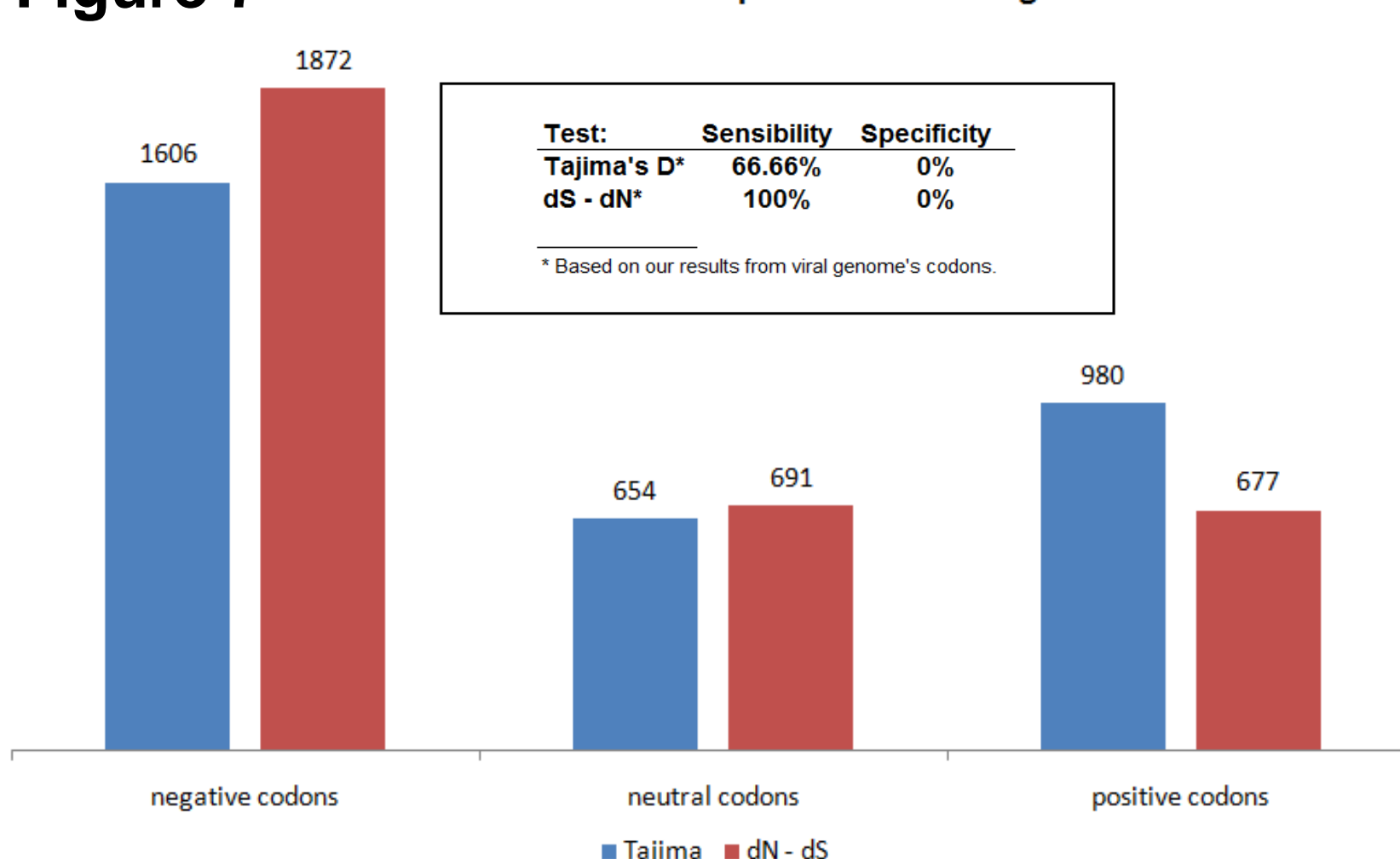
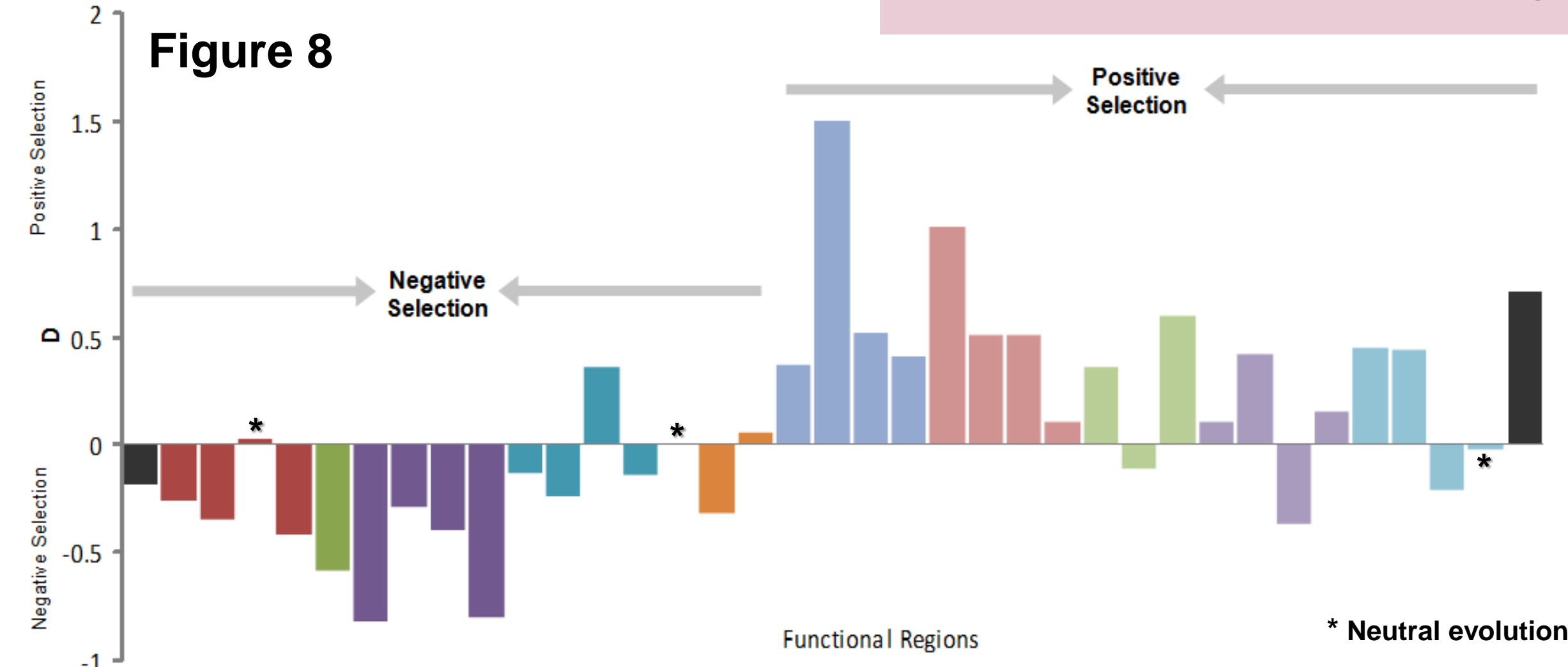


Figure 8



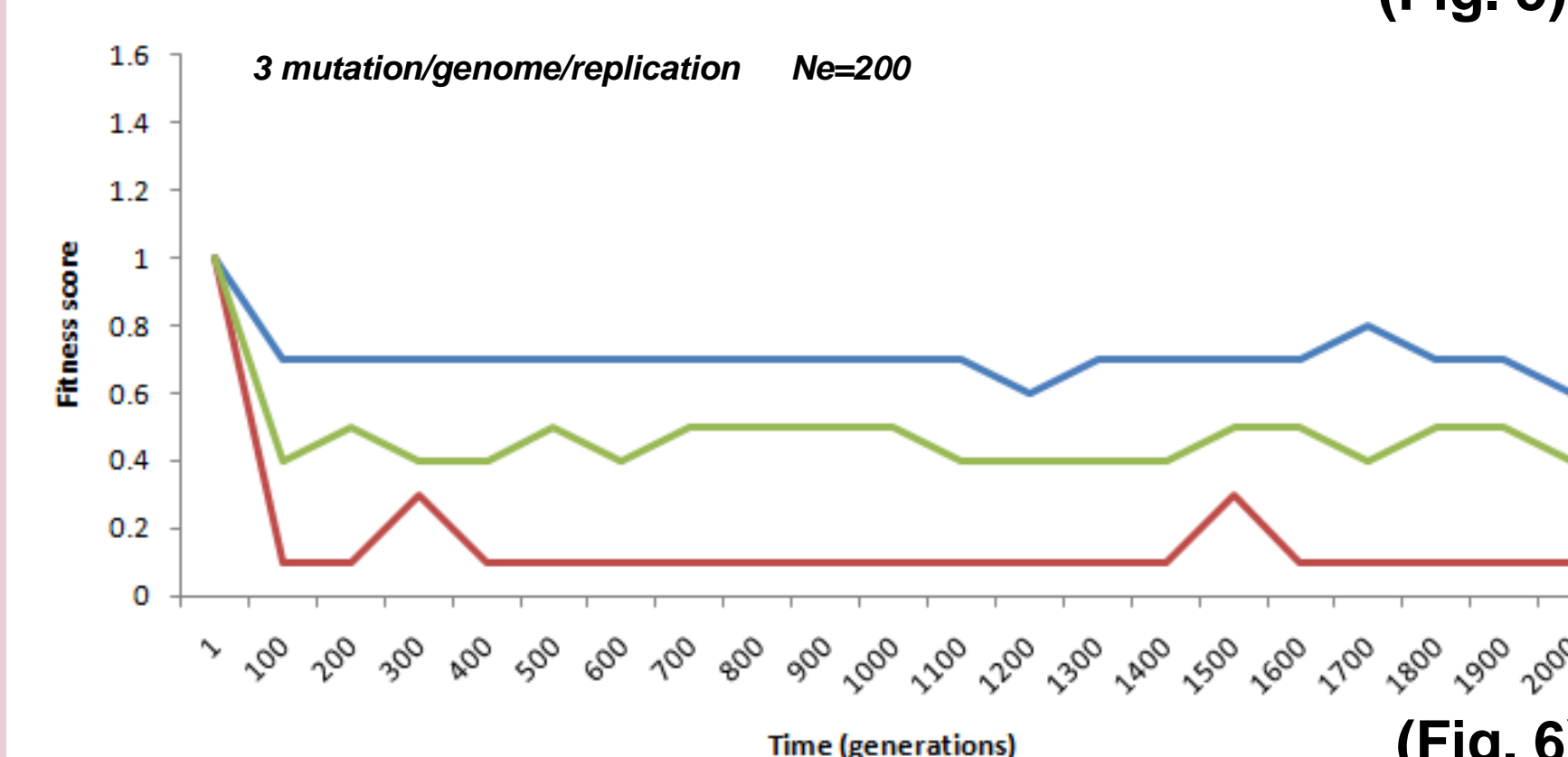
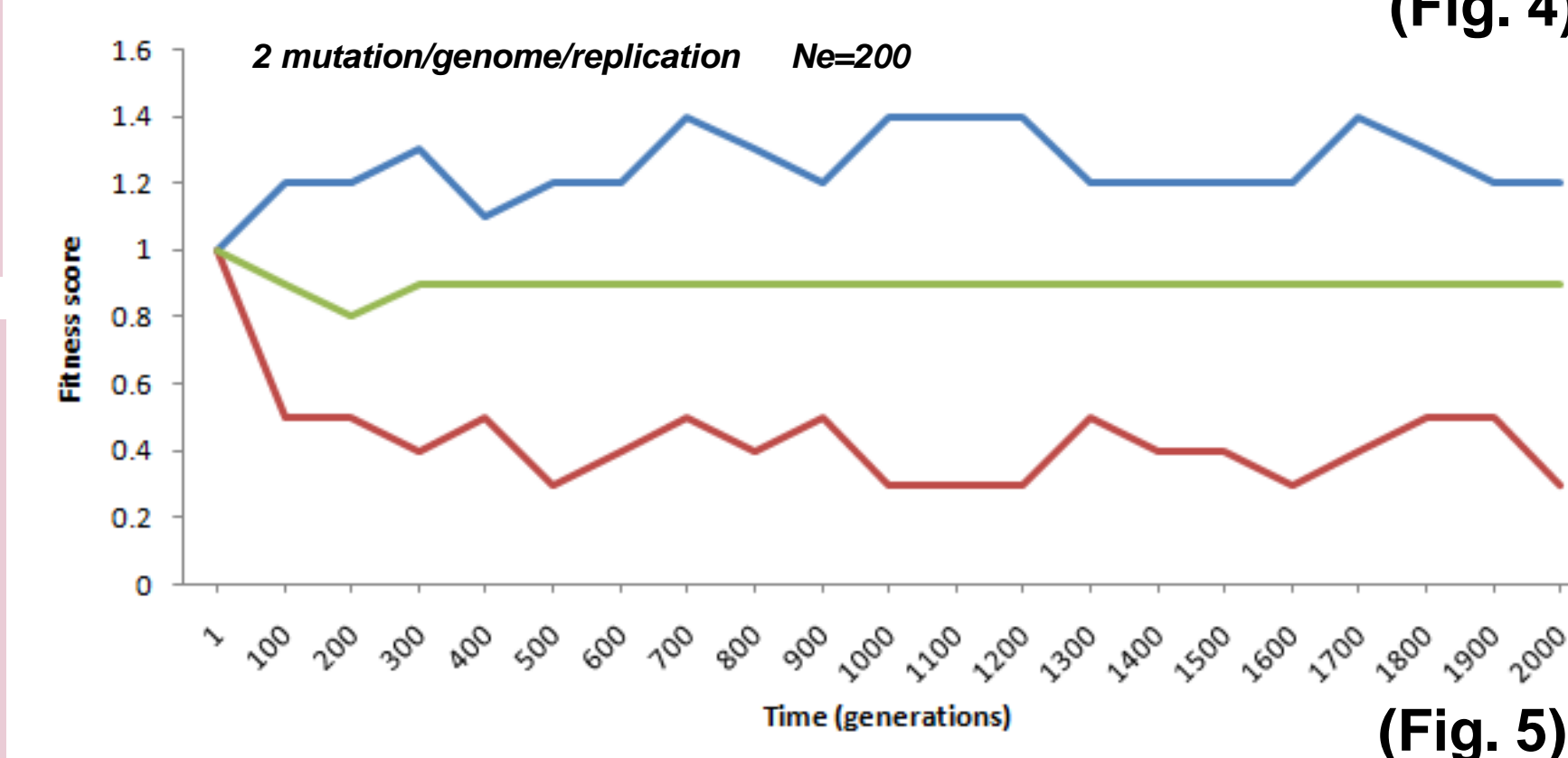
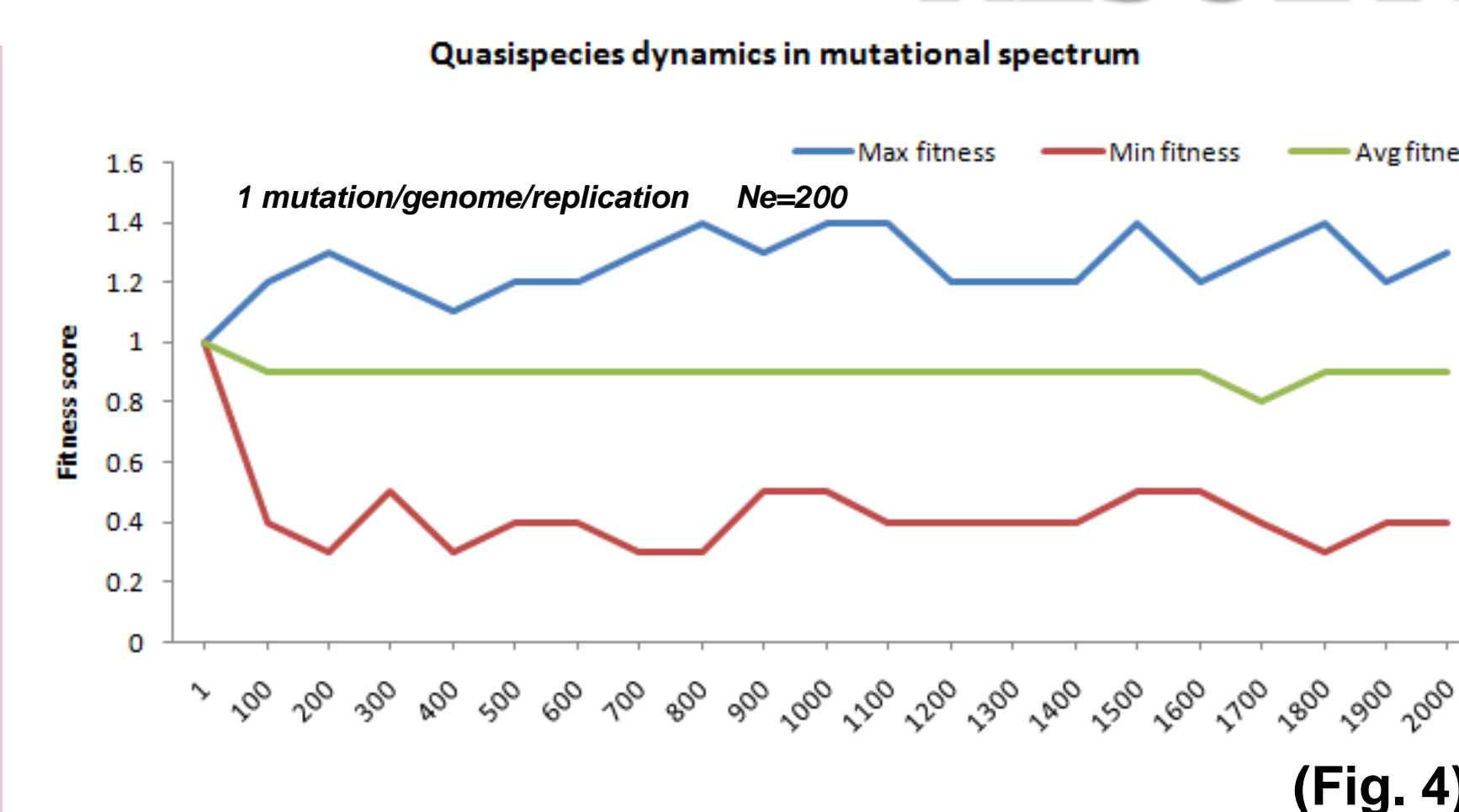
REMARK 1

From our simulations, the presence of an **error threshold** during the evolutionary dynamic of HIV-1 can be observed. A consensus sequence (from the ancestral genome) is maintained with a high fitness at 1 and 2 mutations per genome/replication (Fig. 4 and 5). However, the consensus sequence is lost when the initial population is submitted to replicate with 3 mutations every cycle. Consequently, a drastic decrease of the viral fitness occurred (Fig. 6).

REMARK 2

We suggest that HIV-1 population cannot maintain the genotype with the highest fitness when it crosses the error threshold. Therefore, an **error catastrophe** had place with the lost of HIV-1 quasispecies behavior, favoring genotypes with lower fitness but greater mutational robustness (Fig. 6). We suggest that RNA genomes can keep small genomes when they are subject to high and extended negative selection (Fig. 7), which in turn, it only maintain the replication of few core functions, and of strategic positive selection in those regions encoding interaction with cellular hosts (Fig. 8).

RESULTS



CONCLUSION

Our main contribution to the quasispecies model is to consider the evolutionary impact of purifying selection on the viral dynamics. Based on our mutational fitness approach, it is possible to observe that a high proportion of negative mutations in the viral genomes (Fig. 7 and 6) can drive the evolutionary dynamic of HIV-1 to reach the quasispecies thresholds and faster.