Rapid Communications

Key role of recombination in evolutionary processes with migration between two habitats

David B. Saakian [1,2,*] Eugene V. Koonin,[3,†] and Kang Hao Cheong[4,‡]

[1]*Laboratory of Applied Physics, Advanced Institute of Materials Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[2]*Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam*
[3]*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894, USA*
[4]*Science and Math Cluster, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372, Singapore*

Recombination is one of the leading forces of evolutionary dynamics. Although the importance of both recombination and migration in evolution is well recognized, there is currently no exact theory of evolutionary dynamics for large genome models that incorporates recombination, mutation, selection (quasispecies model with recombination), and spatial dynamics. To address this problem, we analyze the simplest spatial evolutionary process, namely, evolution of haploid populations with mutation, selection, recombination, and unidirectional migration, in its exact analytical form. This model is based on the quasispecies theory with recombination, but with replicators migrating from one habitat to another. In standard evolutionary models involving one habitat, the evolutionary processes depend on the ratios of fitness for different sequences. In the case of migration, we consider the absolute fitness values because there is no competition for resources between the population of different habitats. In the standard model without epistasis, recombination does not affect the mean fitness of the population. When migration is introduced, the situation changes drastically such that recombination can affect the mean fitness as strongly as mutation, as has been observed by Li and Nei for a few loci model without mutations. We have solved our model in the limit of large genome size for the fitness landscapes having different peaks in the first and second habitats and obtained the total population sizes for both habitats as well as the proportion of the population around two peak sequences in the second habitat. We identify four phases in the model and present the exact solutions for three of them.

All evolving genomes recombine via different mechanisms, and in sexually reproducing organisms, multiple recombination events between homologous chromosomes occur in each meiotic division [1–6]. Although recombination clearly is one of the central processes in evolution, the evolutionary advantages it provides remain a matter of active study and debate [2–6]. There is an exact analytical solution for recombination processes with a single crossover on symmetric fitness landscapes [7–11]. However, there is no exact theory for recombination with mutation combined with migration for symmetric landscapes, although it has been suggested [12] that recombination is central to the evolutionary process in space. The evolutionary dynamics in space is complicated, but can be emulated as a simple migration between two habitats. Evolution with both recombination and migration was first investigated in a series of papers [13–23]. The mathematical biologists discovered that even in the absence of epistasis, migration can still create an effective epistasis [13]. Most of these studies describe the diploid evolution and do not take into account mutation. For the recombination version of a quasispecies model, we investigate the haploid case with mutation and selection for a large genome length. The quasispecies model with migration (without recombination)

has been investigated in Refs. [24,25] using a modified quasispecies model [26–37]. The migration model can capture some features of evolution in space while allowing an exact analytical solution. In this Rapid Communication, we formulate a model of evolution with migration and recombination, based on the Crow-Kimura model, and solve this model exactly. Our exact solution supports the hypothesis [12] on a particularly important role of recombination for the evolution in space.

Our results can be viewed as both a solution to the spatial aspects of recombination and a different direction of quasispecies analysis that involves space dependence. The space dependence of evolutionary dynamics is usually ignored in quasispecies research, but there are some situations where space dependence strongly affects evolution. The space structure is assumed to play a crucial role in the case of the hypercycle via compartmentalization of replicators that circumvents the error threshold problem [38]. In Refs. [39,40], evolution models with migration are used to model cancer development. The quasispecies model has been numerically investigated with a simplified mutation scheme [24]. The quasispecies version of Ref. [24] has also been solved exactly [25]. In Ref. [41] the Eigen model version of migration between two habitats has been solved in the case of a single-peak fitness landscape. The general case of a symmetric fitness landscape analytically is unsolvable.

The conventional quasispecies model includes selection, mutation, and dilution processes. The frequencies of all types in the population decrease at the same rate and are related to

---

[*]david.saakian@tdtu.edu.vn
[†]koonin@ncbi.nlm.nih.gov
[‡]kanghao_cheong@sutd.edu.sg

the average fitness of the population. The population size is constant. Dilution corresponds to nonlinear interaction terms in the evolutionary equation: Every type interacts with the others competing for the resources.

