Physica A 474 (2017) 32–38

Contents lists available at ScienceDirect

Physica A

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

Biological evolution model with conditional mutation rates

David B. Saakian [a,b], Makar Ghazaryan [c,d], Alexander Bratus [e,f], Chin-Kun Hu [a,g,h,*]

[a] Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan
[b] A.I. Alikhanyan National Science Laboratory (Yerevan Physics Institute) Foundation, Alikhanian Brothers St.2, Yerevan 375036, Armenia
[c] State Engineering University of Armenia, Teryan St. 105, Yerevan 375036, Armenia
[d] Yerevan State University, Alek Manougian 1, Yerevan 375036, Armenia
[e] Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow 119992, Russia
[f] Applied Mathematics department, Moscow State University of Railway Engineering, Moscow 127994, Russia
[g] National Center for Theoretical Sciences, National Tsing Hua University, Hsinchu 30013, Taiwan
[h] Business School, University of Shanghai for Science and Technology, Shanghai 200093, China

HIGHLIGHTS

- Most biological evolution models have constant mutation rates.
- Here mutation rates from lower populated sequences to higher ones are reduced.
- The calculated analytic results are confirmed by numerical solutions.
- The model increases the heterogeneity and the mean fitness of the population.
- The model can qualitatively explain some experimental results of a viral population.

ARTICLE INFO

Article history:
Received 10 March 2015
Received in revised form 29 November 2016
Available online 4 January 2017

Keywords:
Crow–Kimura model
Modulated mutation rates
Complex system

ABSTRACT

We consider an evolution model, in which the mutation rates depend on the structure of population: the mutation rates from lower populated sequences to higher populated sequences are reduced. We have applied the Hamilton–Jacobi equation method to solve the model and calculate the mean fitness. We have found that the modulated mutation rates, directed to increase the mean fitness.

1. Introduction

Ideas and methods of statistical physics have been applied to study various interesting interdisciplinary research problems, such as literary authorship disputes [1–3], financial fluctuations [4–9], and biological evolution [10–14]. In this paper, we will address an interesting problem in molecular models of biological evolution.

In recent decades, there was much progress in the study of asexual biological evolution models [10–26] with fixed fitness landscape and constant mutation rates. In such models, a genome with L genes is represented by a chain of L spins (alleles) and every spin takes the values ± 1, similar to the Ising model [27]. There are 2^L different types of sequences

* Corresponding author at: Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan.
E-mail address: huck@phys.sinica.edu.tw (C.-K. Hu).

http://dx.doi.org/10.1016/j.physa.2017.01.026
0378-4371/

$S_i \equiv (\sigma_1^{(i)}, \sigma_2^{(i)}, \ldots, \sigma_L^{(i)})$, $0 \leq i \leq 2^L - 1$ with corresponding probabilities p_i and the fitness r_i. The Hamming distance between S_i and S_j is given by

$$d_{ij} \equiv \left(L - \sum_{k=1}^{L} s_i^{(k)} s_j^{(k)} \right) /2.$$

Without the loss of generality, we can choose the sequence with the largest r_i value as a reference sequence and denote it as S_0 with all spin components being $+1$. During a short period of time dt, any allele can change the type (from $+1$ to -1 or from -1 to $+1$) with the probability μdt.

There are two famous molecular models of biological evolution. One is the Eigen model [10,14,20,22] with coupled mutation–selection scheme, in which mutation and reproduction appear in the same term in the equations for p_i; another is the Crow–Kimura model [11,17,21] with parallel mutation–selection scheme, in which mutation and reproduction appear in different terms in the equations for p_i. The Crow–Kimura model has been mapped into the quantum statistical model of the Ising model in the transverse magnetic field [17]; in such a mapping, the genome length L is corresponding to the lattice size of the lattice spin model [27]. A lattice model may have a phase transition when the lattice size approaches to infinite. To get different phases for evolutionary dynamics, we need rather large genome length [21] and population size.

