RELATIONSHIP BETWEEN THE SELECTION PRESSURE AND THE RATE OF MUTATION ACCUMULATION

KATARZYNA BOŃKOWSKA, PRZEMYSŁAW BIECEK,
AGNIESZKA ŁASZKIEWICZ and STANISŁAW CEBRAT

Department of Genomics, Wrocław University
Przybyszewskiego 63/77, 51-148 Wrocław, Poland
E-mail: cebrat@smorfland.uni.wroc.pl

Abstract. We use the diploid, sexual Penna ageing model and its modification with noise and environment fluctuations to analyse the influence of random death on the accumulation of defective genes in the genetic pool of populations evolving under different environmental conditions.

1. Introduction. Medawar [4] claimed that the force of natural selection weakens with increasing age. A weaker selection pressure enables the accumulation of mutations in the individuals' genomes in the late expressed genes. Consequences of genetic defects, which appear late in the lifetime, may be unimportant from the evolutionary point of view since they affect a small number of individuals. Nevertheless, they shorten the individuals' lifespans. This phenomenon is known as the mutation accumulation theory of ageing.

The Penna model [5] confirms this theory. The model assumes the chronological activation of genes during individuals' lives, which leads to the accumulation of defective genes expressed late during life. For all organisms, living in the natural world is connected with inevitable deaths due to restrictions of resources or diseases and predators. In the Penna model, the size of the population is regulated by a parameter which can operate only on newborns or the whole population. In the first strategy, there are no random deaths during lifetime—the Medawar's assumption, that random deaths are responsible for the weaker selection of individuals due to their older age, disappears. The gradient of accumulated defects and specific age structure of population still persists [3].

In the modified Penna model [1] the health status of an individual is described by variance of its fluctuations. Each expressed deleterious mutation increases the fluctuations. Additionally, the fluctuations of the environment are superimposed on the fluctuations of individuals in the population. The individual dies if the combination of both stochastic processes trespasses the limit (level of homeostasis) set as the model parameter. The genes are switched on chronologically leading to the accumulation of defective genes expressed during the late periods of life in the genetic pool of the population. A decrease of the variation of the environmental noise, corresponding to the decrease in the selection pressure, increases the average expected lifespan of individuals.

2. Model. In the sexual version of the Penna model, individuals are represented by two strings of bits of declared length L. The genes are represented by single bits. There are two possible states for a gene. 0 means that the gene is correct, while a mutated one is represented by 1. We have assumed that all defects are recessive. There are no reversions.

In the model genes are switched on chronologically which means that in the first Monte Carlo (MC) step after birth, the first locus of the genome is expressed and checked whether it is correct or not. In the second MC step, the second locus is expressed and so on. If the number of expressed defective loci reaches the value of a parameter T - the individual dies. Additionally, each individual may be killed by the environmental restrictions. In the model, the space and food limits are introduced as the Verhulst factor:

$$V = 1 - N_t/N_{max}, \tag{1}$$

where V is a probability of survival, N_t is a current population size while N_{max} is a maximum population size, called also the maximum capacity of the environment.

If the individual stays alive until the minimum reproduction age R, it generates an offspring with declared probability B. In the female genome, one locus is randomly chosen with probability C, bitstrings at this locus are cut and homologous fragments are exchanged. A new mutation is introduced to the gamete with probability M. Then, the male individual is randomly chosen and the same processes happen during its gamete production. The zygote is formed by joining female and male gametes. Sex of the newborn is established with an equal probability for male and female.

Let N_{Rt} denote the number of individuals in the current population in the reproduction age while $N_{ft} = \max(0, N_{max} - N_t)$ denotes the number of "free slots" in the population. The number of newborns follows the distribution

$$\mathcal{B}\left(\mathcal{B}(N_{Rt}, B), \frac{N_{ft}}{N_{max}}\right) \underset{N_{max}\to\infty}{\longrightarrow} \mathcal{N}\left(N_{Rt} * \frac{B * N_{ft}}{N_{max}}, N_{Rt} * \frac{B * N_{ft}}{N_{max}} * \left(1 - \frac{B * N_{ft}}{N_{max}}\right)\right), \tag{2}$$

where $\mathcal{B}$ signifies the binomial and $\mathcal{N}$ the Gaussian distribution.

Thus, the expected number of newborns is $N_{Rt} * \frac{B*N_{ft}}{N_{max}}$. If $N_{Rt} << N_{max}$ the population will become smaller and smaller and in the end it may extinct.

Because the Penna model assumes chronological activation of genes during an individual's lifespan, it leads to the accumulation of defective genes expressed late during the lifespan. Genes switched on before reproductive age are under strong selection pressure and the fraction of defects in this set of genes is low. Genes switched on after the mini-

mum reproductive age are under weaker selection pressure and an increasing fraction of defects is observed.