In the evolutionary models with migration, there is competition for the resources within each habitat, and migration from one habitat to another. In such a case, there is no reason to assume a constant population size. Hence, the notion of saturation is introduced instead of dilution [24].

Consider an evolutionary model with two types of nucleotides, where L is the genome length. In the Crow-Kimura model with recombination, we have probabilities $\hat{p}_i$ for the 2^L sequences. There are mutation processes with transition rates from the jth type to the ith type μ_{ij} (nonzero only between neighbor states i and j). The balance condition is $\mu_{ij} = 0$. There is also a recombination process when the sequences i and j meet and exchange alleles at some position in the genome.

Let us consider a symmetric fitness landscape, where fitness depends on the total number of mutations [7,8]. In such a model, the population is distributed among Hamming classes, i.e., sets of sequences with the same number of mutations. We denote P_l the total probability of all the sequences with l mutations. We set the mutation rate to be equal to 1, and the recombination rate $c/2$. Recombination with the rate $c/2$ is equivalent to horizontal gene transfer (HGT) with the rate c. In models of HGT, an allele at some position in the genome is replaced by the corresponding allele from another sequence that is randomly chosen from the entire population. Assuming the fitness of a sequence with l mutations to be r_l, we derive

$$
\begin{aligned}
\frac{dP_l}{dt} &= P_l(r_l - 1 - R) - cP_l\left[\left(1 - \frac{\bar{l}}{L}\right)\frac{l}{L} + \left(1 - \frac{l}{L}\right)\frac{\bar{l}}{L}\right] \\
&+ \frac{L - l + 1}{L}P_{l-1}\left[1 + c\frac{\bar{l}}{L}\right] \\
&+ \frac{l + 1}{L}P_{l+1}\left[1 + c\left(1 - \frac{\bar{l}}{L}\right)\right],
\end{aligned} \tag{1}
$$

where $0 \leqslant l \leqslant L$, $\bar{l} = \sum_l lP_l$ is the mean number of mutations, and $R \sum_l P_l(t)r_l$ is the mean fitness. The term $\frac{L-l+1}{L}P_{l-1}c\frac{\bar{l}}{L}$ has an explicit meaning: In the $l-1$th Hamming class, a $+1$ allele (with the probability $\frac{L-l+1}{L}$) is replaced with the -1 allele (with the probability $\frac{\bar{l}}{L}$). In the same way, we can deduce the terms $\sim P_{l+1}$. Then, the term $\sim P_l$ is derived from the balance condition.

For the case of $l = 0$, $l = N$, we drop the terms P_{-1}, P_{N+1}. A balance condition is usually required at the initial $t = 0$,

$$
\sum_i P_l = 1,
$$

and this condition holds for any moment of time. The balance condition is a reasonable mathematical constraint, and also directly describes serial transfer experiments with viruses in a chemical reactor. The model described by Eqs. (1) and (2) has an important property: Only the ratio of the Wrightian fitnesses (the difference of the Malthusian fitnesses) is important for the evolution processes. In the case of two habitats, the situation changes: As mutation, migration is a kinetic process,

and the mean fitness depends on the absolute fitnesses as well as the mutation and migration rates.

Two-habitat model with recombination. Consider now the two-habitat model, where P_l is the total frequency of the replicator of the lth Hamming class from the first habitat, Q_l is the frequency of the same class in the second habitat, and r_l and $\hat{r}_l$ are the corresponding fitness values,

$$
\begin{aligned}
\frac{dP_l}{dt} &= P_l(r_l - 1 - \hat{P} - \alpha) - cP_l\left[\left(1 - \frac{\bar{l}}{L}\right)\frac{l}{L} + \left(1 - \frac{l}{L}\right)\frac{\bar{l}}{L}\right] \\
&+ \frac{L - l + 1}{L}P_{l-1}\left[1 + c\frac{\bar{l}}{L}\right] \\
&+ \frac{l + 1}{L}P_{l+1}\left[1 + c\left(1 - \frac{\bar{l}}{L}\right)\right], \tag{2}
\end{aligned}
$$

