Interfering Waves of Adaptation Promote Spatial Mixing

Erik A. Martens[1] and Oskar Hallatschek

Biophysics and Evolutionary Dynamics Group, Max Planck Institute for Dynamics and Self-Organization,
37073 Göttingen, Germany

ABSTRACT A fundamental problem of asexual adaptation is that beneficial substitutions are not efficiently accumulated in large populations: Beneficial mutations often go extinct because they compete with one another in going to fixation. It has been argued that such clonal interference may have led to the evolution of sex and recombination in well-mixed populations. Here, we study clonal interference, and mechanisms of its mitigation, in an evolutionary model of spatially structured populations with uniform selection pressure. Clonal interference is much more prevalent with spatial structure than without, due to the slow wave-like spread of beneficial mutations through space. We find that the adaptation speed of asexuals saturates when the linear habitat size exceeds a characteristic interference length, which becomes shorter with smaller migration and larger mutation rate. The limiting speed is proportional to $\mu^{1/2}$ and $\mu^{1/3}$ in linear and planar habitats, respectively, where the mutational supply μ is the product of mutation rate and local population density. This scaling and the existence of a speed limit should be amenable to experimental tests as they fall far below predicted adaptation speeds for well-mixed populations (that scale as the logarithm of population size). Finally, we show that not only recombination, but also long-range migration is a highly efficient mechanism of relaxing clonal competition in structured populations. Our conservative estimates of the interference length predict prevalent clonal interference in microbial colonies and biofilms, so clonal competition should be a strong driver of both genetic and spatial mixing in those contexts.

O NE of the most basic questions of evolutionary biology that can be studied in controlled evolution experiments is: How fast do microbial populations adapt to new environments by the accumulation of beneficial mutations? Traditionally, it was thought that the accumulation process is limited by the supply of beneficial mutations (Novick and Szilard 1950; Atwood *et al.* 1951). If a rare beneficial mutation arises and becomes sufficiently frequent, it will expand rapidly until it is present in all individuals of the population. After the completion of such a selective sweep, the population is stationary again until the next beneficial mutation arises. The accumulation rate of beneficial mutation should thus be controlled by the appearance of new beneficial mutations, which is proportional to the population size and the mutation rate (Sniegowski and Gerrish 2010). This scenario of "periodic selection" crucially rests

on the assumption that mutation rates are so small that beneficial mutations occur strictly sequentially. Evolution experiments of the last decade have shown, however, that beneficial mutation rates in microbes can be as large as 5×10^{-5} per genome per generation in the bacterium *Escherichia coli* with typical fitness effects in the range of a few percent (Desai *et al.* 2007; Perfeito *et al.* 2007; Sniegowski and Gerrish 2010). These high mutation rates have severely restricted the parameter range of periodic selection. Multiple selective sweeps, simultaneously in progress, are characteristic of a well-mixed microbial population containing $>10^4$ cells (see, *e.g.*, estimates in *Appendix D*). In the past, most evolution experiments had effective population sizes exceeding this threshold, including the Lenski experiment, in which the bacterium *E. coli* was evolved for $>50,000$ generations to adapt to minimal medium (De Visser *et al.* 1999; Miralles *et al.* 1999; Barrick *et al.* 2009). Accordingly, these experiments did not reveal a linear relation between adaptation speed and population size, but instead a much weaker dependence indicative of a mechanism of "diminishing returns" (De Visser *et al.* 1999; Shaver *et al.* 2002). The reason for this behavior is that multiple beneficial mutations

arise on different genetic backgrounds simultaneously and compete with one another for sweeping through the population. As a consequence, only a small number of arising beneficial mutations can go to fixation, and most of them are lost in the competition between clones. It was first noted by Fisher (1930) and Muller (1932) that sex relaxes clonal competition and speeds up the process of adaptation as it allows beneficial mutations to be combined in a single genome even if they first appeared in different lineages. Today, this Fisher–Muller advantage of sex is one of the most important explanation for why most organisms engage in some form of genetic exchange.

In recent years, extensive research efforts have been invested in modeling clonal interference (Rouzine *et al.* 2003; Desai and Fisher 2007; Park and Krug 2007; Hallatschek 2011) with the goal to quantitatively explain adaptation speeds and genetic diversities measured for well-mixed microbial populations (Gerrish and Lenski 1998; Desai *et al.* 2007). Although theoretical models depend on largely unknown effect distributions of beneficial mutations, they all predict that key quantities of interest, including the speed of adaptation and the genetic diversity, depend only weakly (logarithmically) on population size and mutation rates. While these predictions shape our current view of adaptation in large asexual populations, they primarily apply to well-mixed test tube populations of microbes. Many naturally occurring microbial populations attach to surfaces in the form of biofilms exhibiting a pronounced spatial structure (Tolker-Nielsen and Molin 2000; Watnick and Kolter 2000). Even in laboratories microbes are routinely grown on agar plates by which they acquire spatial structure. More complicated spatial structure is relevant to the evolution of viral populations such as influenza and severe acute respiratory syndrome (SARS), which spread in space via human transportation networks (Hufnagel *et al.* 2004; Eggo *et al.* 2011). Moreover, there is increasing evidence that spatial structure is also important to the evolutionary processes involved in cancer progression (Greaves *et al.* 2006; Merlo *et al.* 2006; Salk *et al.* 2009). One might thus wonder whether clonal interference is relevant to spatially extended populations and, if so, whether its effect on adaptation differs from that in well-mixed populations.

Although some simulation studies have considered the case of structured populations with one individual per deme (Gordo and Campos 2006; Gonçalves *et al.* 2007), a crucial fact about spatially continuous populations has been ignored so far: Beneficial mutations spread in the form of waves, first described by Fisher (1937) and Kolmogorov *et al.* (1937). Since these adaptation waves spread at a constant speed, mutant clones grow linearly with time and hence much slower than they would if they grew in well-mixed populations. The slowness of selective sweeps has the consequence that growing clones are much more likely to interfere than in well-mixed populations. The link between slow adaptive waves and the potential importance of clonal interference was in fact anticipated by Fisher (1937) in his seminal article
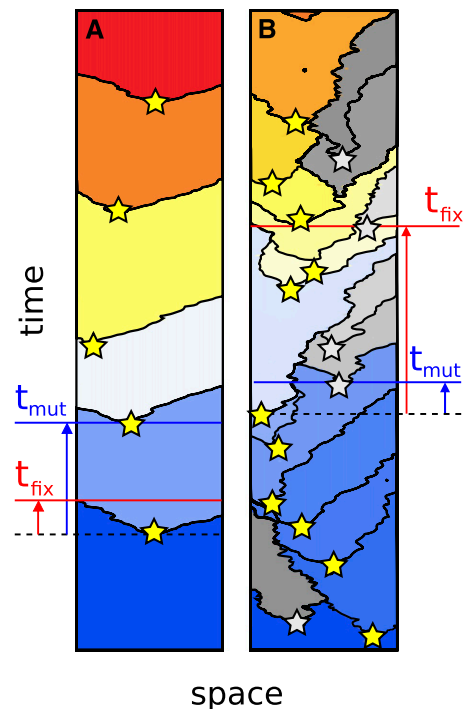

Figure 1 Simulations of asexual adaptation in a spatially extended population with absorbing boundary conditions. Stars denote the spatial location (horizontal axis) and the instant of time (vertical axis) at which new beneficial mutations become established. These events give rise to new mutant subpopulations ("clones"), which expand by waves traveling at a constant mean speed. Two drastically different adaptation scenarios are shown: (A) Periodic selection. When beneficial mutations arise sufficiently rarely, mutations sweep strictly sequentially because the waiting time for a new beneficial mutation exceeds the "fixation" time needed for a mutation to spread through the entire population. (B) Clonal interference. When the waiting time for beneficial mutations is much smaller than the fixation time, different clones compete for fixation. As a consequence, beneficial mutations are wasted (gray) unless they happen to arrive on the winning clone. Simulations were carried out with parameters $L = 100$, $s_0 = 0.25$, and $\mu = 10^{-5}$ in A and $\mu = 10^{-4}$ in B; with exponentially distributed selective coefficients; and with deterministic adaptation waves.

on "The wave of advance of advantageous genes." Therein, he considers the concrete example of a mutation with selective advantage $s = 1\%$ spreading along a continuously occupied shoreline. He estimates that spreading over 100 miles might take 10,000–100,000 generations and concludes that ". . . at any one time, the number of such waves of selective advance, simultaneously in progress, must be large" (Fisher 1937, p. 367). This suggests to consider a scenario of clonal interference characterized by many interfering adaptation waves, as illustrated in Figure 1B.

In the following, we will use simulations and analytical arguments to answer the following key questions: What is the effect of interfering Fisher waves on the speed of adaptation and the genetic diversity in an asexual population? Are there simple mechanisms of mitigating clonal interference and thus accelerating adaptation? Are these effects relevant to microbial colonies and biofilms, and perhaps measurable in evolution experiments?

Model of Interfering Adaptation Waves

To investigate clonal interference in spatially extended populations, our model is set up to allow for an efficient simulation of many adaptation waves simultaneously in progress. The habitat is modeled by a lattice with periodic boundary conditions and is always fully populated. Each lattice site records the genetic identity of the locally dominating clone in a subpopulation. This approximate description of the gene pool of the whole population, which accelerates our simulations, is appropriate if subpopulations are strongly dominated by single clones; see *Appendix A* and *Discussion* for an explicit range of validity.

We first consider a linear habitat, where the genetic state of the population is represented by a lattice of length L. Natural selection is implemented so that adaptation waves run across sites and shift the spatial extent of different clones: In each generation, the clone at site i is replaced by the neighboring clone at site j with a probability proportional to $1 + \text{sgn}(\Delta W)c(|\Delta W|)$, where c is the speed of a Fisher wave driven by a fitness difference $\Delta W = W_j - W_i$, where the sign of ΔW defines the direction of the wave.

In effect, the replacement rule generates adaptation waves traveling at an average velocity given by $c = c(\Delta W)$. The function $c(\Delta W)$ is chosen to represent the classical Fisher wave speed $c(\Delta W) = 2\sqrt{m\Delta W}$ (with migration rate $m = \frac{1}{4}$), valid for large populations with negligible genetic drift (Fisher 1937; Kolmogorov *et al.* 1937). Alternatively, we use the wave speed given by $c \sim 2\rho m\Delta W$ for strong noise (Doering *et al.* 2003; Hallatschek and Korolev 2009) with $\rho m = 1$, where ρ represents the population density (see *Appendix A* for further details).

To simulate planar habitats, we represent the population on a hexagonal lattice of size $L \times L$. Natural selection is implemented such that the clone at site i is replaced by one of the neighboring clones j with a probability proportional to its fitness, W_j (normalized by the fitness summed over all neighbors). This replacement rule generates adaptation waves traveling at an average velocity $c = 2\sqrt{m\Delta W}$ with a migration rate of $m = \frac{1}{4}$, consistent with the classical Fisher–Kolmogorov wave theory (see also *Appendix A*).

