

Scaling, genetic drift, and clonal interference in the extinction pattern of asexual population

Alexandre Rosas*

Departamento de Física, CCEN, Universidade Federal da Paraíba, Caixa Postal 5008, 58051-970, João Pessoa, Paraíba, Brazil

Isabel Gordo†

Instituto Gulbenkian de Ciências, P-2781-901 Oeiras, Portugal

Paulo R. A. Campos‡

Departamento de Física e Matemática, Universidade Federal Rural de Pernambuco, 52171-900, Dois Irmãos, Recife, Pernambuco, Brazil

(Received 1 February 2005; published 5 July 2005)

We investigate the dynamics of loss of favorable mutations in an asexual haploid population. In the current work, we consider homogeneous as well as spatially structured population models. We focus our analysis on statistical measurements of the probability distribution of the maximum population size N_{sb} achieved by those mutations that have not reached fixation. Our results show a crossover behavior which demonstrates the occurrence of two evolutionary regimes. In the first regime, which takes place for small N_{sb} , the probability distribution is described by a power law with characteristic exponent $\theta_d = 1.86 \pm 0.01$. This power law is not influenced by the rate of beneficial mutations. The second regime, which occurs for intermediate to large values of N_{sb} , has a characteristic exponent θ_c which increases as the rate of beneficial mutations grows. These results establish where genetic drift and clonal interference become the main underlying mechanism in the extinction of advantageous mutations.

DOI: [10.1103/PhysRevE.72.012901](https://doi.org/10.1103/PhysRevE.72.012901)

PACS number(s): 87.23.Kg, 89.75.Da, 02.50.Ey

The role that beneficial mutations play in long-time adaptation [1–7] and in the evolution of sexual reproduction [8–10] has been extensively investigated. It is well established that adaptation proceeds through the fixation of advantageous mutations. The contributions from theoretical investigations as well as the recent developments in the field of experimental evolutionary biology have favored a better comprehension of the dynamics fixation of those mutations. Those experiments have demonstrated the occurrence of the clonal interference phenomenon [1,3,11,12]. Clonal interference takes place when two or more favorable mutations compete for fixation. In asexual populations, beneficial mutations in distinct lineages compete with each other in order to reach fixation, which results in the ultimate loss of the remaining ones.

Most beneficial mutations are lost in the earlier stages of their appearance [13], on account of genetic drift—random fluctuations which drive new mutations to extinction even if they are beneficial. Those mutations with larger selective effects hold a larger chance of surviving these random effects. After surviving genetic drift, the beneficial mutations also have to eliminate other competitors, when they exist, in order to reach fixation. In general, the favorable mutations with larger selective effects outcompete the other ones with smaller benefit.

As opposed to the standard investigations, in this contribution we focus our study on the favorable mutations that are

lost. We wish to investigate the dynamical properties of those mutations that do not reach fixation. As we already mentioned, the extinction of a given advantageous mutation has two possible reasons: genetic drift and competition among mutations (clonal interference). We aim to distinguish the two causes by studying the distribution of the population size reached by these mutations before shrinking and subsequently going extinct.

For this purpose, we study the evolution of populations of asexual haploid organisms. We consider the Wright-Fisher model [9], in a homogeneous population and in a population with spatial structure, where each individual occupies one cell on an $L \times L$ square lattice [6]. We assume nonoverlapping generations, so that the individuals at time t generate the population at time $t+1$. Each individual competes with others to generate as many descendants as possible and hence pass its genetic information to future generations. It is precisely in the competition scheme that we distinguish the two versions of the model: while in the homogeneous form each individual competes with every individual in the population, the spatially structured model assumes local competition. Those best adapted organisms are most likely to produce offsprings. The likelihood that an individual i at time $t+1$ is the offspring of an individual j at time t is

$$P_{ij} = \frac{\pi_j}{\sum_k \pi_k}$$

where π_j is the fitness of individual j and the sum runs over all individuals for the homogeneous population and over the Moore neighborhood for the spatially structured one.