$$
\begin{aligned}
\frac{dQ_l}{dt} &= \alpha P_l + Q_l(r_l - \mu - \hat{Q}) \\
&- cQ_l\left[\left(1 - \frac{\hat{l}}{L}\right)\frac{l}{L} + \left(1 - \frac{l}{L}\right)\frac{\hat{l}}{L}\right] \\
&+ \frac{L - l + 1}{L}Q_{l-1}\left[\mu + c\frac{\hat{l}}{L}\right] \\
&+ \frac{l + 1}{L}Q_{l+1}\left[\mu + c\left(1 - \frac{\hat{l}}{L}\right)\right], \tag{3}
\end{aligned}
$$

where α is defined as the rate of migration from the first habitat to the second habitat with 1 and μ as the mutation rates. There is no motivation for maintaining the balance of the total population, so saturation functions $-P_l\hat{P}$ and $-Q_l\hat{Q}$ have to be introduced to restrict the population growth (this is equivalent to degradation depending on the population size). The population sizes in both habitats become important characteristics of the state of the model and there is an ambiguity in the choice of the saturation function.

Following Ref. [24], we choose for $\hat{P}$ and $\hat{Q}$,

$$
\hat{P} = \sum_l P_l, \quad \hat{Q} = \sum_l Q_l, \tag{4}
$$

which implies saturation of the population growth. We can actually choose

$$
P = \phi\left(\sum_l P_l\right), \quad \hat{Q} = \phi\left(\sum_l Q_l\right), \tag{5}
$$

where $\phi(P)$ is a polynomial function with the property $\phi(0) = 0$.

Symmetric fitness landscape. Consider Eq. (3) where the fitnesses can be described by the smooth functions,

$$
r_l = f(x_l), \quad \hat{r}_l = g(x_l),
$$

and $x_l = 1 - 2l/L$.

A trivial solution is

$$
\hat{P} = 0. \tag{6}
$$

Consider the nontrivial solutions, first, for the zero recombination case. Figure 1 illustrates the dependence of the second habitat population size on the migration rate. We observe plateaus at the very low and very high migration rates but a concave function in the midrange (Fig. 1). The concave curve
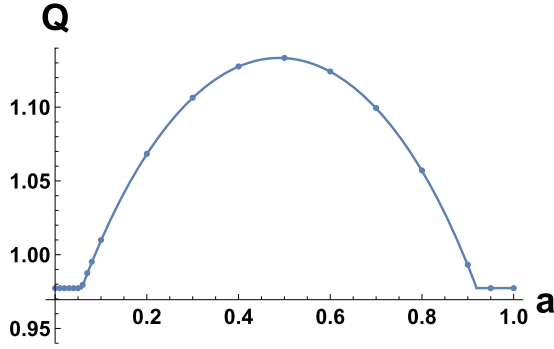
FIG. 1. The population size in the second habitat $Q = \hat{Q}$ vs the transition rate α for $N = 100$, $f(x) = 0.05x + 0.95$, $g(x) = 0.95 - 0.05x$, $\alpha = 1$, $\mu = 0.032$, $c = 0$. The smooth line is the analytical solution, and the solid dots show the numerical results. There are three intervals of solutions. In the interval $[0, x_1]$, there is a two-peak distribution in the second habitat, with the solution given by Eqs. (14)–(17) with $c = 0$; in the interval $[x_1, x_2]$, there is a single-peak distribution in the second habitat, with the solution given by Eq. (12); and in the third interval $[x_2, 1]$, there is a nonzero population only in the second habitat, with the solution given by the standard Crow-Kimura model. $x_1 = 0.0581$, $x_2 = 0.9191$.

corresponds to the one-peak distribution, whereas the plateaus reflect the two-peak distribution.

Let us now return to the nonzero recombination case. The equation for the first habitat is independent of the second habitat. Let us assume

$$P_l = \exp[Nu(1 - 2l/L)].$$

Using the formulas $P_{l\pm1} = P_l e^{-2u'}$, we derive

$$\hat{P} = f(x) + \frac{1+x}{2}\left(1 + c\frac{1-s_1}{2}\right)e^{2u'}$$
$$+ \frac{1-x}{2}e^{-2u'}\left(1 + c\frac{1-s_1}{2}\right) - \alpha - 1 - c\frac{1-xs_1}{2}. \quad (7)$$

where $u' = du/dx$.

