Bacterial motility governs the evolution of antibiotic resistance in spatially heterogeneous environments

Vit Piskovsky[1,2] and Nuno M. Oliveira[1,3,*]

[1] *Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, University of Cambridge, Cambridge CB3 0WA, United Kingdom*

[2] *Mathematical Institute, Woodstock Rd, University of Oxford, Oxford OX2 6GG, United Kingdom*

[3] *Department of Veterinary Medicine, Madingley Road, University of Cambridge, Cambridge CB3 0ES, United Kingdom*

(Dated: October 23, 2022)

Bacteria evolving in natural and clinical settings experience spatial fluctuations of multiple factors and this heterogeneity is expected to affect bacterial adaptation. Notably, spatial heterogeneity in antibiotic concentrations is believed to accelerate the evolution of antibiotic resistance. However, current literature overlooks the role of cell motility, which is key for bacterial survival and reproduction. Here, we consider a quantitative model for bacterial evolution in antibiotic gradients, where bacteria evolve under the stochastic processes of proliferation, death, mutation and migration. Numerical and analytical results show that cell motility has major effects on bacterial adaptation. If migration is relatively rare, it accelerates adaptation because resistant mutants can colonize neighbouring patches of increasing antibiotic concentration avoiding competition with wild-type cells; but if migration is common throughout the lifespan of bacteria, it decelerates adaptation by promoting genotypic mixing and ecological competition. If migration is sufficiently high, it can limit bacterial survival, and we derive conditions for such a regime. Similar patterns are observed in more complex scenarios, namely where bacteria can bias their motion or switch between motility phenotypes either stochastically or in a density-dependent manner. Overall, our work reveals limits to bacterial adaptation in antibiotic landscapes that are set by cell motility.

Active motility is a defining feature of many cell types, governing their ecology and physiological functions. Immune cells such as neutrophils patrol multicellular organisms in their relentless search for invading microbes. Sperm cells actively swim and search ovules to fuse with. Growth cones of neurons seek their synaptic targets. Active motility is also fundamental for many unicellular organisms ranging from bacteria to amoeba and algae. It allows them to find nutrients, light or a host, and avoid toxic compounds, predators or parasites. In particular, bacterial motility has been thoroughly studied once it became understood as key for the reproductive success of bacteria and, more specifically, for their ability to cause disease [1–7]. Yet, despite these important realizations, there is surprisingly little evidence about how motility contributes to bacterial adaptation, namely to the evolution of antibiotic resistance, which is ranked among other major threats such as climate change and terrorism [8].

Laboratory experiments suggest that bacterial motility is important for bacterial adaptation in antibiotic landscapes where cells can move through different concentrations of antibiotics [9–11]. In particular, it was found that in spatially heterogeneous environments, bacteria evolve antibiotic resistance faster than in homogeneous conditions [9]. However, cell motility was not controlled in these experiments, and its precise contribution to the evolutionary dynamics found is not known. Notably, other authors did not find an accelerated adaptation in their antibiotic landscapes and argued that the discrepancy likely derived from the ability of bacteria to experience the diverse antibiotic concentrations, which was different in their experiments [10].

Arguably, the best understood link between cell motility and the ability of bacteria to cope with antibiotics comes from bacterial swarming, a form of group motility on surfaces where cells are particularly resilient to antibiotic stress [12, 13]. More precisely, in swarming conditions, cell motility is thought to reduce exposure of individual cells to antibiotics, which leads to antibiotic tolerance [14]. While there are many studies relating bacterial swarming and phenotypic resistance [12, 14–21], the effect of swarming motility on the evolution of genetic resistance has not been explored. In addition to swarming, other links between bacterial motility and antibiotic landscapes have been reported. In particular, it has been shown that antibiotics trigger motile responses [22–24]. However, the impact of such bacterial behaviour on the evolution of antibiotic resistance was not addressed. In short, one can find works that study bacterial evolution in antibiotic landscapes, but these do not explore the role of cell motility explicitly; and one can find works that study bacterial motility upon antibiotic exposure, but these do not explore evolutionary timescales.