To introduce the role of the environment, we modified the diploid Penna model in which each individual is represented by a pair of bitstrings. The state of the organism and the environment are determined by the value of their fluctuation [1];

$$P_i(t) \sim \mathcal{N}(0, \sigma_i^2(t)), \qquad E(t) \sim \mathcal{N}(0, \sigma_e^2). \tag{3}$$

The amplitude of the fluctuation depends on the genetic background of the organism

$$\sigma_i^2(t) = \sigma_0^2 + \sum_{j=1}^{age(i,t)} g_{ij}\sigma_d^2. \tag{4}$$

where σ_d is the increase of variation caused by a defective gene; σ_0 is the base variation of the individual state (base penalty) while g_{ij} is the number of defective pairs of alleles in loci j for individual i. Thus, the more defective loci the higher is the variance of the individual's state and higher the risk of the individual's death, since individual i is killed if the sum of both states crosses the homeostasis limit F [1];

$$P_i(t) + E(t) \geq F. \tag{5}$$

There is a similarity between parameter T in the standard Penna model and parameter F in the considered model. Both parameters define the maximum acceptable harm in the individuals, in the standard Penna model it is simply the number of allowed defects, since in the modified model it is the maximum deviation from the expected environmental state.

3. Results and discussion. As it was mentioned above, the role of the Verhulst factor is to prevent the unlimited growth of the population. The Verhulst factor is responsible for the random deaths of individuals. In the first version, the Verhulst factor operates during the whole life independently of age (strategy $Vall$). Nevertheless, these deaths are not completely random considering the genetic structure of individuals. Individuals with a better genetic structure (better fitted to the environment) live longer and they are exposed to the Verhulst factor for a longer time. If the Verhulst factor determines the survival of newborns only (strategy Vb), the individuals are killed with the same probability, independently of their genetic structure. This strategy generates a better genetic structure of the population than the $Vall$ strategy [2] (Fig. 1). We have also observed longer life spans of individuals' and higher fraction of individuals in the reproductive age (Fig. 2).

In the noisy model, the state of the environment randomly fluctuates, thus, even individuals with very good phenotypes may die, but individuals with defective alleles die more frequently due to the higher variation of their states. Therefore, this randomness and the one represented by the Verhulst factor are different. The noisy Penna model allows us to investigate the role of the environment on the population evolution.

We considered two populations, one which exists in the mild environment and one living in the more severe environment. The difference between populations occurs through
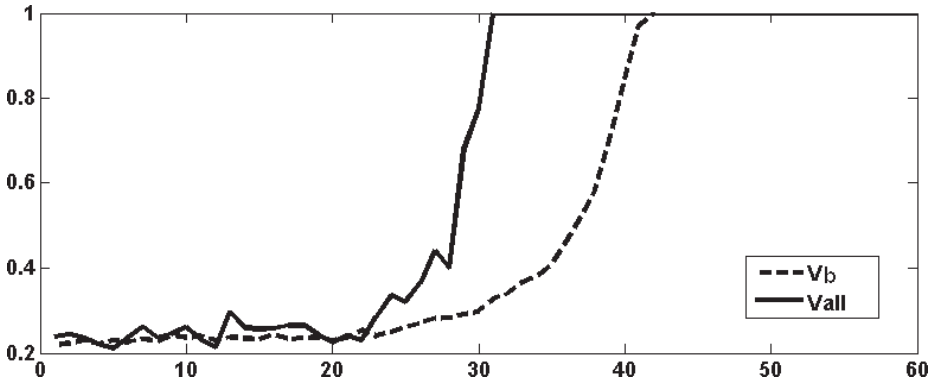
Fig. 1. Fraction of defective genes observed if the Verhulst factor operates only on newborns (Vb) or on the entire population ($Vall$). Parameters of simulations: $L = 64$, $R = 20$, $B = 1$, $T = 3$.
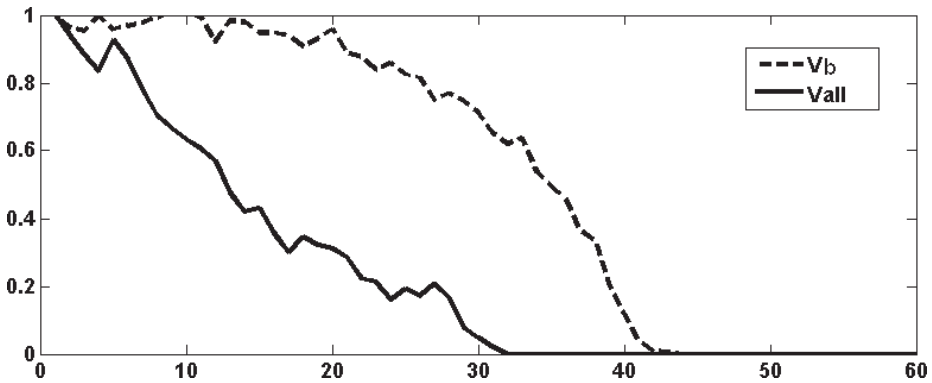

Fig. 2. Age distribution generated by Verhulst factor operating only on newborns (Vb) or on the entire population ($Vall$). Parameters of simulations are as in Fig. 1.

values of σ_e and σ_d. In the more severe environment both values are larger, thus the individuals with the same number of defects die more often in the unfriendly environment.