New beneficial mutations appear at a rate μ per site (μ thus represents the product of population density ρ and beneficial mutation rate per genome) and survive genetic drift with probability $2s$ (Maruyama 1974), where s is the selective fitness advantage. When a mutation becomes established at position i, the fitness W_i of the i-th clone is updated according to $W_i(t) = W_i(t-1) \cdot (1+s)$. Effects of epistasis are absent, and the habitat is homogeneous such that selective pressures are the same throughout the habitat. The selective fitness advantage s of a mutation is chosen to be either constant s_0 or drawn from a distribution with mean value s_0. We choose either an exponential distribution or a hump-shaped gamma distribution with shape parameter $k = 2$, which are both frequently used in theoretical evolu-

tion models (Eyre-Walker and Keightley 2007). See inset in Figure 3 for the shape of these distributions.

In our simulations of interfering Fisher waves, we vary the following parameters: the beneficial mutation rate μ per site, the habitat size L, the relation between speed c of traveling waves and fitness differences (linear habitat only), and the mean selective fitness advantage s_0.

Results

Speed limit for asexuals

Starting from a population devoid of genetic variation, our model generates a population of genotypes that differ in the amount and type of beneficial mutations they carry. After a transient period, in which genetic diversity builds up, the mean fitness in the population increases at a steady pace by the fixation of beneficial mutations. We quantify the speed V of adaptation by the mean fitness increase per generation.

Figure 2 shows the adaptation speed as a function of the habitat size L for various mutation rates and selection coefficients. For small habitat sizes, the adaptation speed is linear in the habitat size, which indicates that adaptation is limited by the occurrence of mutations: Doubling the habitat size, and thus the influx of beneficial mutations, doubles the adaptation speed. We thus recover the classical regime of periodic selection, in which beneficial mutations sweep strictly sequentially, as illustrated in Figure 1A. In this regime, the dynamics of individual sweeps are irrelevant, and the mean fitness increase per generation is given by $2NU_b s_0^2$, just as in the corresponding well-mixed case. However, as the habitat size exceeds a characteristic length scale L_c, we observe that the adaptation speed begins to saturate due to clonal interference. For very large systems, the adaptation speed approaches a limiting value $V_{\max}$.

These results call for an explanation of the "interference" scale L_c at which clonal interference sets in and of how the magnitude of the "speed limit" $V_{\max}$ depends on the parameters, in particular mutation rates and Fisher wave speed. It turns out that—for constant fitness effects—relatively simple estimations can be given for both quantities.

The condition for clonal interference, and hence the characteristic scale L_c, can be determined by comparing two important timescales. The first timescale estimates how long it takes for a single adaptation wave to run across the length L of the habitat. This fixation time is given by $t_{\text{fix}} \approx L/c_0$, where the traveling speed c_0 depends on the selection coefficient s_0. The second timescale is the waiting time t_{mut} for a new beneficial mutation to become established, which is inversely proportional to both the establishment probability $2s$ and the rate $\mu_b L^d$ at which beneficial mutations appear. The latter depends on the dimension d of the habitat ($d = 1$ and $d = 2$ for a linear and a planar habitat, respectively). We thus obtain the mutation waiting time $t_{\text{mut}} \approx (2s_0 L^d \mu)^{-1}$. Now, the mode of adaptation depends on the relative size of t_{fix} and t_{mut}: If the fixation time is smaller than the mutation waiting time ($t_{\text{mut}} < t_{\text{fix}}$), we expect periodic
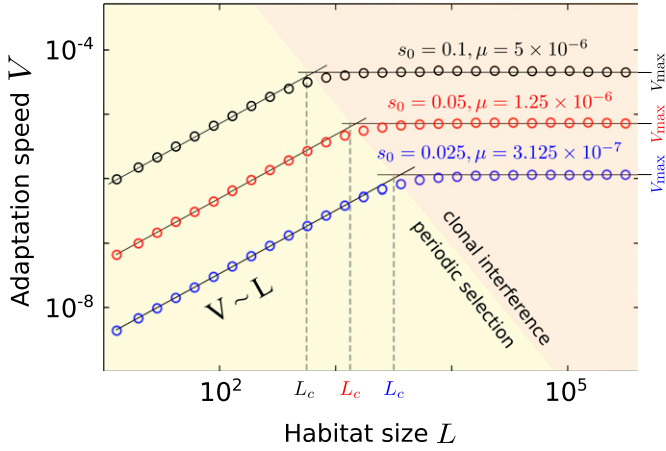
Figure 2 The speed of adaptation in a linearly extended population. For small habitat sizes, the adaptation speed V (mean fitness increase per generation) grows linearly with the habitat size L, as expected for the periodic selection regime. As the habitat exceeds a characteristic size L_c, the adaptation speed saturates at a limiting value V_{max}. Both the crossover scale and the adaptation speed limit depend on the selection advantage s_0 of beneficial mutations and the mutation supply rate μ. The simulations were run assuming a constant selection coefficient and deterministic adaptation waves. The same saturation behavior is observed when selection coefficients are distributed (exponential distribution and gamma distribution). Data represent the average over $\sim$2000 fixation events.

selection. In the opposite case, we expect clonal competition. The crossover from periodic selection to clonal interference occurs just when $t_{fix} \sim t_{mut}$, implying a characteristic "interference scale" of

$$L_c = \left(\frac{c_0}{2\mu s_0}\right)^{1/(d+1)} \propto \begin{cases} \mu^{-1/2}, & \text{linear habitat} \\ \mu^{-1/3}, & \text{planar habitat.} \end{cases} \quad (1)$$

Next, we use this estimate to find an approximate expression for the adaptation speed limit V_{max} observed in Figure 2. To this end, it is convenient to express the adaptation speed in the following general form:

$$V = 2s_0^2 L^d \mu \cdot F\left(\frac{L}{L_c}\right). \quad (2)$$

The factor $2s_0^2 L^d \mu$ on the right-hand side is the rate at which beneficial mutations accumulate in the absence of clonal interference. The function F represents the probability that a mutation reaches fixation once established. When mutations arise sequentially, each established mutation also reaches fixation; hence $F \sim 1$ for small habitats, $L \ll L_c$. On the other hand, for large habitats ($L \gg L_c$) clones interfere and F becomes very small. How small? Given that the adaptation speed saturates at large system sizes, as inferred from our simulation results in Figure 2, we must require that the habitat length L drops out of Equation 2 for large habitats. This can occur only if $F(L/L_c) \sim (L_c/L)^d$ for large L. Hence, we estimate the speed limit for large systems by

$$V_{max} \sim s_0^2 L_c^d \mu \propto \begin{cases} \mu^{1/2}, & \text{linear habitat} \\ \mu^{1/3}, & \text{planar habitat.} \end{cases} \quad (3)$$

It can be seen from Equation 2 that the scaled adaptation speed V/V_0 with $V_0 = s_0 c_0/L$ should be a unique function of the scaled mutation rate μ/μ_0 with $\mu_0 = c_0/(2s_0 L^{d+1})$. Therefore, if we plot our data sets for different parameters in one figure using axes V/V_0 and μ/μ_0, they should all collapse on a single master curve. The resulting scaling plots for linear and planar habitats are shown in Figure 3. Data corresponding to constant selection coefficients (Figure 3, black line) do indeed collapse on a single master curve, even though we varied μ, s_0, and the type of adaptation wave (weak/strong genetic drift). Note also the transition from a linear regime to the sublinear regimes with power law exponents $\frac{1}{2}$ and $\frac{1}{3}$ in one and two dimensions, respectively. These exponents are consistent with our prediction in Equation 3.

Moreover, Figure 3 displays results for simulations in which the fitness effects of new mutations were drawn from an exponential and a gamma distribution (with shape parameter $k = 2$), respectively. Note that although data for the different distributions follow our scaling predictions individually, they are slightly shifted with respect to each other. This indicates that the prefactors of the adaptation speed depend on the tails of the distribution. Broader distributions tend to yield larger adaptation speeds, apparently because they will more frequently give rise to the sampling of unusually large-s clones that outcompete average-s clones. As a consequence, fixations are more frequent and the adaptation speed is higher (see also *Appendix F*). The effect is stronger in the clonal interference than in the periodic selection regime.

Both the existence of a speed limit V_{max} for large habitats and its dependence on mutation rates contrast with the well-mixed case, where the adaptation speed depends logarithmically on both population size and mutation rates (see *Discussion*).

To further characterize the clonal interference regime, we have measured fixation times and fitness correlations in linear habitats as a function of habitat size; see *Appendixes B* and *C*. We find that $t_{fix} \sim L^{3/2}$ for clonal interference and that the fitness variance scales linearly with the size of the habitat. Both observations suggest that our model shares some universal features with certain types of crystal growth models that have been extensively studied in physics (Kardar *et al.* 1986).

Genetic and spatial mixing

We investigate two strategies of mitigating clonal interference and increasing adaptation speeds. The classical solution to clonal interference is recombination, which allows us to combine beneficial mutations arising on different genetic backgrounds. For well-mixed populations, it has been predicted that the adaptation speed increases with the
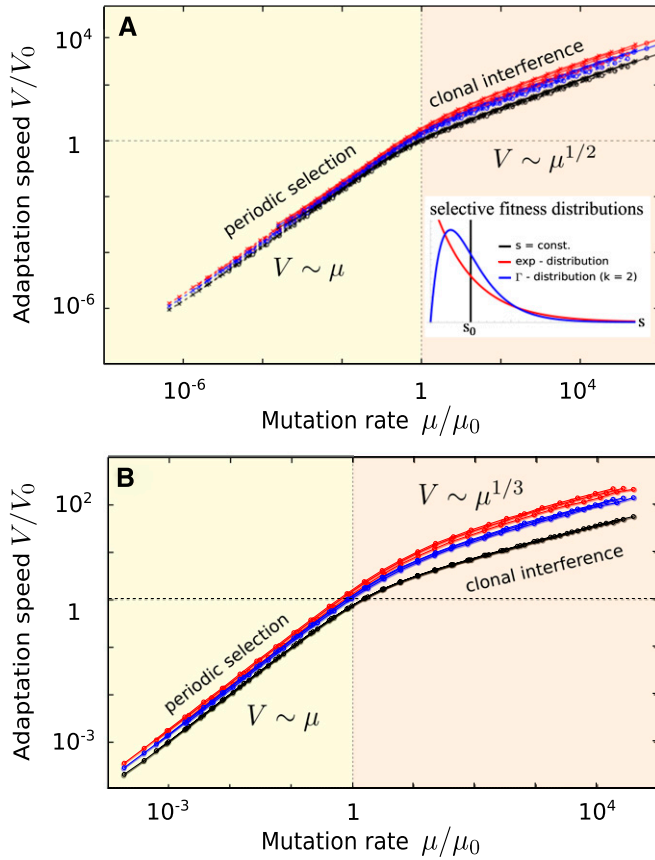
Figure 3 Scaling plots summarizing simulation data. The rescaled adaptation speed V/V_0 is shown as a function of the rescaled mutation rate μ/μ_0 (see text for the characteristic speed V_0 and mutation rate μ_0). For the periodic selection regime (small μ), we observe a linear relation between the adaptation and mutation rate; for clonal interference (large μ), we find the power laws $V \sim \mu^{1/2}$ in linear habitats (A) and $V \sim \mu^{1/3}$ in planar habitats (B). Simulations were carried out for mean selection coefficients $s_0 = 0.025$, 0.05, and 0.1 and mutation supply rates $\mu = 5 \times 10^{-6}$, 1.25×10^{-6}, and 3.125×10^{-7}. We also varied the type of adaptation wave (deterministic/noisy) and the distribution of fitness effects (black, constant selection coefficients; red, exponential distribution; blue, gamma distribution with shape parameter $k = 2$). Note that data for each type of fitness distributions fall on a single master curve, consistent with our scaling argument.