The calculated quantities in biological evolution models include the mean fitness [17,19,22], the steady state distribution [23,24,28], and population dynamics [25,29]. These solutions supported the idea that there is something more than a "climbing of fitness hills" [30] (the population moves in the genome space to the genome with the maximal fitness) and there are essentially collective (emergent) phenomena in evolution, including the error threshold [10] (the phase transition from the phase where the majority of population is around the high peak to the phase with uniform distribution of population) and selection via the flatness phenomenon [13] (the group of sequences with the equal fitness, the flat peak, can attract more population than single sequence with a higher fitness). The refereed phenomena have a collective behavior, while "hill climbing" can be organized simply, without any collective interaction. The collective phenomenon is a result of the statistical physics aspects of evolution models, as has been realized by Tarazona [15].

The mentioned phenomenon was found in the evolution with the fixed fitness landscape and looks like a cooperation between replicators with different genomes. The fixed fitness landscape had been modified to take into account some more realistic situation. Bratus, et al. [31] considered explicit space and global regulation of the Eigen model to study the diffusive stability of the model. A spatial quasispecies model was studied in [32]. The experimental results reported in [33] support the idea that there are more involved collective effects in evolution, when viruses of different types (quasispecies) interact with each other during the evolution processes, getting some advantage for the whole population. During the experiments two virus populations have been isolated, the wild-type, and the second virus population with the suppressed mutation rate due to a special point mutation. As a result, the second virus population has approximately the same fitness landscape as the first one, while carries 6 times less mutations. After putting the virus population in the new environment, the first, more heterogenous population, was much more effective at infecting new cells than the second one. It has been suggested that there are some cooperative interactions between viruses with different genomes. It is a highly involved phenomenon. While the mutation process is mainly random, its strength is somehow modulated according to the current structure of population. The phenomenon observed in [34] has been identified either as a second level selection (the high fitness does not mean that such a sequence will attract the majority of population), or as a selection via evolvability, when the evolving population tends to have an evolution advantage in changing environments. It is impossible to describe such a phenomena using a simple evolution scheme with a constant mutation rate and fitness landscape. There are good experimental confirmations that the mutation rate has been well modulated for different parts of genome [35]. The cooperative phenomenon in case of cancer cells clonal evolution is even stronger than in case of viruses [36].

In the current work we construct a simple generalization of the traditional quasispecies model with the mutation rate modulated by the sequence distribution in the population. The mutation rate between the adjacent Hamming classes (groups of sequences with the same number of mutations from the reference sequence) depends on the ratio of the number of viruses on these chases, therefore we have somehow modulated asymmetry of mutation rates. The asymmetry of mutation rates is well confirmed experimentally [37]. In our model the population itself modulates the mutation rates, while in [33] it is done artificially. What is common in both cases, the existence of different mutations rates, the heterogeneity of population, brings to the evolutionary advantage. The advantage of our model is that it is still exactly solvable. The dependence of the fitness on the population distribution is well known phenomenon in evolutionary game models [38,39]. Another well known case of the changing of the mutation rate by the virus population is a mutator phenomenon, well confirmed by experiments [40]. Our model assumes the modulation of mutation rate by population distribution, and it is much more involved to solve the current model than the mutator model [29,41].

Here we consider the parallel mutation selection scheme of the Crow–Kimura model [11,17,21], the selection and mutations are two parallel processes, contrary to the Eigen model where the selection is coupled with the mutation [14].

We consider the case of symmetric fitness landscape when the fitness is a function of number of mutations from the reference sequence. For such a model with symmetric original distribution of viruses, it is possible to get a short set of $L + 1$ equations [12,16,18] for the probabilities. The probabilities of the sequences at the same Hamming distance from the reference sequence (number of mutations from the reference sequence to the given sequence) are the same, there are $N_l = \frac{L!}{l!(L-l)!}$ sequences in the lth Hamming class (the collection of all the sequences with the l mutations from the reference sequence).