This is, simply, the equation for the Crow-Kimura recombination model [11], the only difference being that fitness is changed by $-\alpha$, and R is replaced with $\hat{P}$. To obtain the value of $\hat{P}$, we turn to the right-hand side of the equation to derive the minimum of the expression as a function of u'. Thus

$$\hat{P} = \text{Max}\left[f(x) + \sqrt{(1-x^2)\left[\left(1 + \frac{c}{2}\right)^2 - \frac{c^2 s_1^2}{4}\right]} \right. $$
$$\left. \left. -1 - c\frac{1-xs_1}{2} - \alpha \right]\right|_x. \quad (8)$$

We introduce the surplus of the distribution,

$$s = \frac{\sum_l P_l(1 - 2l/L)}{\sum_l P_l}.$$

At the maximum point, we have $u' = 0$ and obtain

$$\hat{P} = f(s_1) - \alpha. \quad (9)$$

Importantly, s_1 does not depend on α.

The correlated phase in the second habitat. We have to choose the maximum among Eqs. (5) and (7). Thus, as before, there are three phases: extinction phase (5), nonselective phase [$s_1 = 0$ in Eq. (7)], and selective phase [$s_1 > 0$ in Eq. (7)].

We first focus on the second habitat. Summing over the index l in the second equation, we derive an exact relation,

$$(\hat{Q})^2 = (\hat{Q})\frac{\sum_l \hat{r}_l Q_l}{\hat{Q}} + \alpha P. \quad (10)$$

When deriving the solution for the second habitat, we encounter two situations: (a) a distribution around the peak sequence of the first habitat, and (b) distributions around the two peaks, where the first peak corresponds to the distribution for the first habitat and the second peak corresponds to the Crow-Kimura model for the second habitat.

Let us denote the population size in the second habitat near the first peak as q_1, and the population size near the second peak as q_2.

We can consider two cases, the first (uncorrelated) one with two peaks,

$$q_1 \sim q_2, \quad (11)$$

and the second (correlated) one with a single peak,

$$q_2 \ll q_1. \quad (12)$$

Consider now the case of Eq. (15) for strong correlation between P_l and Q_l with $q_2 = 0$. Equation (2) gives

$$[g(s) - \hat{Q}]\hat{Q} + [f(s) - \alpha]\alpha = 0. \quad (13)$$

This is the first solution for the population size in the second habitat.

The uncorrelated phase with two-peak distribution. Let us assume that the distribution has two peaks. The first one is just s_1 and the second one is some unknown s_2.

We assume that Q_l has the following representation,

$$Q_l = \exp[L\bar{u}(1 - 2l/L)].$$

Then we derive, far from s_1,

$$\hat{Q} = g(x) + \frac{1+x}{2}\left(1 + c\frac{1-s_1}{2}\right)e^{2\bar{u}'}$$
$$+ \frac{1-x}{2}e^{-2\bar{u}'}\left(1 + c\frac{1-s}{2}\right) - \alpha - 1 - c\frac{1-xs}{2}, \quad (14)$$

where s_2 is the point of maximum, and we obtain

$$\hat{Q} = g(s_2). \quad (15)$$

Near the s_1, Eq. (2) gives

$$q_1 = \frac{\alpha\hat{P}}{g(s_2) - g(s_1)}, \quad (16)$$

where $\hat{P}$ is given by Eq. (6).

Now s in Eq. (13) is a linear combination of s_1 and s_2,

$$s = \frac{q_1 s_1 + q_2 s_2}{q_1 + q_2} = s_2 + \frac{q_1(s_1 - s_2)}{g(s_2)}. \quad (17)$$
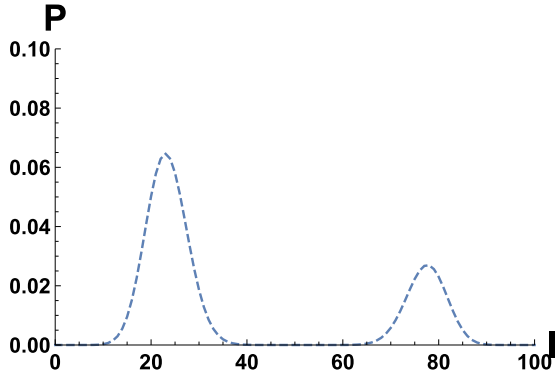
FIG. 2. The probability distribution in the second habitat P vs l for $N = 100$, $f(x) = 0.05x + 0.95$, $g(x) = 0.95 - 0.05x$, $\alpha = 1$, $\mu = 0.032$, $\alpha = 0.04$, $c = 0$, the case of the first horizontal line in Fig. 1. A two-peak distribution is observed.

