Journal of Theoretical Biology 260 (2009) 412–421

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Mutate now, die later. Evolutionary dynamics with delayed selection

Andreea Munteanu [a,*], Peter F. Stadler [b,c,d,e,f]

[a] ICREA-GRIB Complex Systems Lab, UPF, Parc de Recerca Biomedica Barcelona Dr Aiguader 88, E-08003 Barcelona, Spain
[b] Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstraße 16-18, D-04107 Leipzig, Germany
[c] Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, D-04103 Leipzig, Germany
[d] Fraunhofer Institut für Zelltherapie und Immunologie – IZI Perlickstraße 1, D-04103 Leipzig, Germany
[e] Department of Theoretical Chemistry, University of Vienna, Währingerstraße 17, A-1090 Wien, Austria
[f] Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, USA

ARTICLE INFO

Article history:
Received 15 February 2009
Received in revised form
15 June 2009
Accepted 24 June 2009
Available online 3 July 2009

Keywords:
Evolution
Neutral networks
Fitness
Extinction
Telomere

ABSTRACT

We analyze here the evolutionary consequences of selection with delay in a population genetics context. In the classical works on evolutionary dynamics, an individual produces off-springs in direct proportion to its fitness, a process in which mutations may occur. In the present scenario of delayed selection, individuals that acquire deleterious mutations can still reproduce unharmed for several generations. During this time delay, the damage passed on to off-springs can potentially be repaired by subsequent compensatory mutations. In the absence of such a repair, the individual becomes sterile. Here we study the population-genetic effects of such a time delay by means of both numerical simulations and theoretical modeling. The results show that delayed selection lowers the extinction threshold, endangering the survival of the population. Surprisingly, however, no traces of this delay effect are encountered in the sequence diversity of the population. These conclusions suggest that delayed selection is hard to detect in genetic data and thus could be a wide-spread but rarely detected phenomenon.

1. Introduction

Darwinian evolution is the interplay of the production of variation and subsequent selection. Due to the complexity of biological organism, selection tends to act at all times, punishing or rewarding small differences among individuals. This is not necessarily the case at the level of (small) genetic subsystems, however. The intuitive rationale for this claim is that an "emergency subsystem", for instance, may not need to be activated for several generations. While unused and inactive, it tends to escape the forces of selection and conceivably, acquire damages. Once conditions change and it is needed again, however, there are severe (fitness) penalties if its functionality has not been maintained or repaired. We expect such "delayed selection" to leave detectable traces in the genome. Hence we study here the dynamical implications of delayed selection in some detail.

It may come as a surprise that the best studied example is a generic component of the eukaryotic replication machinery, namely the reconstruction of telomere ends. Mice deficient for the mouse telomerase RNA (mTR-/-) are fertile and show initially little if any pathologies. However, they can breed only for about six generations due to decreased male and female fertility and to an increased embryonic lethality in later generations. Even late generation (mTR-/-) mice are viable to adulthood, only showing a decrease in viability in old age (Lee et al., 1998; Herrera et al., 1999). These effects appear to be linked to the shortening of the telomeres (Verdun and Karlseder, 2007). Similar effects can be observed in cell culture, again establishing a relationship between viability and telomere length: Terc-deficient embryonic stem cells show gradual reduction of growth rate after about 300 divisions, and proliferation virtually stops after 450 generations (Niida et al., 1998). At the same time, telomerase RNA exhibits extremely high rates of evolution (Xie et al., 2008). The speculation that delayed selection may be part of the explanation for the unexpected evolutionary plasticity of telomerase RNA motivated this work.

Delayed selection is also likely to occur in species for which environmental conditions vary periodically at timescales longer than generation time. A spectacular example is the monarch butterfly (*Danaus plexippus*) (Urquhart, 1960). The "migratory" generation migrates from Eastern North America to overwintering sites in Mexico. This long-lived generation is characterized by reproductive diapause persisting until next spring, when the butterflies reproduce and start the journey back north. Another two to three generations of reproductively competent, short-lived "summer" butterflies follow the progressive, northward emergence of milkweed. Significant differences in gene expression between summer and migratory butterflies (Zhu et al., 2008) suggest that some parts of the butterflies genetic system may be unused over a few generations. Whether this is indeed the case

* Corresponding author.
E-mail addresses: andreea.munteanu@upf.edu (A. Munteanu), studla@bioinf.uni-leipzig.de (P.F. Stadler).

0022-5193/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2009.06.022

could be tested directly if characteristic genomic fingerprints of delayed selection can be detected.

A more subtle context in which delayed selection may play a role is that of synthetically lethal genes. A pair (or a larger set) of genes is called *synthetically lethal* if knocking out the entire set is lethal, while the knockout of all smaller subsets retains viability (Hartman et al., 2001; Kaelin, 2005; Le Meur and Gentleman, 2008). Note that synthetically lethal gene pairs typically share their primary function but cannot be redundant in all their functional aspects. The reason is that exact redundancy is evolutionarily unstable: it is quickly resolved by the loss of one copy (Force et al., 1999). This type of genetic buffering may, however, delay the detrimental effects of functional loss in one partner until a rarely employed secondary function of the affected gene is required. Again, a recognizable signal in the genomic DNA would be of utmost interest.