Figure 4 Genetic and spatial mixing speed up adaptation. (A) Effect of recombination on the adaptation speed in a linear habitat. Simulation runs are shown with various recombination rates (see inset). Note that the adaptation speed is positively correlated with recombination rate r in the regime of clonal interference, consistent with the Fisher–Muller hypothesis. Simulations were carried out with $s = $ const., $s_0 = 0.25$, $\mu = 3.125 \times 10^{-5}$, and deterministic adaptation waves. (B) Adaptation speeds *vs.* habitat size for varying rates m_l of long-range jumps. The adaptation speed in the periodic selection regime is unchanged, but even a small rate 0.001 of long-range jumps led to fivefold increased adaptation speeds in the clonal interference regime. Simulations were carried out for $s = $ const., $s_0 = 0.1$, $\mu = 5 \times 10^{-6}$, and deterministic adaptation waves.

recombination rate as $V \propto r^2$ for large populations where clones interfere (Neher *et al.* 2010). This result demonstrates that increasing recombination rates can strongly speed up evolution. The effectiveness of this strategy can be appreciated from Figure 4A, which displays the adaptation speed as a function of habitat size for various recombination rates. To generate these data, we have incorporated recombination in our simulations as follows. Recombinants are produced at a rate r per site from two (haploid) individuals randomly selected from the nearest neighborhood in the parent generation (selfing is allowed). Recombinant genotypes are formed from the parental genotypes by one-point recombination: The parent chromosomes are paired and a point is randomly chosen by which both chromosomes are s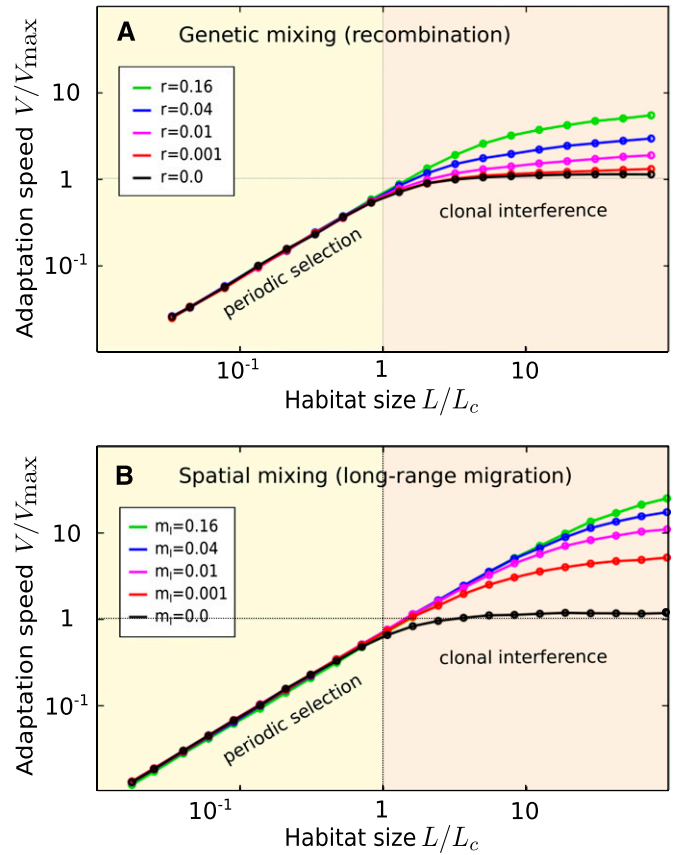plit. The mutations to both sides of the recombination point are then exchanged between the two chromosomes, forming the recombinant genotype (see *Appendix A* for further details).

Next, we study spatial mixing as an alternative way of mitigating clonal interference. Spatial mixing is implemented so that any given clone competes at a small rate m_l with a randomly chosen clone rather than with its immediate neighbor (see *Appendix A* for details of the implementation). The effect of these long-range jumps on the adaptation speed is shown in Figure 4B. Long-range migration has no effect on the periodic selection regime, because it does not change the limiting supply of beneficial mutations. In the clonal interference regime, however, adaptation speeds are continually increasing as a function of jump rate. The effect is large: Allowing for long-range jumps only every 1000 generations approximately yields a fivefold increased adaptation speed in the clonal interference regime.

Long-range migration seems to be an efficient way to mitigate clonal interference, which has significant consequences for evolution (see *Discussion*). The basic mechanism can be understood as follows: Long-range migration allows clones to replicate themselves at different locations, which enables them to effectively grow at a faster rate. In other words, a clone that has led to a second "seed" by means of long-range migration will grow twice as fast as a clone with only one seed. A third seed triples the growth rate, etc. In effect, when clonal interference is absent, one obtains exponential growth with the rate

$$s_e = \sqrt{2m_l c_0},\qquad(4)$$

which depends on the selection coefficient through the wave velocity relation $c_0 = c(s_0)$. A derivation for this rate is given in *Appendix E*. Note that the resulting exponential growth of mutants mirrors selective sweeps in well-mixed populations, however, with a quite different *effective* selection coefficient s_e (see also *Discussion*). Importantly, the resulting fast Malthusian growth accelerates the fixation process of beneficial mutations. As a consequence, mutations that would interfere without long-range migration are now more likely to arise sequentially. Consistent with this view, Figure 4 shows that the linear regime of periodic selection is extended to a larger parameter range as the long-range migration rate is increased. Of course, for sufficiently large habitat sizes, clonal interference also occurs in the presence of long-range migration. However, the number of simultaneously competing clones is reduced due to the faster sweeps. This leads to generally larger adaptation rates than for clonal interference without long-range migration, as seen in Figure 4.

Discussion

A crucial feature of adaptation in spatially extended populations is that advantageous mutations spread through the population by means of traveling waves (Fisher 1937; Kolmogorov *et al.* 1937). This wave-like motion implies that clones grow at a constant speed, in strong contrast to the accelerating logistic growth characteristic of well-mixed populations. We have established a model of adaptation by waves to investigate the rate at which beneficial mutations are accumulated in spatially extended populations. Our model represents the gene pool of local subpopulations by the locally dominating genotype. This representation is appropriate when clonal interference is absent within these subpopulations. This sets an upper bound to the beneficial mutation rate $U_b \ll \sqrt{s_0/m}/K$ (see *Appendix A*). When the beneficial mutation rate is still larger, one obtains a hybrid scenario of clonal interference on a local well-mixed scale and clonal interference by adaptation waves (E. A. Martens and O. Hallatschek, unpublished results). With the assumptions of a homogeneous habitat (constant selection pressure), short-range migration, and negligible epistasis, our model is a first step toward understanding clonal interfer-

ence in spatially structured populations, such as microbial colonies or biofilms. We expect that similar phenomena to those reported here may be observed in other rapidly evolving populations. Many natural populations are, however, characterized by more complex migration patterns than the stepping-stone model used in our work, for instance, influenza or SARS (Hufnagel *et al.* 2004; Kaluza *et al.* 2010; Eggo *et al.* 2011). For concrete predictions, the existing model needs to be extended to include these migration networks. There is also increasing evidence that spatial structure is relevant in evolutionary processes related to cancer progression (Salk *et al.* 2009), suggesting that similar modeling approaches may apply (Martens *et al.* 2011). In the following, we estimate under which conditions clonal interference might be observable in microbial colonies and discuss our key predictions and how they may be verified in microbial evolution experiments.

Clonal interference generally occurs when fixation times are larger than the waiting time for new beneficial mutations to arise in the population. Since adaptation waves travel at a constant speed and thus advance much more slowly than Malthusian sweeps, fixation times are much larger in spatially extended than in well-mixed populations. As a consequence, the regime of clonal interference is inflated in structured populations compared to well-mixed ones. The condition for clonal interference is made explicit by comparing the typical fixation time of a single beneficial mutation with the waiting time for it to arise. We found that clonal interference occurs when the habitat size is larger than a characteristic size L_c. The interference scale L_c not only defines a crossover, but also measures the typical distance an adaptation wave travels freely before it collides with another wave that arose independently of the first one. The tug-of-war between adaptation waves and the role of the interference scale L_c can be discerned from the simulation results in Figure 5.

Our simple quantitative expression in Equation 1 for the interference scale was found to increase with the speed of adaptation waves and decrease with the rate of beneficial mutations per site. A similar characteristic length appears in the context of soft sweeps, where μ has to be interpreted as the establishment rate for a particular beneficial mutation (Ralph and Coop 2010). For $L > L_c$, multiple clones carrying the same mutation collide and subdivide the habitat into patches of size L_c (see Figure 5A). This situation is similar to our case of $s = $ const. when μ is interpreted as the occurrence rate of a particular adaptation; in our model, however, waves belonging to different adaptations keep interfering with each other until one clone has reached fixation in the habitat.