Solving for the minimum of Eq. (13) via u' and the maximum via x, we obtain

$$g(s_2) = \text{Max}\left[g(x) + \sqrt{(1-x^2)\left[\left(\mu + \frac{c}{2}\right)^2 - \frac{c^2 s^2}{4}\right]}\right.$$
$$\left.\left. - \mu - c\frac{1-xs}{2} - \alpha\right]\right|_x. \qquad (18)$$

Substituting Eq. (16) into Eq. (17), we obtain a system of equations for the variables x and s_2.

Different phases for the population distribution. For the case without recombination, we obtain a concave graph with local plateaus and a global maximum point (see Fig. 1 for more information). There are three regimes for the population size in the second habitat. The plateaus correspond to the two-peak (bimodal) distribution in Fig. 2, and the single-peak distribution in Fig. 3 corresponds to the central part of the curve in Fig. 1. In the case of nonzero recombination, we found four regimes (see Fig. 4) where we observe a global maximum and two local minima, see Fig. 5 for the fourth phase.
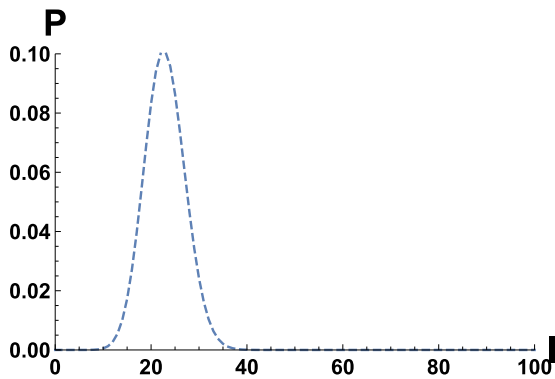

FIG. 3. The probability distribution in the second habitat P vs l for $N = 100$, $f(x) = 0.05x + 0.9$, $g(x) = 0.9 - 0.05x$, $\alpha = 0.02$, $\mu = 0.032$, $\alpha = 0.2$, $c = 0$, the central region of the graph in Fig. 1. A single-peak distribution is observed.
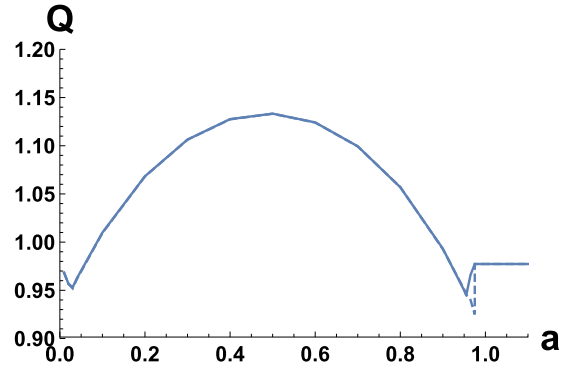

FIG. 4. The population size in the second habitat $Q = \hat{Q}$ vs migration rate $a = \alpha$ for $N = 100$, $f(x) = 0.05x + 0.9$, $g(x) = 0.9 - 0.05x$, $\alpha = 0.6$, $\mu = 0.032$, $c = 0.1$. The dashed line represents the analytical solution, and the solid line represents the numerical result. Four intervals of the solution are observed. In the interval [0,0.03], there is a two-peak distribution in the second habitat, Q is given by Eqs. (14)–(17); in the interval [0.03,0.955], there is a single-peak distribution in the second habitat given by Eq. (12); in the third interval [0.955,0.975], increase of the mean fitness with α is observed, and there is no solution; in the fourth interval $0.975 < \alpha$, there is a nonzero population only in the second habitat, with the simple solution given by the standard Crow-Kimura model, see Fig. 5.