*Electronic address: arosas@fisica.ufpb.br

†Electronic address: igordo@igc.gulbenkian.pt

‡Electronic address: prac@ufrpe.br

In a recent investigation, Gordo and Campos introduced a spatially structured model for adaptive evolution in an asexual population [6]. The main motivation of that formulation is that the vast majority of species are to some extent structured into local populations where individuals compete with nearby individuals. Gordo and Campos have shown that the effectiveness of selection is reduced in the spatially structured model and this effect is more pronounced when we consider a smaller neighborhood size. In addition, they found that the substitution rate of beneficial mutations is smaller in the spatially structured model than that obtained for homogeneous populations [4]. These previous investigations use the standard formulation of studying those mutations that ultimately fix in the population, and no further information is raised about the mutations that are not successful. In the current work, we focus our study on the unsuccessful mutations in order to get some insight about the evolutionary mechanisms acting on those populations.

Each individual is represented by an infinite sequence of digits $\mathbf{S}=(s_1, s_2, \dots)$, where each bit s_α denotes the state of nucleotide α which is allowed to take two possible states $s_\alpha=0, 1$. Once we have assumed infinitely large sequences, the occurrence of back mutations is negligible. The aforementioned model corresponds to the infinite-site model [14,15]. During replication, each newborn individual inherits the beneficial (increase the fitness) and deleterious (decrease the fitness) mutations from its progenitor and an additional amount of n deleterious mutations (n taken from a Poisson distribution with, on average, U_d mutations per individual per generation). Each deleterious mutation decreases fitness by a multiplicative factor $1-s_d$. On the other hand, beneficial mutations take place at a constant rate U_b and improve fitness by a multiplicative factor $1+s_b$. Their selective effects s_b are assumed to be exponentially distributed, with average \bar{s}_b , as observed in recent experiments in bacteria [16]. It is also assumed that subsequent beneficial mutations in the same genome only occur with the fixation of the previous beneficial mutations. Hence the fitness of a given individual is given by [4,6]

$$\pi = (1 + s_b^j)(1 - s_d)^{k_d^j} \pi_{fix}, \quad (1)$$

where k_d^j is the number of deleterious mutations and π_{fix} takes into account the already fixed beneficial mutations, which are the same for every individual in the population. A given beneficial mutation is said to reach fixation when the first individual to acquire it becomes the common ancestor of the entire population, which means that every individual shares that advantageous mutation.

As mentioned above, two different effects are responsible for the disappearance of new mutations with beneficial effects on fitness. The first is the genetic drift [13,17] and the second is the clonal interference [12], competition among clones carrying different mutations. Genetic drift is expected to be more significant in small populations, since the rate of drift is proportional to $1/N$, where N is the population size. It is also expected to be more important in the earlier stages of the mutations' occurrence, since they occur at lower frequencies. On the other hand, the competition among clones which