To see whether clonal interference is relevant to microbial colonies, we estimate the order of magnitude of L_c (for details see *Appendix D*). We assume fitness effects on the order of 1% and mutation rates in the range of 10^{-6}–10^{-4} per genome and generation. Given the large variance of the thickness of biofilms and microbial colonies, we consider
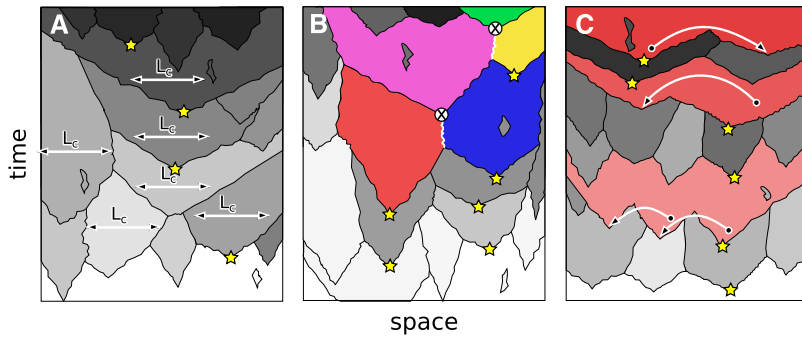
Figure 5 How can clonal interference and its deleterious effects on adaptation be overcome in spatially extended populations? (A) Standard clonal interference. The scale L_c is the typical distance at which adaptation waves collide in the clonal interference regime. (B) Recombination speeds up adaptation by combining beneficial variants (Fisher–Muller hypothesis). Two clones (red and blue) share interfaces (highlighted white) until recombination occurs ($\otimes$), and a more fit recombinant arises (violet); another recombination occurs later (yellow and violet result in the green mutant). (C) An alternative mechanism of mitigating clonal interference is provided by spatial mixing, *e.g.*, through long-range migration. Long-range jumps (arrows) accelerate selective sweeps and therefore increase both the interference scale and the fixation probability of beneficial mutations. If clonal interference is strong, spatial mixing is thus predicted to be under strong selection. Stars denote only mutations that eventually reach fixation.

a range of cell densities between 1 and 1000 per square cell diameter. This yields a mutation supply rate μ that lies within the range of 10^{-6} and 10^{-1} per square cell diameter. For $s = 1\%$, we estimate a deterministic Fisher wave speed c_0 in the range of 0.1–10 cell diameters per generation by assuming that microbes in dense colonies disperse between 1 and 100 cell diameters per generation. Thus, we estimate $L_c = (c_0/(2s_0\mu))^{1/3}$ to lie within 4 and 800 cell diameters. The characteristic interference length scale $L_c = (c_0/(2s_0\mu))^{1/2}$ for a linear habitat could be relevant for expanding microcolonies. For instance, it has been shown that the expansion of a colony is driven by a thin layer of pioneer cells (Hallatschek and Nelson 2010; Nadell *et al.* 2010). This pioneer population at the expanding front evolves in an effectively linear habitat. Applying similar estimates to the above to the context of an expanding microbial colony, we estimate an interference scale between 50 and 500 cell diameters. Importantly, the length scales both for planar and for linear habitats are much smaller than typical sizes of microbial colonies or biofilms. We thus expect clonal interference to be a widespread phenomenon—rather than the exception—in growing dense cellular clusters that conserve their spatial structure over time.

Our most important predictions for the regime of clonal interference may be summarized as follows: When the habitat size exceeds a characteristic length L_c, multiple adaptation waves are simultaneously in progress at any one time because beneficial mutations frequently arise on different genetic backgrounds. Which of the competing clones prevails and reaches fixation depends on a tug-of-war between interfering Fisher waves. Remarkably, we find that in this "frustrated" situation, increasing the habitat size does not only lead to an increased loss of beneficial mutation: The adaptation speed actually even saturates to a limiting value V_{max}, which is seen for both constant and distributed selective coefficients (see *Appendix F*). This speed limit of adaptation and its dependence on parameters were not revealed in a previous study (Gordo and Campos 2006), seemingly because too small habitats were simulated. A simple approximate expression for the speed limit is provided by Equation 3. In particular, we found that V_{max} is proportional to $\mu^{1/2}$

and $\mu^{1/3}$ in linear and planar habitats, respectively. This scaling was shown to hold for various distributions of fitness effects, including constant and exponentially distributed fitness effects, for very noisy and for deterministic Fisher waves. We may thus conclude that the predicted power laws are very robust toward details of the underlying model.

Both the speed limit and the robust power law scaling with the mutation rate are in stark contrast to well-mixed standard models, which predict that adaptation speeds never saturate and depend logarithmically on both population size and mutation rates. An experimental comparison of adaptation in well-mixed and spatially structured habitats should therefore yield qualitatively and quantitatively different results. Experiments have already found qualitative differences in the adaptation dynamics of well-mixed and structured *E. coli* populations (Habets *et al.* 2006, 2007; Perfeito *et al.* 2008). Our model suggests that the speed limit of adaptation, V_{max}, may be detected and thus quantified by varying the size of the microbial colony. The power law dependence of the limiting adaptation speed on both mutation rate and population density could be tested by comparing wild type with a mutator strain (Shaver *et al.* 2002) and by manipulating the thickness of the colony, respectively.

The prevalence of clonal interference and the associated speed limit raises the question of how clonal interference may be overcome in spatially extended populations. Certainly, genetic exchange [in bacteria through lateral gene transfer (Cooper 2007)] provides a mechanism that lessens clonal interference as it allows one to bring together beneficial mutations that arose on different genetic backgrounds; see Figure 5B. This hypothesis, originally formulated by Fisher (1930) and Muller (1932), is corroborated by our simulation results in Figure 4A. However, our simulations also show that spatial mixing is a very efficient alternative mechanism of mitigating clonal interference in structured populations. Spatial mixing increases the growth rate of fitter clones and thus lowers fixation times. Lowering fixation times increases the likelihood for beneficial mutations to occur sequentially on the same background, and therefore clonal interference is attenuated. In our simulations, spatial mixing was introduced by allowing for a small rate of long-

range jumps. These jumps enable clones to replicate themselves at distant locations in the habitat, which was shown to lead to accelerated clonal growth, as illustrated in Figure 5C. When clones do not interfere, the growth is exponential. The rate of this exponential growth, or *effective* selection coefficient s_e, was found to depend on both the actual selection coefficient and the rate of long-range jumps (*cf*. Equation 4). Importantly, we found that, due to these fast Malthusian sweeps, even very small rates of long-range jumps strongly increased adaptation speeds in the clonal interference regime (*cf*. Figure 4B). Thus, long-range migration seems highly beneficial in spatially extended populations with homogeneous selection pressures. Nevertheless, it remains to be seen in future work whether alleles conferring the ability of long-range migration are actually selected, even if large jumps impose a danger (for instance, a long-range jump could be fatal to a cell if the cell is swept away to a harmful environment). This evolutionary mechanism could be relevant for biofilm-forming bacteria and select for occasional switching from the biofilm state to the more efficiently dispersing planctonic phenotype.

In summary, we have analyzed a model of interfering waves. Our model suggests that clonal interference is widespread in biofilms and leads to markedly different adaptation dynamics than in well-mixed populations. In particular, the speed limit V_{max} and its power law dependence on mutation rates should be quantifiable in evolving microbial colonies. Finally, we found spatial mixing to be highly beneficial in structured populations as it relaxes clonal interference and speeds up adaptation.

Acknowledgments

Literature Cited

Atwood, K. C., L. K. Schneider, and F. J. Ryan, 1951   Periodic selection in Escherichia coli. Proc. Natl. Acad. Sci. USA 37(3): 146.

Barrick, J. E., D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh *et al.*, 2009   Genome evolution and adaptation in a long-term experiment with E. coli. Nature 461(7268): 1243–1247.

Cooper, T. F., 2007   Recombination speeds adaptation by reducing competition between beneficial mutations in populations of Escherichia coli. PLoS Biol. 5(9): e225.

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

Desai, M. M., and D. S. Fisher, 2007   Beneficial mutation selection balance and the effect of linkage on positive selection. Genetics 176: 1759–1798.

Desai, M. M., D. S. Fisher, and A. W. Murray, 2007   The speed of evolution and maintenance of variation in asexual populations. Curr. Biol. 17(5): 385–394.

de Visser, J. A. G. M., C. W. Zeyl, J. L. Blanchard, P. J. Gerrish, and R. Lenski, 1999   Diminishing returns from mutation supply rate in asexual populations. Science 283(5400): 404–406.

Doering, C., C. Mueller, and P. Smereka, 2003   Interacting particles, the stochastic FKPP equation, and duality. Physica A 325(1–2): 243–259.

Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow, 1998   Rates of spontaneous mutation. Genetics 148: 1667–1686.

Drossel, B., 2001   Biological evolution and statistical physics. Adv. Phys. 50(2): 209–295.

Eggo, R. M., S. Cauchemez, and N. M. Ferguson, 2011   Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. J. R. Soc. Interface/R. Soc. 8(55): 233–243.

Elena, S. F., V. S. Cooper, and R. E. Lenski, 1996   Punctuated evolution caused by selection of rare beneficial mutations. Science 272(5269): 1802–1804.

Eyre-Walker, A., and P. D. Keightley, 2007   The distribution of fitness effects of new mutations. Nat. Rev. Genet. 8(8): 610–618.

Fisher, R. A., 1930   *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Fisher, R. A., 1937   The wave of advance of advantageous genes. Ann. Eugen. 7: 355–369.

Gerrish, P. J., and R. E. Lenski, 1998   The fate of competing beneficial mutations in an asexual population. Genetica 102: 127–144.

Gonçalves, E. D. A., V. M. de Oliveira, A. Rosas, and P. R. Campos, 2007   Speed of adaptation in structured populations. Eur. Phys. J. B 59(1): 127–132.

Gordo, I., and P. R. Campos, 2006   Adaptive evolution in a spatially structured asexual population. Genetica 127(1–3): 217–229.

Greaves, L. C., S. L. Preston, P. J. Tadrous, R. W. Taylor, M. J. Barron *et al.*, 2006   Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission. Proc. Natl. Acad. Sci. USA 103(3): 714–719.

Habets, M. G. J. L., D. E. Rozen, R. F. Hoekstra, and J. A. G. M. de Visser, 2006   The effect of population structure on the adaptive radiation of microbial populations evolving in spatially structured environments. Ecol. Lett. 9(9): 1041–1048.

Habets, M. G. J. L., T. Czárán, R. F. Hoekstra, and J. A. G. M. de Visser, 2007   Spatial structure inhibits the rate of invasion of beneficial mutations in asexual populations. Proc. Biol. Sci. 274 (1622): 2139–2143.

Haldane, J. B. S., 1927   A mathematical theory of natural and artificial selection. Proc. Camb. Philos. Soc. 23: 838–844.

Hallatschek, O., 2011   The noisy edge of traveling waves. Proc. Natl. Acad. Sci. USA 108(5): 1783.

Hallatschek, O., and K. Korolev, 2009   Fisher waves in the strong noise limit. Phys. Rev. Lett. 103(10): 1–4.

Hallatschek, O., and D. R. Nelson, 2010   Life at the front of an expanding population. Evolution 64(1): 193–206.

Hufnagel, L., D. Brockmann, and T. Geisel, 2004   Forecast and control of epidemics in a globalized world. Proc. Natl. Acad. Sci. USA 101: 15124–15129.

Imhof, M., and C. Schlötterer, 2001   Fitness effects of advantageous mutations in evolving E. coli populations. Proc. Natl. Acad. Sci. USA 98(3): 1113.

Kardar, M., G. Parisi, and Y. Zhang, 1986   Dynamic scaling of growing interfaces. Phys. Rev. Lett. 56(9): 889–892.