The paper is organized as follows: in Section 2 we introduce the methodology and the results of the stochastic simulations for a population of RNA molecules. In Section 3, we confront these results with a mean-field model that captures the evolution of the population in a delayed-selection scenario. We quantify the amount of diffusion in the sequence space for various time delays in search of a signature on the evolutionary rates of such altered selection. Finally, we discuss the findings, with special emphasis on the lack of such an inequivocable signature, in the context of genomic studies in Section 4.

2. RNA-based simulations

2.1. A simple model of telomere damage

The simulation framework used in this contribution is motivated by the telomerase RNA (TR) system briefly discussed in the Introduction. For simplicity we distinguished only between fitness-neutral and lethal mutations. Each individual is characterized by its TR gene and the length of its telomere. Off-springs with intact TR have full-length telomeres, while telomeres shrink by a constant amount with each replication step in which the telomerase is inactive. Individuals whose telomeres have shrunk to zero are sterile, i.e., their fitness is set to 0.

In order to include a genetic component with a realistic genotype–phenotype map, we use RNA secondary structures to represent phenotypes. In this approach, each sequence s is folded into its minimum energy secondary structure $\varphi(s)$ and then fitness is evaluated by comparing $\varphi(s)$ with a target structure φ^* (see e.g., Fontana et al., 1989; Schuster et al., 1994; Huynen et al., 1996a). Here, we stipulate that only the target secondary structure is functional. The fitness f of an individual with genotype s and telomere length k is given thus by

$$f(s,k)=\begin{cases}1 & \text{if} \quad \varphi(s)=\varphi^* \text{or} \quad k>0\\ 0 & \text{if} \quad \varphi(s)\neq\varphi^* \text{and} \quad k=0\end{cases} \tag{1}$$

Since the computational effort for RNA folding computations is cubic in sequence length, vertebrate TR gene with 300–500nt are too long to be practical for our simulations. Instead of a real TR structures, we defined an arbitrarily chosen target structure of length 100 to represent the viable phenotype. RNA secondary structure predictions are performed using the `Vienna RNA Package` (Hofacker et al., 1994).

We simulate a population of N individuals in a flow reactor under stochastically controlled constant organization as described in e.g., Fontana et al. (1989). Individuals replicate proportional to their fitness. During replication, each letter is mutated with a probability μ. Then the structure $\varphi(s')$ of the offspring s' is computed. If $\varphi(s')=\varphi^*$, we set $k'=K$, otherwise $k'=k'-1$, where K is the number of generations for which a defective TR is tolerated. In other words, if after K replications, such an incorrect fold has not encountered the neutral network, its fitness becomes 0 and thus loses the capacity of replication.

In the following, we shall discuss the results of the simulations. Based on these data, we introduce a theoretical model associated to the simulation framework, a model that provides a reasonably good fit of the simulation results and also a tool to better understand the implications of the delayed-selection effect.

2.2. Extinction threshold

As a first observation, we notice that one of the consequences of the delayed selection is a reduced critical value of the mutation rate at which the population goes extinct. An erroneously replicating haploid population shows the so-called *error threshold* phenomenon, by which the population loses coherence and quickly approaches a uniform distribution in sequence space as soon as the mutation rate exceeds a critical value. Originally described on single-peak landscapes (Eigen, 1971; Eigen et al., 1989), an analogous phenomenon can be observed at the phenotypic level (Forst et al., 1995; Huynen et al., 1996a; Wilke, 2001). With instantaneous fitness effects, the critical value of μ can be estimated from a μ-dependence of the equilibrium concentration of the "poor" phenotypes.

Before we comment in detail the results obtained for the current framework, we wish to discuss the distinction between error threshold and extinction threshold, a distinction often disregarded in the literature. Extinction can be the consequence of a process such as lethal mutagenesis (Bull et al., 2007), with the latter being a demographic process occurring, for example, in the context of within-host population of viruses that become extinct with an elevated mutation rate. In this case, the population is overwhelmed by deleterious mutations and cannot sustain itself. Eigen's error catastrophe or error threshold, although inspired by the idea of lethal mutagenesis, is a distinct process. The error threshold is defined as the mutation rate beyond which the mean fitness of the population does not decrease exponentially with the mutation rate but remains constant, as all genotypes are insensitive to mutations (the information is lost from the population). Contrary to the intuition, Eigen two-class fitness landscape (population of only high- and low-fitness genotypes) actually retards population extinction (Bull et al., 2007). In the light of these comments, we witness in the current framework the process of stochastic extinction rather than an error threshold, and thus we refer to the mutation threshold as extinction threshold.

The upper panels of Fig. 1 represent the mean fitness (i.e., the fraction of reproducing or fit individuals) for several examples of simulations, showing that the stochastic extinction of the population at finite times is largely driven by an increase of the fluctuations. That is, for a fixed mutation rate μ, the average fraction of reproducing individuals is the same for various values of K, but the standard deviation increasing with K. For large K, due to these large excursions, the reproducing population may reach a threshold value at which extinction occurs. The main effect of delayed selection is thus a strong increase in fluctuations, that causes stochastic extinction in finite populations at mutation rates significantly lower than the non-delayed selection ($K=1$). This can be seen in the lower panels of Fig. 1 where the extinction threshold or survival probability (panel b) is illustrated as resulting from the simulations. A rough estimation of the survival probability was considered to be the fraction of the