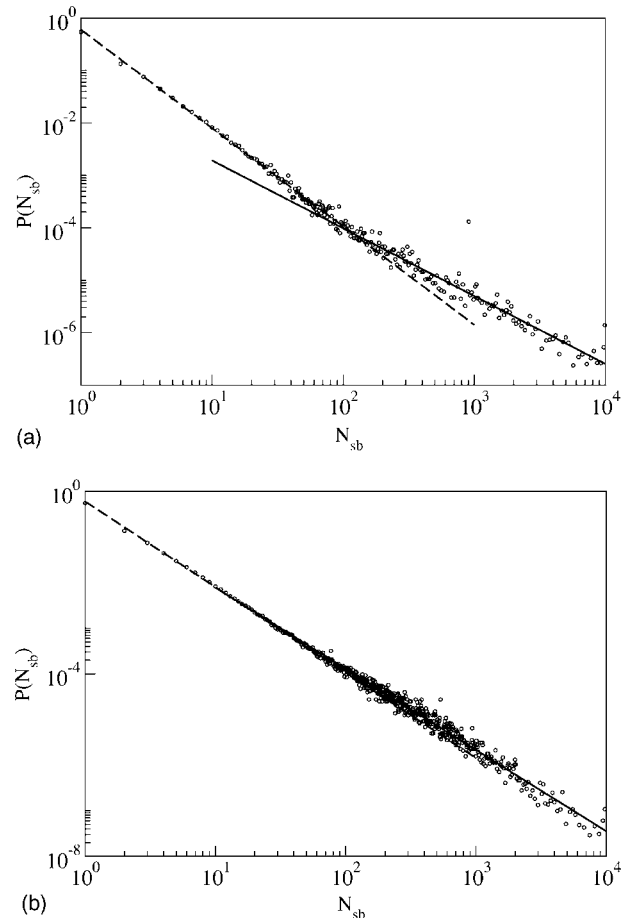


FIG. 1. Density of extinct populations $P(N_{sb})$ as a function of the maximum size of the population for the spatially structured model in a 100×100 square lattice with $\bar{s}_b=0.05$, $s_d=0.1$, and $U_d=0.1$. (a) $U_b=0.0001$ and (b) $U_b=0.002$.

carry different mutations arises when their frequencies increase as they compete for ultimate fixation. Hence, clonal interference is expected to be more significant after beneficial mutations have overcome genetic drift and have increased in size. In both cases, we expect that the probability of extinction decays with the size of the clone population that carries a beneficial mutation.

In Fig. 1 we plot the probability distribution $P(N_{sb})$ of the maximum population size N_{sb} achieved by clones that carry beneficial mutations which have not reached fixation. The data for the plots were obtained for the spatially structured population model, the homogeneous one being qualitatively similar. From the figure we clearly notice the existence of two distinct regimes: one for small values of N_{sb} , and a second regime for intermediate to large N_{sb} . The crossover behavior becomes less pronounced as we increase the rate of beneficial mutations U_b (see Fig. 1). When U_b is large, we barely distinguish the two straight lines [Fig. 1(b)]. In the former regime, genetic drift plays an essential role. In this stage, the newly arising beneficial mutations try to increase their frequency in order to escape stochastic loss. After that, they become less sensitive to random sampling in the reproduction procedure. When N_{sb} becomes large and so the clone population becomes less sensitive to random sampling; com-

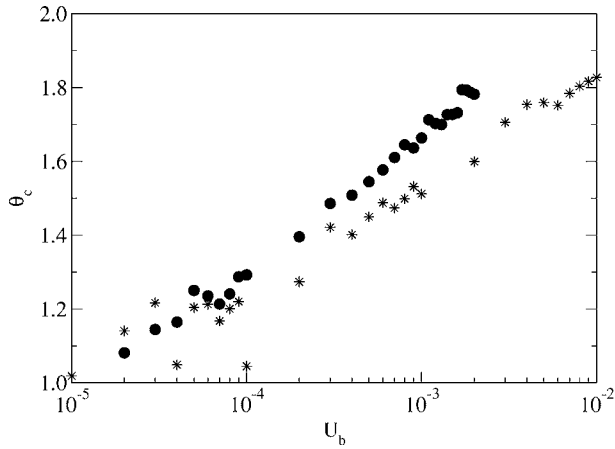


FIG. 2. Exponent θ_c , $P(s) \equiv N_{sb}^{-\theta_c}$, as a function of U_b for the spatially structured model (circles) and the homogeneous one (stars). The simulation parameters are the same as in Fig. 1. For the homogeneous model, the population is composed of 10 000 individuals.

petition among other favorable mutants becomes the dominating force that has to be overcome when beneficial mutations try to reach fixation. In general, the advantageous mutation which confers a higher benefit outcompetes the others. In the figure, we notice that each regime exhibits power law behavior with distinct characteristic exponents θ_d and θ_c , respectively. We also see that genetic drift is the main evolutionary mechanism responsible for the loss of beneficial mutations for population size up to $N_{sb} \approx 100$ [$P(N_{sb}) \sim N_{sb}^{-\theta_d}$], while clonal interference governs extinction of clone population of large size s [$P(N_{sb}) \sim N_{sb}^{-\theta_c}$]. Moreover, the critical value of changing regimes $N_{sb} \approx 100$ does not display any noticeable dependence on U_b .