Kaluza, P., A Kölzsch, M. T. Gerstner, and B. Blasius, 2010   The complex network of global cargo ship movements. J. R. Soc. Interface 7(48): 1093–1103.

Kimura, M., 1962   On the probability of fixation of mutant genes in a population. Genetics 47: 713–719.

Kolmogorov, A. N., I. G. Petrovskii, and N. S. Piskunov, 1937   A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. Bull. Moscow Univ. Math. Ser. A 1(1): 713–719.

Martens, E. A., R. Kostadinov, C. C. Maley, and O. Hallatschek, 2011   Spatial structure increases the waiting for cancer. arXiv:1106.3005. New J. Phys. 13 (in press).

Maruyama, T., 1974 A simple proof that certain quantities are independent of the geographical structure of population. Theor. Popul. Biol. 5(2): 148–154.

Merlo, L. M. F., J. W. Pepper, B. J. Reid, and C. C. Maley, 2006 Cancer as an evolutionary and ecological process. Nat. Rev. Cancer 6(12): 924–935.

Miralles, R., P. J. Gerrish, A. Moya, and S. F. Elena, 1999 Clonal interference and the evolution of RNA viruses. Science 285 (5434): 1745.

Muller, H., 1932 Some genetic aspects of sex. Am. Nat. 66(703): 118–138.

Nadell, C. D., K. R. Foster, and J. B. Xavier, 2010 Emergence of spatial structure in cell groups and the evolution of cooperation. PLoS Comput. Biol. 6(3): e1000716.

Neher, R. A., B. I. Shraiman, and D. S. Fisher, 2010 Rate of adaptation in large sexual populations. Genetics 184: 467–481.

Novick, A., and L. Szilard, 1950 Experiments with the chemostat on spontaneous mutations of bacteria. Proc. Natl. Acad. Sci. USA 36(12): 708.

Park, S. C., and J. Krug, 2007 Clonal interference in large populations. Proc. Natl. Acad. Sci. USA 104(46): 18135.

Patwa, Z., and L. M. Wahl, 2008 The fixation probability of beneficial mutations. J. R. Soc. Interface/R. Soc. 5(28): 1279–1289.

Perfeito, L., L. Fernandes, C. Mota, and I. Gordo, 2007 Adaptive mutations in bacteria: high rate and small effects. Science 317 (5839): 813–815.

Perfeito, L., M. I. Pereira, P. R. A. Campos, and I. Gordo, 2008 The effect of spatial structure on adaptation in E. coli. Biol. Lett. 4 (1): 57–59.

Prähofer, M., and H. Spohn, 2000 Universal distributions for growth processes in 1+1 dimensions and random matrices. Phys. Rev. Lett. 84(21): 4882–4885.

Ralph, P. L., and G. Coop, 2010 Parallel adaptation: One or many waves of advance of an advantageous allele? Genetics 52: 647–668.

Rouzine, I. M., J. Wakeley, and J. M. Coffin, 2003 The solitary wave of asexual evolution. Proc. Natl. Acad. Sci. USA 100(2): 587–592.

Rozen, D. E., J. A. G. M. de Visser, and P. J. Gerrish, 2002 Fitness effects of fixed beneficial mutations in microbial populations. Curr. Biol. 12(12): 1040–1045.

Salk, J. J., S. J. Salipante, R. A. Risques, D. Crispin, L. Li et al., 2009 Clonal expansions in ulcerative colitis identify patients with neoplasia. Proc. Natl. Acad. Sci. USA 106(49): 20871–20876.

Shaver, A. C., P. G. Dombrowski, J. Y. Sweeney, T. Treis, R. M. Zappala et al., 2002 Fitness evolution and the rise of mutator alleles in experimental E. coli populations. Genetics 162: 557–566.

Sniegowski, P. D., and P. J. Gerrish, 2010 Beneficial mutations and the dynamics of adaptation in asexual populations. Philos. Trans. R. Soc. B 365(1544): 1255–1263.

Tolker-Nielsen, T., and S. Molin, 2000 Spatial organization of microbial biofilm communities. Microb. Ecol. 40(2): 75–84.

Watnick, P., and R. Kolter, 2000 Biofilm, city of microbes. J. Bacteriol. 182(10): 2675.

Wloch, D. M., K. Szafraniec, R. H. Borts, and R. Korona, 2001 Direct estimate of the mutation rate and the distribution of fitness effects in the yeast Saccharomyces cerevisiae. Genetics 159: 441–452.

Appendix A: Model Description

We represent the habitat by a linear lattice of length L in one dimension and a planar hexagonal lattice of side length L in two dimensions. In both cases we choose periodic boundary conditions. Each site represents the identity of the dominating clone (represented by one haploid genome). The habitat is homogeneous and the selection pressure is the same everywhere. The habitat is fully populated at any time. Generations are nonoverlapping, such that the genome at site i descends from the genomes of the neighboring sites of the parent generation.

The algorithm stores at each lattice site the genotype of the locally dominating clone, which represents a subpopulation of K individuals in a well-mixed deme. This representation is valid when (i) clones *within* the deme do not interfere and (ii) the growth of a clone within the deme generates expanding clonal waves. These conditions are quantified as follows. If a beneficial mutation becomes established, it grows into an expanding wave after a timescale of $t_{est} = \mathcal{O}(1/s)$. During this establishment time, the clone spreads diffusively over a length scale

$$\xi = \sqrt{m\, t_{est}}. \tag{A1}$$

The establishment time t_{est} can be estimated by $t_{est} \approx \log(K\xi s_0)/s_0$, which is the time it takes for a beneficial mutation to fix in a well-mixed population of size $K\xi$. As a consequence, the length scale ξ is determined self-consistently via

$$\xi = \sqrt{\frac{m\, \log(K\xi s_0)}{s_0}}. \tag{A2}$$

Up to logarithmic corrections, we thus have $t_{est} \approx 1/s_0$ and $\xi \sim \sqrt{m/s_0}$, which corresponds to the width of a Fisher wave (Fisher 1937). For Fisher waves to occur, we have to require that the habitat size $L > \xi$ and that $\xi \gg 1$, *i.e.*, that wave fronts are much broader than a single deme size. The representation in terms of locally dominating genotypes at each lattice site is appropriate only when clones do not interfere on the length scale ξ. In other words, our model is valid when the time between the establishment of two successive beneficial mutations in a habitat region of size ξ, estimated by $(K\xi U_b 2s_0)^{-1}$, is larger than the expected establishment time t_{est}. This condition is equivalent to

$$K\xi U_b \log(K\xi s_0) \ll 1, \tag{A3}$$

which—up to logarithmic corrections—amounts to $K\sqrt{m/s_0}U_b \ll 1$. This condition defines an upper bound for the mutation rate, given by

$$U_b \ll \frac{\sqrt{s_0/m}}{K}. \tag{A4}$$

Simulations are started with a population devoid of genetic variation. As time passes, beneficial mutations occur at a constant rate μ per site and generation. The selective
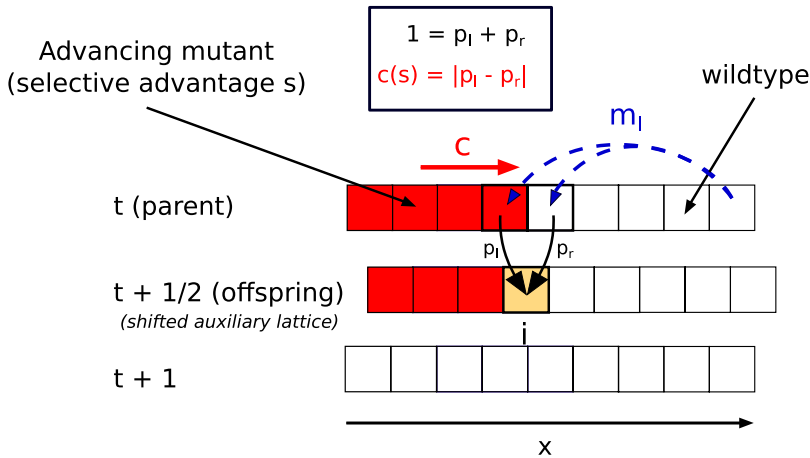
Figure A1 Simulation scheme in the linear habitat. Mutant waves advance with a prescribed velocity relation $c = c(\Delta W)$, corresponding to weak or strong selection. The fitness difference $\Delta W = W_l - W_r$ defines a velocity according to the wave velocity relation $c(\Delta W)$ (chosen for either strong or weak selection). The offspring is picked from the two immediate neighbors in the parent generation at t. An auxiliary lattice is used every half generation $(t + \frac{1}{2})$, shifted by half a lattice site (corresponding to a hexagonal lattice in space–time coordinates). The probabilities of picking either left or right neighbor are defined via $|p_l - p_r| = |c(\Delta W)|$ and $p_l + p_r = 1$. Long-range migration occurs at a rate m_l (per site and generation), whereas the distances of the jumps are drawn uniformly. Selection occurs in the parent generation between one of the neighboring parent individuals and the randomly chosen distant parent (blue arrow).

fitness advantage s of a mutation is chosen to be either constant or drawn from a distribution with mean value s_0. We choose either an exponential distribution or a hump-shaped gamma distribution with shape parameter $k = 2$, which are both frequently used in theoretical evolution models (Eyre-Walker and Keightley 2007).

Fitness is updated according to $W_i \rightarrow W_i(1 + s)$ when a mutation with effect s is drawn at site i (hence epistasis is neglected). To avoid numerical overflow we compute the logarithmic fitness. The speed of adaptation on the population is defined by the rate of change of logarithmic fitness; *i.e.*, $V = \log[\bar{W}(t + 1)/\bar{W}(t)]$, where $\bar{W} = 1/N\sum_{i=1}^{N} W_i$, where N is the number of sites.

Simulation of the Linear Habitat

For the simulation of the linear habitat, we devised an algorithm that enables us to control the speed of advancing clonal waves. This allows us to check the robustness of our scaling hypotheses toward the different wave velocity relations for strong selection, $c = 2\sqrt{m\Delta W}$ (Fisher 1937; Kolmogorov *et al.* 1937) (m represents the migration rate between sites), and for weak selection, $c \sim 2\rho m\Delta W$, respectively (Doering *et al.* 2003; Hallatschek and Korolev 2009). Natural selection is implemented such that waves run across sites with the velocity $c = c(\Delta W)$ (ΔW is the fitness difference between the clones on both sides of the wave front)

and alter the spatial extent of different clones. We use an auxiliary lattice every half generation, shifted by half a lattice site, as shown in Figure A1. As a consequence, an individual at location i in the offspring generation $(t + \frac{1}{2})$ may descend from only two individuals in the parent generation (t), *i.e.*, the neighbor to the left or to the right (see Figure A1). The fitness difference between the left and right sites is then defined as $\Delta W = W_l - W_r$. The lattice formed by full and half generations forms a hexagonal lattice in space–time coordinates.