The most important finding of this work is the decrease of mean fitness due to recombination in our model without epistasis for both habitats and with different selection coefficients. We hypothesize that this is a general phenomenon whereby migration followed by recombination results in effective epistasis.

The case of a general saturation function. The results derived thus far have been for a specific choice of the saturation function given by Eq. (3). Let us now consider the general case given by Eq. (4). We obtain an exact relation,

$$(\hat{Q})\phi(\hat{Q}) = (\hat{Q})\frac{\sum_l \hat{r}_l Q_l}{\hat{Q}} + \alpha P. \qquad (19)$$

Then, we obtain the following equation by taking the limit at large L for the correlated phase,

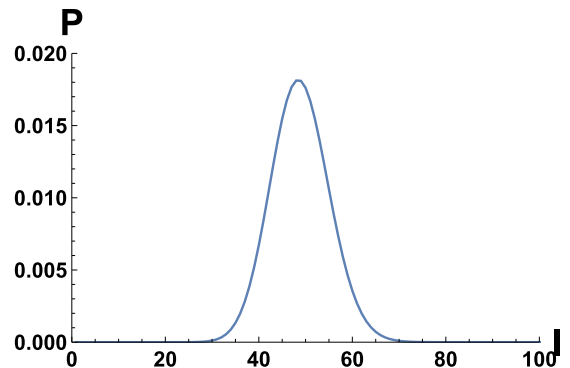$$[g(s) - \phi(\hat{Q})]\hat{Q} + [f(s) - \alpha]\alpha = 0. \qquad (20)$$

FIG. 5. The probability distribution in the second habitat P vs l for $N = 100$, $f(x) = 0.05x + 0.9$, $g(x) = 0.9 - 0.05x$, $\alpha = 0.965$, $\mu = 0.032$, $c = 0.1$. A single-peak distribution is observed.

This equation is nonlinear, and a mathematical problem arises for the proper choice of $\hat{Q}$. The solution with the maximum value of $\hat{Q}$ can be chosen. We can use our formulas to obtain solutions for different schemes of saturation given that there is no single scheme for such a phenomenon.

In conclusion, recombination is one of the leading forces of evolutionary dynamics, and it is crucial for exploring and understanding different aspects of evolution. Real-world evolutionary dynamics unravel in space, with migration between different habitats occurring at a certain rate. This includes evolutionary dynamics which involve recombination. Therefore, it is essential to investigate the space dependence of such evolutionary processes and to identify the boundaries within which space dependence can be safely ignored. We have analyzed the space dependence of evolutionary dynamics with recombination by using a simple quasispecies model with migration between two habitats. We obtained the first exact solution of such dynamics. While a series of results about recombination in space has been documented, our work closes an important lacuna by incorporating mutation, recombination, selection, and migration. Our proposed quasispecies model can describe the critical features of evolutionary dynamics while allowing for an exact analytical solution. Numerical experiments corroborate our analytical solutions, which are exact in the large genome length limit. We have also proved that the mean fitness in the steady state depends on the absolute values of Wrightian fitnesses, rather than their ratio (as is the case for evolution without space dependence). While in the zero epistasis case, recombination affects only

the dynamics of the evolutionary process without changing the mean fitness, but the situation drastically changes when asymmetric migration between two habitats is introduced. In this case, recombination can affect mean fitness as much as mutation. When solving our model in the steady state, we observe either a distribution grouped around single peak or a distribution grouped around two peaks. The order parameters of the model are the total population sizes in the habitats as well as the fractions of the populations around the peaks. There are four phases in the model, and we obtained the exact solutions for three of these. We derived the steady-state distribution and the size of populations in the two habitats in the exact form in the limit of large L. We also obtained the solution for the correlated case in which the distributions in the two habitats have the same peak position, defined by the peak in the first habitat. Solving for the third phase in our classification, which is the situation just before the extinction of the population in the first habitat, motivates future work. Our results are important as we have identified different phases of evolutionary dynamics which are universal in nature.

[1] M. Smith, *The Evolution of Sex* (Cambridge University Press, Cambridge, UK, 1978).

[2] S. P. Otto and Th. Lenormand, Nat. Rev. Genet. **3**, 252 (2002).

[3] J. Arjan, G. M. de Visser, and S. F. Elena, Nat. Rev. Genet. **8**, 139 (2007).