In the simulations, we have found that the rate of production of beneficial mutations U_b does not affect the extinction pattern of small populations—for U_b ranging from 0.000 02 to 0.002, the scaling exponent is $\theta_d = 1.86 \pm 0.01$ for the spatial model and 1.89 ± 0.02 for the homogeneous one. This is consistent with the fact that in this regime random fluctuations play the major role in driving beneficial mutations to extinction. On the other hand, the scaling exponent for the clonal interference regime increases as U_b grows, but it is always smaller than the scaling exponent θ_d .

Figure 2 shows the dependence of the scaling exponent θ_c on the rate of beneficial mutations U_b . The scaling exponent θ_c grows slightly with the augmentation of U_b . This dependence is nearly logarithmic as we can infer from the semi-logarithmic scale in the figure. The qualitative behavior is the same for the spatially structured model as well as for homogeneous populations, although it is also worth mentioning that θ_c is systematically greater for the former. This effect has been observed previously in a model for the evolution of the protein synthesis [18,19] and is a consequence of the spatial structure that prevents the quick disappearance of clones carrying beneficial mutations by large fluctuations, since the competition among distinct mutants only takes place when the boundaries of their clusters interact, while, in homogeneous populations, as soon as a new favorable muta-

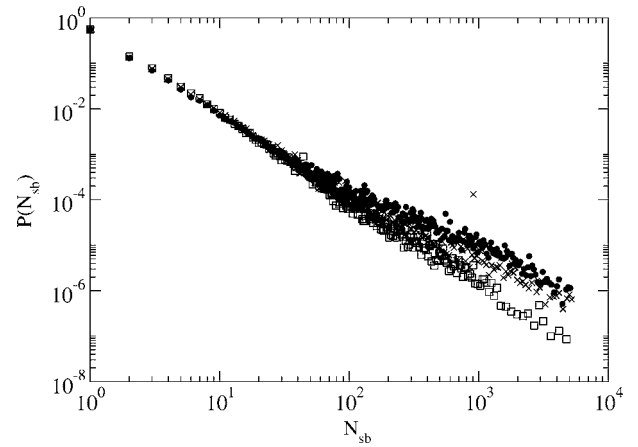


FIG. 3. Density of extinct populations $P(N_{sb})$ as a function of the maximum size of the population for the spatially structured model in a 100×100 square lattice with $\bar{s}_b = 0.05$, $s_d = 0.1$, $U_b = 0.01$, and different values of U_d : 0.01 (filled circles), 0.1 (crosses), and 0.4 (open squares).

tion arises, it immediately interacts with all other mutants because in this case competition is global.

Likewise, we investigate the dependence of the crossover on U_d . Figure 3 shows that, as U_d increases, the crossover becomes less and less pronounced. We argue that this is a consequence of the weakness of clonal interference as U_d increases, making genetic drift the sole beneficial mutation extinction mechanism.

Since the probability of extinction of beneficial mutations depends on their selection coefficient, we now turn our attention to Fig. 4, where we display the expected maximum fraction of the population $\phi(s_b)$ reached by the beneficial clone population (that does not fix) as a function of their mutational advantage s_b . The data reveal a peaked function near \bar{s}_b . We observe that mutations of small effect are really eliminated by drift in the first generations of their appearance. Mutations of large effect also hold a small value of $\phi(s_b)$. This result indicates that either these mutations are lost as soon as they arise or, when they escape stochastic loss