In Figure A1, a more advantageous clone is shown on the left (red) and a less fit wild type on the right (white). The fitter clone is more likely to reproduce and therefore, on average, tends to move to the right. The genome to be passed on to site i in the offspring generation $t + \frac{1}{2}$ descends from either the left or the right immediate neighbor in the parent generation t, depending on its fitness difference ΔW. To determine from which site the genome is passed on, we assign two probabilities p_l and p_r to the left and right parents. The probabilities must sum to unity; *i.e.*, $p_l + p_r = 1$. To obtain a resulting motion of the advantageous clone according to a given velocity relation $c(\Delta W)$, we let the difference of the two probabilities be equal to the wave velocity; *i.e.*, $|c(\Delta W)| = |p_l - p_r|$. As a consequence, the advantageous clone moves with an average velocity of $c(\Delta W)$.

To confirm that our algorithm reproduces the correct wave velocity relations, we have measured ensemble averages of
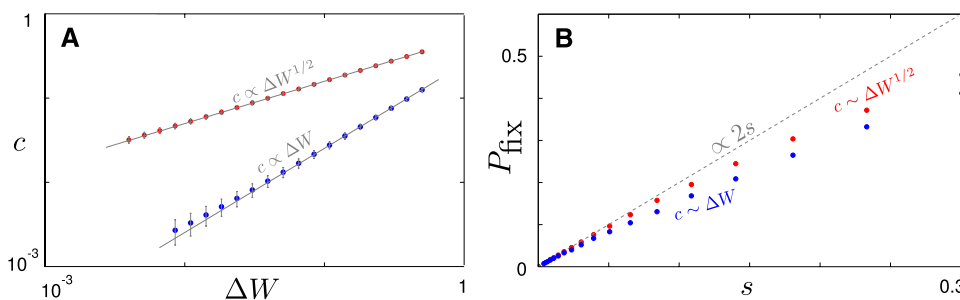

Figure A2 Wave velocity relations $c = c(\Delta W)$ and fixation probabilities of arising mutations (for the linear habitat) were measured to check the expected wave velocity in our simulations. (A) Data match the linear and square-root behavior, corresponding to the regimes of strong genetic drift and strong selection, respectively. (B) The fixation probability is given by $2s$ for small selective advantages s. Simulations were carried out for a habitat size of $L = 1000$, and measurements represent averages over 4000 fixation events.
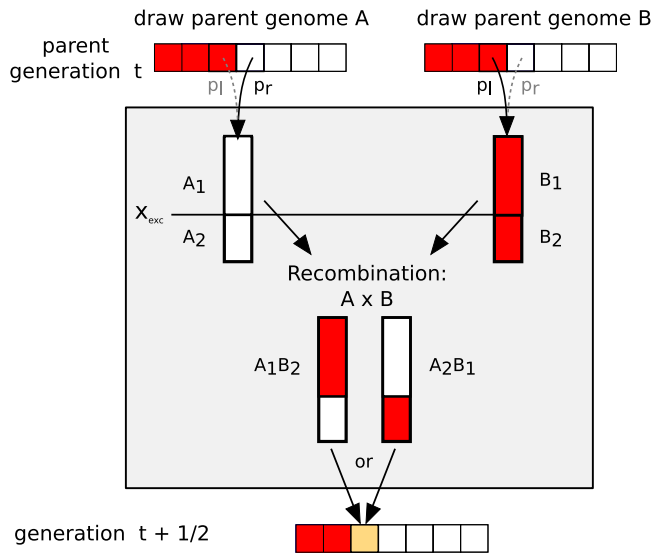
Figure A3 Recombination occurs at a rate r per site and generation. Sites occupied by the wild type are shown in white and sites occupied by advantageous mutations are shown in red. Two parent genomes are selected from the parent generation. Subsequently, their genomes are separated at a random position x_{exc} and the resulting genome parts are swapped to form the recombined genomes A_1B_2 and A_2B_1. One of them is chosen randomly and passed on to the next generation.

the wave velocity c in the absence of clonal interference; see Figure A2A. For the deterministic Fisher wave speed the migration rate is $m = 1/4$, and for the noisy Fisher wave speed the product of migration rate and population density is $\rho m = 1$.

For the case of nonstructured populations in the absence of clonal interference, it is known that the probability of fixation of a mutant with a small fitness effect s is equal to $2s$ (Haldane 1927; Kimura 1962). Maruyama showed that the fixation probability is the same for spatially structured populations (Maruyama 1974), under the assumptions of a structurally subdivided population with symmetric migration and haploid individuals that evolve according to the Moran process in each subpopulation. This result was later also confirmed in simulations of spatially structured habitats (Gordo and Campos 2006; Gonçalves *et al.* 2007) (see also Drossel 2001 and Patwa and Wahl 2008 for detailed reviews of fixation probabilities). However, when we choose to use the linear wave velocity relation for weak selection, we find that our algorithm yields a fixation probability that deviates from $2s$.

We therefore made an adjustment to our simulation to compensate for this deviation, such that mutations fixate with probability $2s$. To accomplish this, we imposed a survival probability on every new mutation. We first measured the fixation probability for a given velocity, $P_{fix}(c)$. Given that a new mutation must survive with probability $2s$, we determine a survival probability via $P_{surv}(s, c) = 2s/P_{fix}(c)$. Accordingly, in the adjusted algorithm, we allow only new mutations to survive the next generation with probability P_{surv}. The resulting fixation probability approximates the re-
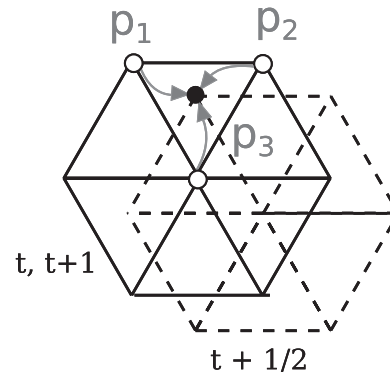

Figure A4 Planar habitats are represented by a hexagonal lattice. An auxiliary shifted lattice is used every half generation just as in the linear case. The offspring site inherits genomes from the three surrounding sites with probabilities $p_k = W_k/(W_1 + W_2 + W_3)$ with $k = 1, 2, 3$.

quired fixation probability of $2s$ well for small s, as shown in Figure A2B.

Long-Range Migration

Spatial mixing is implemented so that any given clone competes at a small rate m_l with a randomly chosen clone rather than with its immediate neighbor. Such a long-range migration event is indicated by the blue arrow in Figure A1. Accordingly, we determine the probabilities for passing on the genomes from the two competing sites by calculating the fitness difference ΔW between the randomly chosen site and a neighboring site.

Recombination

The implementation of recombination is shown in Figure A3. Recombinants are produced at a rate r per site from two (haploid) individuals randomly selected from the nearest neighborhood in the parent generation (selfing is allowed) by determining the probabilities p_l and p_r, according to the rules described above without recombination. Recombinant genotypes are formed from the parental genotypes A and B by one-point recombination: The parent chromosomes are paired and a point x_{exc} is randomly chosen by which both chromosomes are split. The mutations to both sides of the recombination point are then exchanged between the two chromosomes, forming the recombinant genotypes A_1B_2 and A_2B_1. One of these two genomes is chosen randomly and passed on to the offspring generation.

To perform recombination of two genomes, we need to keep track of all mutations and any time. Doing so is computationally expensive when many clones are sweeping simultaneously in the population. In this situation, the algorithm exceeds the complexity of $\mathcal{O}(L)$ that is expected otherwise and scales also with the number of present clones. As a consequence, simulation times explode when we try to obtain results deep in the clonal interference regime, *i.e.*, for $L \gg L_c$.
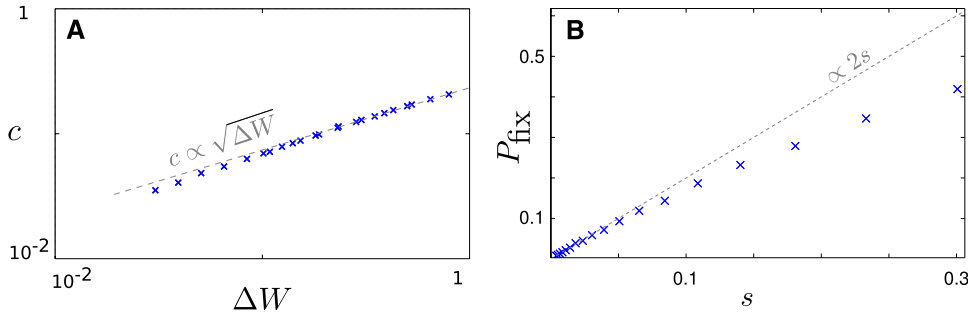
Figure A5 Measurement of the wave velocity relation and fixation probability in the simulation model of the planar habitat. (A) The wave velocity matches the behavior of the deterministic Fisher wave speed $c = 2\sqrt{\Delta W}$ with a migration rate of $m = \frac{1}{4}$. (B) The fixation probability behaves like $2s$ for small selective coefficients s. The wave velocity was measured for a habitat of $N = L \times L = 1024 \times 1024$ sites over two fixation events. The fixation probability was measured in a smaller habitat with $L = 256$.

Simulation of the Planar Habitat

For planar habitats we do not directly control the wave velocity relation, but use a simplified scheme. The habitat is represented using a hexagonal planar lattice. As in the simulation of the linear habitat, we use an auxiliary shifted lattice at every half time step $(t + \frac{1}{2})$, as shown in Figure A4. Thus we have three sites in the parent generation t surrounding the offspring site at $t + \frac{1}{2}$. Selection is performed such that the offspring site inherits the genome of one of the neighboring sites proportional to their fitness differences. This means that we need to compute three probabilities $p_k = W_k/(W_1 + W_2 + W_3)$ for inheriting the genome from the parent at site k.

We measured the velocity relation $c(\Delta W)$ resulting from this algorithm, shown in Figure A5A. The relation matches the behavior expected from a deterministic Fisher wave, $c = k\sqrt{\Delta W}$, with a prefactor $k = \frac{1}{4}$. We also measured the fixation probabilities, shown Figure A5B. The fixation probability assumes the expected behavior of $2s$ for small selective coefficients (Maruyama 1974; Gordo and Campos 2006).