The distributions of defects in the genomes are presented in Fig. 3. The average lifespan for a population in the friendly environment is longer due to the lower probability of death. On the other hand, the population in the severe environment has a better genetic structure, i.e. fewer defective genes are expressed before reproduction age.

We checked the results of competition between the two populations. After 100000 MC steps, during which these populations evolve separately, we allowed individuals to migrate between the environments. With probability 0.01 each individual could decide to switch its environment and live in different environment than their parents. Mating between individuals from the mild and the severe environments was forbidden, thus, we were able to check which population was better adapted. We expected that the population from the mild environment had smaller chance to survive in the unfriendly conditions, but this population lost (i.e. all individuals in the population died after 100000 MC steps) even
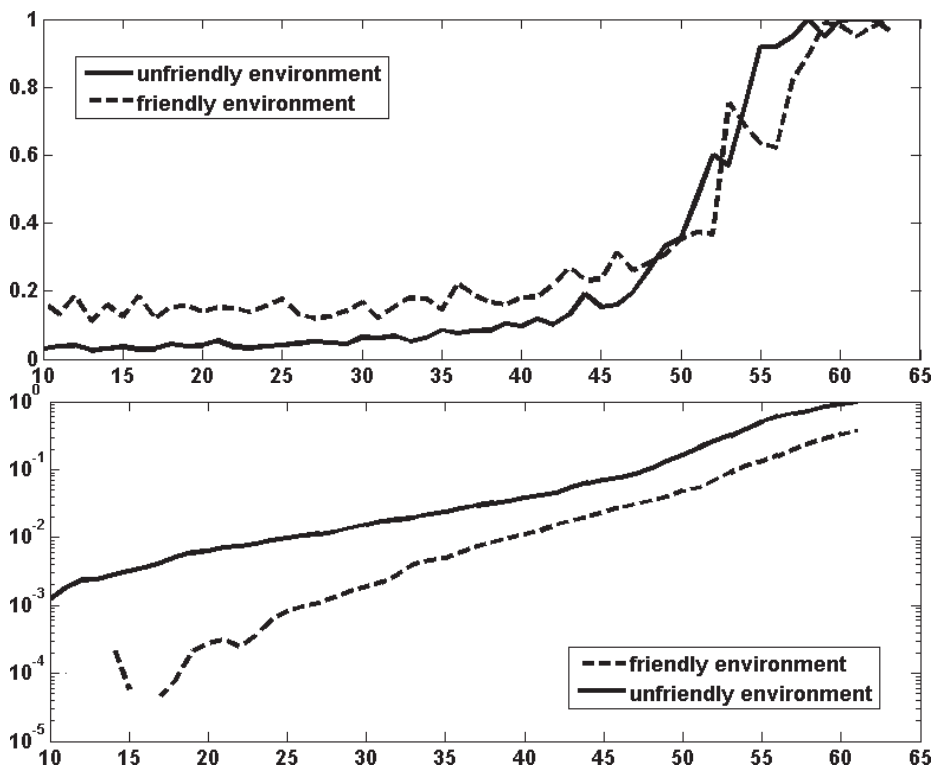
Fig. 3. The distribution of defective genes (top) and mortality (bottom) in populations living in friendly and unfriendly environments.

in the friendly environment in which it was used to live. We repeated these simulations 100 times and in 99 cases the population from friendly environment was eliminated from both environments.

4. Conclusions. Our model confirms the trivial rules that the structures of individuals' genomes in the evolving populations are the results of two counterbalancing forces: selection and mutation pressure. Nevertheless, environment can eliminate individuals randomly or selectively. Random eliminations decrease the speed of adaptation of population and its evolution and its negative impact depends on the age of affected individuals. The most efficient way for the evolution of the fittest genomes is sacrificing the youngest individuals in the random process and imposing the directional selective force on the older members of population.

<div align="center">

References

</div>

[1]   P. Biecek and S. Cebrat, *Fluctuations, environment, mutations accumulation and ageing*, Int. Journal of Modern Physics C, 6 (17) (2006), 923–931.

[2]   A. Łaszkiewicz, *Biological interpretation of the Penna model parameters*, PhD dissertation, 2006.

[3]   J. S. Martins and S. Cebrat, *Random deaths in a computational model for age-structured populations*, Theory in Bioscien. 119 (2000), 156–165.

[4]   P. B. Medewar, *An unsolved problem of biology, an inaugural lecture delivered at University College London*, H. K. Lewis, London, 1952.

[5]   T. J. P. Penna, *A bit-string model for biological aging.* J. Stat. Phys. 78 (1995), 1629–1633.