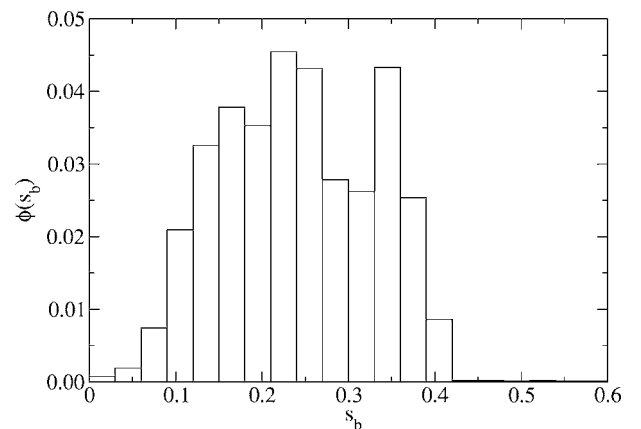


FIG. 4. Average maximum fraction of the population carrying favorable mutations that have not reached fixation $\phi(s_b)$ as a function of their selective effect s_b . $U_b = 0.01$. Other parameters as in Fig. 1.

in the earlier generations, they reach fixation very rapidly such that clonal interference is not the main underlying force that governs their fate. Thence, clonal interference acts only over a limited range of s_b ; otherwise genetic drift is the main cause of extinction of beneficial mutations.

In summary, we have quantified how genetic drift and clonal interference cause the extinction of clonal populations carrying beneficial mutations at different rates. While genetic drift affects the small size populations, clonal interference is the mechanism that drives large populations to extinction. In addition, we have observed a clear crossover in the population size where one or another of these effects takes place. Furthermore, comparing the extinction pattern on the homo-

geneous and the spatially structured models, we have noticed that large clusters in the spatial model are more stable than their counterpart in the homogeneous model. Finally, our results indicate that the clonal interference acts in a limited range of beneficial advantage while the genetic drift causes indiscriminate extinction of clone populations with any beneficial advantage.

A.R. and P.R.A.C. acknowledge financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). I.G. is supported by Project No. POCTI/BSE/46856/2002 through Fund. para a Ciência e Tecnologia (FCT), FEDER and FCT Grant No. SFRH/BPD/8104/2002.

-
- [1] S. F. Elena, *Antonie van Leeuwenhoek* **81**, 135 (2002).
 [2] R. E. Lenski, M. R. Rose, S. C. Simpson, and S. C. Tadler, *Am. Nat.* **138**, 1315 (1991).
 [3] A. M. R. Sanjuan and S. F. Elena, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8396 (2004).
 [4] P. R. A. Campos and V. M. de Oliveira, *Evolution (Lawrence, Kans.)* **58**, 932 (2004).
 [5] H. A. Orr, *Genetics* **155**, 961 (2000).
 [6] I. Gordo and P. R. A. Campos (unpublished).
 [7] V. M. de Oliveira and P. R. A. Campos, *Physica A* **337**, 546 (2004).
 [8] H. J. Muller, *Mutat Res.* **1**, 2 (1964).
 [9] R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, U.K., 1930).
 [10] J. Felsenstein, *Genetics* **78**, 737 (1974).
 [11] W. G. Hill and A. Robertson, *Genet. Res.* **8**, 269 (1966).
 [12] P. J. Gerrish and R. E. Lenski, *Genetica (The Hague, Neth.)* **102**, 127 (1998).
 [13] J. B. S. Haldane, *Proc. Cambridge Philos. Soc.* **23**, 838 (1927).
 [14] M. Kimura and J. F. Crow, *Genetics* **49**, 725 (1964).
 [15] G. A. Watterson, *Theor Popul. Biol.* **10**, 256 (1975).
 [16] M. Imhof and C. Schlotterer, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1113 (2001).
 [17] P. R. A. Campos and J. F. Fontanari, *J. Phys. A* **32**, L1 (1999).
 [18] A. Rosas, C. P. Ferreira, and J. F. Fontanari, *Phys. Rev. Lett.* **89**, 188101 (2002).
 [19] A. Rosas and J. F. Fontanari, *Origins Life Evol. Biosphere* **33**, 357 (2003).