Appendix B: Fixation Times in the Clonal Interference Regime

The time to fixation in the periodic selection regime is given by $t_{\text{fix}} \sim L/c$. For the clonal interference regime, we expect that in general, fixation times should be larger, but it is not

obvious how much larger they should be. For the linear habitat we measured the mean fixation time in dependence of the relative habitat size L/L_c, shown in Figure B1. Simulations have been carried out for a large range of habitat sizes L, using the deterministic wave speed $c \sim 2\sqrt{m\Delta W}$ and $s = \text{const}$. Different data sets for a range of beneficial mutation rates μ and selective coefficients s were generated (see Figure B1 legend). Rescaling the fixation time, $t_{\text{fix}}c/L_c$, we find that all data sets collapse on a single master curve and recover the following power law behavior:

$$t_{fix}\frac{c}{L_c} \sim \begin{cases} L/L_c & L \ll L_c \\ (L/L_c)^{3/2} & L \gg L_c. \end{cases} \tag{B1}$$

This result suggests the following fixation scenario: On length scales smaller than L_c, the growth of clones is *ballistic*; that is, they grow linearly in time at a constant speed. As soon as clones reach the size L_c, they start to collide with neighboring clones, which may have higher or lower fitness depending on the mutations they have accumulated in the past. As a consequence, the size of a clone follows a *random walk* on length scales larger than L_c. The variance of the size of the winning clones increases over time until it becomes of order L^2, upon which fixation occurs. Now, if the random walk of the clone size was a normal random walk (similar to, *e.g.*, Brownian motion), we would expect the clone size variance to grow linear with time. Hence, fixation should occur after a time of order L^2. By contrast, we find much
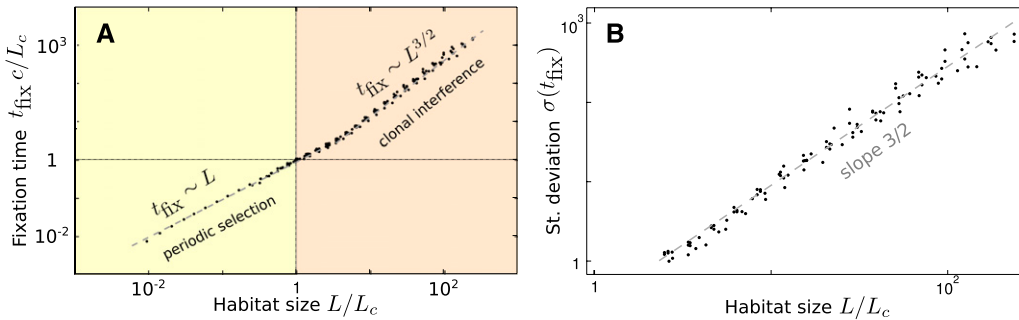

Figure B1 Measurements of fixation times in a linear habitat. (A) Simulations reveal that fixation times in the clonal interference regime scale like $t_{\text{fix}} \sim L^{3/2}$. All data collapse when using rescaled variables $t_{\text{fix}} \cdot c/L_c$ and L/L_c. (B) Standard deviation of the fixation times in dependence of the relative habitat size L/L_c. Simulations are carried out with $s_0 = 0.25$ with $\mu = 3.125 \times 10^{-3}$, 5.6×10^{-3}, and 7.8×10^{-4} and $s_0 = 0.1$ with $\mu = 1.25 \times 10^{-6}$, 7.8×10^{-4}, 5.6×10^{-3}, and 3.125×10^{-3}; clonal waves travel with the deterministic Fisher wave speed $c = 2\sqrt{m\Delta W}$; and selective coefficients are constant.

earlier fixation after a time of order $L^{3/2}$. This indicates that the random walk of the clone size is *superdiffusive* and (with the dynamical exponent of $\frac{3}{2}$) belongs to the Kardar-Parisi-Zhang universality class that describes a wide range of surface growth models (Kardar *et al.* 1986). Indeed, if fitness as a function of time is interpreted as the height of a crystal, then our asexual model with constant selective advantage s_0 and deterministic adaptation waves maps to the so-called polynuclear growth model (Prähofer and Spohn 2000).

Appendix C: Fitness Correlation Function and Interference Scale L_C

We also measured the space-dependent mean-squared fitness variance (MSFV),

$$G(x) = \left\langle \left(\log W(x + \xi) - \log W(\xi) \right)^2 \right\rangle_\xi, \qquad \text{(C1)}$$

which correlates the log fitness W at a location separated by a distance x. The angle brackets indicate that the average is taken over all positions ξ in the habitat. As shown in Figure C1, the MSFV is linear in distance for large habitats and obeys the scaling relation

$$G(x) \sim \frac{x s_0^2}{L_c}, \qquad \text{(C2)}$$

depending on the mean selective fitness advantage s_0 and the interference scale L_c. Since Equation C2 predicts $G(L_c) = s_0^2$, it is consistent with the view that clones expand over a characteristic length L_c until they collide with neighboring clones, which differ in log fitness (typically) by s_0. Furthermore, Equation C2 allows us to calculate the global variance in log fitness as

$$\text{Var}[\log W] \sim \int_0^L dx\, G(x) \sim s_0^2 \frac{L}{L_c}, \qquad \text{(C3)}$$

in large systems. The fitness variance in large systems is thus proportional to the total habitat size and is inversely proportional to the interference scale L_c.

Appendix D: Estimates of Critical Habitat Sizes for Clonal Interference

Mutation Rate and Selective Coefficient in Microbes

While many studies provide estimates for the mutation rates in microbes, fewer results are available regarding the beneficial mutation rate U (per generation and per genome) (Elena *et al.* 1996; Desai *et al.* 2007; Perfeito *et al.* 2007) and their selective coefficients. Some results that are relevant to us are compiled in Table D1 (for a review see Sniegowski and Gerrish 2010).

Some studies suggest mutation rates as low as $U \sim 10^{-9}$ or 10^{-8} per genome per generation (Gerrish and Lenski 1998; Imhof and Schlötterer 2001; Rozen *et al.* 2002). How-
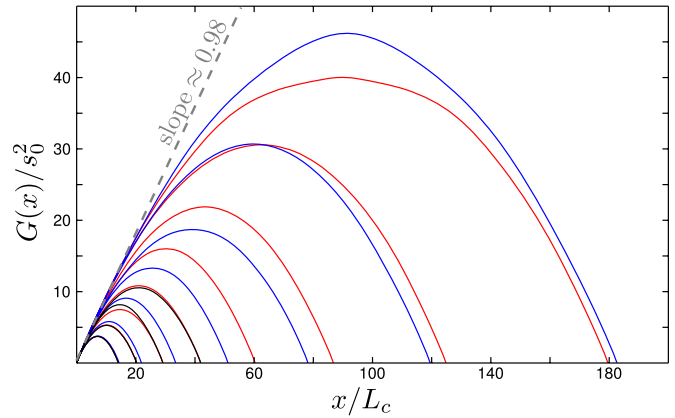

Figure C1 The distance-dependent mean squared variance $G(x)$ in log fitness for a population in a linear habitat when clonal interference is rampant. Note that the data corresponding to different habitat sizes (asymptotically) approach a straight line (dashed line) with slope of order $O(1)$ for large habitats $L \ll L_c$. Simulations were carried out for a range of habitat sizes, using the deterministic Fisher wave velocity $c = 2\sqrt{m\Delta W}$; $s = \text{const.}$; $s_0 = 0.1$; and $\mu = 10^{-4}$, 5×10^{-5}, and 10^{-5} and the corresponding values for the interference scales are $L_c \approx 89$ (red lines), 125 (blue lines), and 281 (black lines).

ever, these studies were carried out for large effective population sizes ($N_e \sim 10^7$) that had not been previously exposed to the used growth environment (and adapted accordingly). Under such circumstances, one expects that the supply rate NU is rather high; at the same time, effects of clonal interference are predicted to be very strong. In this case, not taking into account effects of genetic drift and clonal interference would lead to an underestimate of μ and an overestimate of the average effect of beneficial mutations (Sniegowski and Gerrish 2010).

More recent studies suggest that the rate of beneficial mutations is much higher. Experiments with *E. coli* (Perfeito *et al.* 2007) and *Saccharomyces cerevisiae* (Desai *et al.* 2007) find for smaller effective population sizes ($N_e \sim 10^4$)—thus avoiding clonal interference—beneficial mutation rates of $U \sim 10^{-5}$. Data from previous studies yielding the lower estimates of $U \sim 10^{-9}$ (Elena *et al.* 1996; Gerrish and Lenski 1998) have been reanalyzed (Sniegowski and Gerrish 2010) by using a likelihood model that accounts for both drift and clonal interference. One then obtains a much higher value of $U \sim 10^{-5}$, which is in line with recent estimates (Wloch *et al.* 2001; Desai *et al.* 2007; Perfeito *et al.* 2007).

Considering an estimate for the rate of all mutations (*i.e.*, including neutral and deleterious mutations) of order 10^{-3} (Drake *et al.* 1998) means that beneficial mutations make up $\sim 1\%$ of the total number of mutations. Still, the matter is not quite that simple. Even if we restrict our attention to habitats on petri dishes, we may expect that different environments may yield different U. Moreover, spatial structure may also play a role. Thus, to estimate L_c, we use a range of $U \sim 10^{-6} \ldots 10^{-4}$ for the mutation rate. For the selective coefficient we use $s_0 = 0.01$, which compares well with the estimates in Table D1.

Characteristic Population Size for Well-Mixed Populations

The crossover condition between periodic selection and clonal interference regimes is $t_{\text{fix}} \sim t_{\text{mut}}$, which allows us to determine a characteristic population size above which clonal interference is expected. For a nonstructured population, the waiting time for successful mutations is

$$t_{\text{mut}} \sim (NU\pi)^{-1}, \tag{D1}$$

where the fixation probability is $\pi = 2s_0$. In nonstructured populations the fixation time differs from the one in spatial structures and is given by

$$t_{\text{fix}} \sim \frac{\log(Ns/2)}{s}. \tag{D2}$$

(Crow and Kimura 1970). The crossover condition becomes the implicit equation $2NU\log(Ns/2) \sim 1$. Using a selective coefficient of order $s_0 \sim 0.01$ and a mutation rate in the range of $U = 10^{-6} \ldots 10^{-4}$, we find a lower and an upper bound for the characteristic population size as follows:

$$N_{\text{min}} \sim 2 \times 10^3, \tag{D3}$$

$$N_{\text{max}} \sim 8 \times 10^4. \tag{D4}$$

Upper and Lower Boundary for L_c in a Planar Habitat

To estimate the interference scale L_c for microbial populations, we are going to need estimates for the population density ρ and the wave velocity c. In the forthcoming, we measure length in units of cell length, which for the rod-shaped *E. coli* are of order 3 μm and for *S. cerevisiae* of order 5–10 μm.