[4] W. J. Ewens, *Mathematical Population Genetics* (Springer, New York, 2004).

[5] R. Burger, *The Mathematical Theory of Selection, Recombination, and Mutation* (Wiley, New York, 2000).

[6] M. W. Feldman, F. B. Christiansen, and L. D. Brooks, Proc. Natl. Acad. Sci. USA **77**, 4838 (1980).

[7] E. Cohen, D. A. Kessler, and H. Levine, Phys. Rev. Lett. **94**, 098102 (2005).

[8] J.-M. Park and M. W. Deem, Phys. Rev. Lett. **98**, 058101 (2007).

[9] E. Muñoz, J. M. Park, and M. W. Deem, Phys. Rev. E **78**, 061921 (2008).

[10] D. B. Saakian, Z. Kirakosyan, and C.-K. Hu, Phys. Rev. E **77**, 061907 (2008).

[11] Zh. Avetisyan and D. B. Saakian, Phys. Rev. E **81**, 051916 (2010).

[12] N. Goldenfeld and C. Woese, Annu. Rev. Condens. Matter Phys. **2**, 375 (2011).

[13] W. H. Li and M. Nei, Theor. Popul. Biol. **6**, 173 (1974).

[14] T. Nagylaki, Theor. Popul. Biol. **76**, 197 (2009).

[15] N. Barton and M. Shpak, Genet. Res. **75**, 179 (2000).

[16] R. Burger, J. Math. Biol. **58**, 979 (2009).

[17] U. von Wangenheim, E. Baake, and M. Baake, J. Math. Biol. **60**, 727 (2010).

[18] T. Wiehe and M. Slatkin, Theor. Popul. Biol. **53**, 75 (1999).

[19] M. Spichtig and T. J. Kawecki, Am. Nat. **164**, 70 (2004).

[20] R. Burger and A. Akerman, Theor. Popul. Biol. **80**, 272 (2011).

[21] C. Bank, R. Burger, and J. Hermisson, Genetics **191**, 845 (2012).

[22] J. R. Chasnov, Theor. Popul. Biol. **81**, 119 (2012).

[23] A. Akerman and R. Burger, J. Math. Biol. **68**, 1135 (2014).

[24] B. J. Waclaw, R. J. Allen, and M. R. Evans, Phys. Rev. Lett. **105**, 268101 (2010).

[25] D. B. Saakian, Phys. Rev. E **98**, 012414 (2018).

[26] E. Baake and W. Gabriel, Annu. Rev. Comput. Phys. **7**, 203 (2000).

[27] B. Drossel, Adv. Phys. **50**, 209 (2001).

[28] M. Eigen, J. Mc Caskill, and P. Schuster, Adv. Chem. Phys. **75**, 149 (1989).

[29] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970).

[30] S. Franz and L. Peliti, J. Phys. A **30**, 4481 (1997).

[31] E. Baake, M. Baake, and H. Wagner, Phys. Rev. Lett. **78**, 559 (1997).

[32] J. Hermisson, O. Redner, H. Wagner, and E. Baake, Theor. Popul. Biol. **62**, 9 (2002).

[33] E. Baake and H. Wagner, Genet. Res. **78**, 93 (2001).

[34] J. M. Park and M. W. Deem, J. Stat. Phys. **125**, 971 (2006).

[35] D. B. Saakian, J. Stat. Phys. **128**, 781 (2007).

[36] K. Sato and K. Kaneko, Phys. Rev. E **75**, 061909 (2007).

[37] D. B. Saakian, O. Rozanova, and A. Akmetzhanov, Phys. Rev. E **78**, 041908 (2008).

[38] P. Schuster and M. Eigen, *The Hypercycle: A Principle of Natural Self-Organization* (Springer, Berlin, 1979).

[39] B. Waclaw, I. Bozic, M. E. Pittman, R. H. Hruban, B. Vogelstein, and M. A. Nowak, Nature (London) **525**, 261 (2017).

[40] K. Danesh, R. Durrett, L. J. Havrilesky, and E. Myersm, J. Theor. Biol. **314**, 10 (2012).

[41] T. Yakushkina and D. B. Saakian, Chin. J. Phys. **57**, 309 (2018).