Population Density

Provided that cells are packed densely, the absolute lower boundary of the population density must be $\rho = 1$ per square cell lengths. To estimate the population density in a planar population of *E. coli*, we refer to experimental data (Perfeito *et al.* 2007), where bacteria are grown on a petri dish as a lawn (diameter 6 cm = 60,000 cell lengths, *i.e.*, a radius of $R = 30,000$ cell lengths). The area of a petri dish occupies $\sim A = \pi R^2 \sim 10^9$ square cell lengths. The population in this experiment grows to a final size of $N_f = 10^{10}$. Thus the population density $\rho \equiv N_f/A$ is in this case $\rho \sim 10$ per square cell lengths. However, microbial colonies on rich medium can certainly grow to larger thicknesses. For instance, a thickness of 1000 cell diameters is not uncommon for a colony of *S. cerevisiae*. To account for a range of situations we consider a range of densities bounded by

$$\rho_{\text{min}} \sim 1, \tag{D5}$$

$$\rho_{\text{max}} \sim 1000. \tag{D6}$$

Table D1 Estimates of beneficial mutation rates and their selective coefficients, obtained from microbial experiments in recent literature

	U (per genome per generation)	s_0	Type
Gerrish and Lenski (1998), Rozen *et al.* (2002), and Imhof and Schlötterer (2001)	$10^{-9} \ldots 10^{-8}$	—	*E. coli*
Elena *et al.* (1996)	2×10^{-9}	0.03	*E. coli*
Desai *et al.* (2007)	2.4×10^{-5}	0.01	*S. cerevisiae*
Perfeito *et al.* (2007)	2×10^{-5}	0.02	*E. coli*
Sniegowski and Gerrish (2010) (data from Elena *et al.* 1996)	5.7×10^{-5}	0.003	*E. coli*

Estimating the Wave Velocity

The wave velocity is a function not only of the selective coefficient s_0 and the population density ρ, but also of the diffusion constant m. The relevant diffusion constant in two dimensions is given by

$$m = \frac{1}{4}\langle\Delta X^2\rangle, \tag{D7}$$

where ΔX is the expected dislocation distance to occur per generation. For a densely packed population of bacteria, it is reasonable to assume that ΔX is at least of the order of 1 cell length and at most of the order of 10 cell lengths (far displacement is possible in street formations of rod-shaped *E. coli*); *i.e.*, $\Delta X = 1 \ldots 100$ cells. Thus, we expect $m = 0.25 \ldots 2500$ square cell diameters per generation. To estimate the Fisher wave velocity we use the deterministic wave velocity relation $c = 2\sqrt{sm}$. We find a lower bound of $c_{\text{min}} = 0.1$ and for the upper bound $c_{\text{max}} = 10$ cells per generation. In conclusion, we find the following upper and lower bounds for the interference scale L_c in a planar habitat:

$$L_c^{\text{min}} = \left(\frac{c_{\text{min}}}{\mu_{\text{b,max}}\, 2s_{0,\text{max}}}\right)^{1/3} \sim 4 \text{ cells}, \tag{D8}$$

$$L_c^{\text{max}} = \left(\frac{c_{\text{max}}}{\mu_{\text{b,min}}\, 2s_{0,\text{min}}}\right)^{1/3} \sim 800 \text{ cells}. \tag{D9}$$

These are the estimates that we provide in the *Discussion* of the main text.

Appendix E: Long-Range Migration

As discussed in the main text, small rates of long-range jumps drastically increase the adaptation speed V. Figure E1 shows the increase of the adaptation speed V relative to the adaptation speed without long-range migration, $V|_{m_l=0}$.

Long-range jumps generate "seeds" with identical genomes far off the ancestral clone and thus multiply the

growth rate of the ancestral clone. Long-range migration has no effect on the periodic selection regime, because it does not change the limiting supply of beneficial mutations. However, generating seeds may reduce the time for the clone to fixate. As a consequence, effects of clonal interference are mitigated (and thus, the adaptation speed is increased). Moreover, the range of parameters over which periodic selection occurs, characterized by $t_{fix} \ll t_{mut}$, becomes larger with increasing rates of long-range migration.

We briefly investigate a simple model that explains how long-range migration may lead to an exponential growth of a clone. We assume that waves propagate deterministically (*i.e.*, on straight lines) and that jumps occur before the first (ancestral) clone reaches fixation. Furthermore we assume that at any time there is no more than one mutation present in the habitat (*i.e.*, absence of clonal interference). To effectively generate a seed, an individual must jump from a region already occupied by the ancestral clone to an unoccupied region in the habitat. We denote the average number of such clonal seeds, present at time t, by the variable $\bar{n}(t)$. Furthermore, $\bar{x}(t)$ is the average length in a linear habitat (or the area in a planar habitat) that is occupied by the clone. The growth of a clone follows from the following considerations: First, the rate of change of the number of seeds is given by the product of the spatial occupation of the clone, $\bar{x}(t)$, and the rate at which long-range migration events occur per site, *i.e.*, m_l. Second, the rate at which the spatial occupation increases, *i.e.*, the *effective* growth rate of the clone, $d\bar{x}/dt$, is twice the velocity of a single wave front times the number of waves present in the population. We denote the velocity of the ancestral clone by c_0. These considerations are summarized in the following ordinary differential equations:

$$\frac{d\bar{x}}{dt} = 2c_0\bar{n} \tag{E1}$$

$$\frac{d\bar{n}}{dt} = m_l\bar{x}. \tag{E2}$$

At time $t = 0$, just after a mutation has occurred in the habitat, space is unoccupied, $\bar{x}(0) = 0$, and the number of clonal waves is $\bar{n}(0) = 1$. With this initial condition, the exact solution of Equations E1 and E2 is

$$\bar{x}(t) = \sqrt{\frac{2c_0}{m_l}}\sinh(s_e t), \tag{E3}$$

where we define the effective growth rate

$$s_e = \sqrt{2c_0 m_l}. \tag{E4}$$

This is Equation 3 in the *Results* section of the main text.

The growth rate s_e is related to the time it takes for a second mutation to appear. Solving for the condition $n(t_2) = 2$ yields
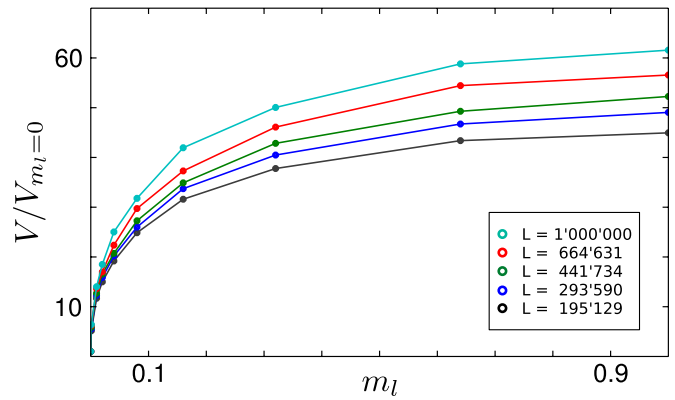

Figure E1 Relative increase of the adaptation speed V due to long-range migration. Even a small rate $m_l = 0.001$ of long-range migration drastically increases the speed of adaptation by a factor of ~5. Simulations are run with $s =$ const. and parameters $s_0 = 0.1$, $\mu = 5 \times 10^{-6}$, and the deterministic wave velocity valid for strong selection.

$$t_2 = \frac{\text{arccosh}(2)}{s_e}. \tag{E5}$$

The growing number of seeds eventually results in fixation of the clone once space is fully occupied by the clone; *i.e.*, $\bar{x}(\tilde{t}_{fix}) = L$. This condition yields a characteristic timescale for the fixation process,

$$\tilde{t}_{fix} = \frac{1}{s_e} \text{arcsinh}\left[\sqrt{\frac{m_l}{2c_0}}L\right]. \tag{E6}$$

Equation E3 describes exponential growth at rate s_e when $t \gg s_e^{-1}$. The solution is a reasonable approximation until $\bar{x}(\tilde{t}_{fix}) \sim L$. For larger times, the assumption that new seeds are born into unoccupied regions cannot be maintained. When the seeds start merging, the effective wave velocity $d\bar{x}/dt$ also slows down because the number of traveling wave fronts decreases until the growth of the clone comes to a complete halt.

Appendix F: Selective Fitness Distributions and Their Effects

We have discussed above that the speed of adaptation saturates for large habitats when clones interfere. Figure F1 demonstrates the same saturation phenomenon when the selective fitness is exponentially distributed. We expect that the saturation behavior documented for constant and exponentially distributed selection coefficients occurs for all subexponential fitness effect distributions. At this point, it is unclear how superexponential distributions might affect the adaptation dynamics.

Compared to the case of the constant selection coefficient, a nonzero variance of the distribution of fitness effects $g(s)$ increases the speed of adaptation. For periodic selection, this can be easily understood as follows. The expectation value of the adaptation speed is given by $\langle V \rangle = \langle NU_b 2s^2 \rangle$ or, more precisely, by
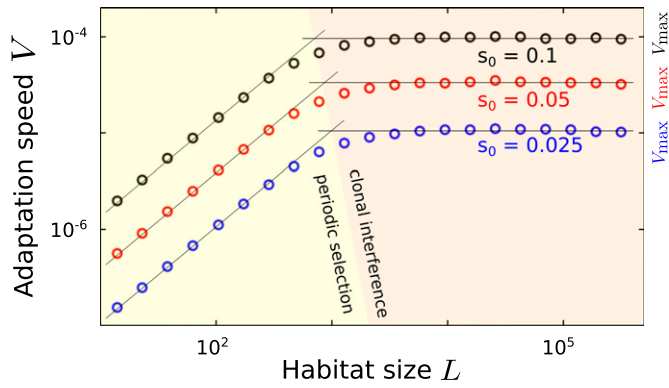
Figure F1 Saturation of the adaptation speed for large habitats for exponentially distributed selective fitness effects. The adaptation speed saturates for large habitat sizes when fitness effects are distributed, just as in Figure 2 where s = const. Simulations were carried out with $\mu_b = 5 \times 10^{-6}$ and the deterministic wave speed for $\sim$2000 fixations.

$$\langle V \rangle = NU_b \langle 2s^2 \rangle_g \equiv 2NU_b \int_0^\infty s^2 g(s) ds. \qquad (F1)$$

The mean value of the distribution is $\mu \equiv \int_0^\infty s\, g(s) ds = s_0$. Because of the identity $\mathrm{var}(s) \equiv \int_0^\infty (s-\mu)^2 g(s) ds = \int_0^\infty s^2 g(s) ds - \mu^2$, we have that

$$\langle V \rangle \equiv 2NU_b \int_0^\infty s^2 g(s) ds = 2NU_b(\mu^2 + \mathrm{var}(s)). \qquad (F2)$$

This implies that a nonzero variance increases the adaptation speed even for the periodic selection regime. When s = const., the variance is zero, but we have $\mathrm{var}(s) = s_0^2$ for the exponential and $\mathrm{var}(s) = s_0^2/k$ for the gamma distribution (in our simulations, we chose a shape parameter of $k = 2$). The shifted values of the adaptation speed $\langle V \rangle$ seen in Figure 3 can thus be explained by the differing variance values of the various selective fitness distributions.