In addition to these experimental works, several mathematical models have been developed to understand how concentration gradients of antibiotics and other drugs affect resistance evolution [25–31], and a key conclusion is

* n.m.oliveira@damtp.cam.ac.uk, corresponding author

that antibiotic gradients accelerate the evolution of antibiotic resistance [25, 26]. These quantitative models consider bacterial movement between environments with different antibiotic concentrations, such as different organs or parts of the human body, but the migration considered is essentially a passive and rare event. This assumption is in sharp contrast with the migration rates of bacteria evolving in natural and experimental settings, where cells can move at speeds up to $\sim 50\mu m/s$, and the fact that in these environments bacteria can experience steep and stable gradients of antibiotics and other drugs [9–11, 22]. Here, we build upon these quantitative models and study how bacterial motility affects the evolution of antibiotic resistance in spatially heterogeneous environments with different concentrations of antibiotics.

I. RESULTS

A. Model overview and basic dynamics.

We are interested to understand how bacterial motility shapes the evolution of resistance in heterogeneous environments where cells experience spatial gradients of antibiotics. We build upon the so-called staircase model [25], which was developed to understand bacterial evolution in drug gradients. The staircase model is a lattice model, which considers cells of specified genotype $g \in \{1, \ldots, L\}$ (of increasing resistance) and spatial compartments $x \in \{1, \ldots, L\}$ (of increasing antibiotic concentration), Fig. 1a. An initial population of susceptible cells $g = 1$ at position $x = 1$ evolves via stochastic processes of death (rate δ), movement to a neighbouring spatial compartment (rate ν), mutation (rates μ_f/μ_b for forward/backward), or division (maximal rate r) [25]. In this model, antibiotics inhibit the growth of susceptible genotypes $g < x$. Therefore, the diagonal of position-genotype lattice in Fig. 1a defines a staircase which separates areas where cells can divide (above the staircase, white) from areas where they cannot (below the staircase, grey). Moreover, the division rate of the N_x cells at position x is logistic with a carrying capacity K:

$$\begin{cases} r(1 - N_x/K), & \text{if } x \leq g \text{ and } N_x \leq K, \\ 0, & \text{otherwise.} \end{cases}$$

The processes of birth, death, mutation and migration drive bacterial adaptation up the antibiotic gradient. Assuming low mutation rates [25, 26], this adaptation corresponds to random jumps $R \to R+1$ between relatively stable population states $R \in \{1, \ldots, L\}$, which we define as the resistance states (Fig. 1b). We note, however, that R can be understood as both the genotype g of highest resistance in the population and the compartment of highest antibiotic concentration x where this genotype can divide because the genotypic and spatial dimensions are interchangeable in our model. Genotypes $g \leq R$ are defined as wild-type and genotype $g = R + 1$ as mutant. Resistance R increases to $R + 1$ when mutants outcompete the wild-type in the overlap region at a position $x = R + 1$. The overlap region is the region where the population first adapts because mutant cells can divide there but wild-type cells cannot (Fig. 1b).

We start by studying how different motility rates ν affect the evolutionary dynamics of bacteria in this model as previous work [25, 26] only considers situations where bacteria are unlikely to move between different antibiotic concentrations during their lifetime ($\nu < \delta$). In particular, we start by describing and comparing bacterial adaptation when cells rarely move during their lifespan ($\nu < \delta$) and situations where motility is common ($\nu > \delta$). Unexpectedly, our simulations show that bacterial adaptation is qualitatively different for low and high motility (Fig. 1c, SI Movie S1, S2). To understand this difference, we first study the qualitative properties of the stable population at a fixed resistance state R and in the next section we focus on the adaptation rate a_R at which the jump in resistance $R \to R + 1$ occurs (Fig. 1b).

Notably, we find that the low and high motility regimes differ in the shape of the wild-type population N_x (SI Theorem 1), which impacts the mutant fitness landscape (Fig. 1d) that is defined by:

$$\begin{cases} r(1 - N_x/K) - \delta, & \text{if } x \leq R + 1, \\ -\delta, & \text{otherwise.} \end{cases}$$

Moreover, we find that at low motility ($\nu < \delta$, SI Movie S1), bacterial adaptation is limited by the movement of the first mutant into the overlap region, where mutant cells have low competition from wild-type cells and high fitness to set up a mutant population (Fig. 1d). In this case, bacterial adaptation is located at the population front (SI Theorem 3), while the rest of the population forms a typical inclined comet tail (Fig. 1c) corresponding to a diversity of strains with different antibiotic susceptibility (Fig. 1d). At high motility ($\nu > \delta$, SI Movie S2), the wild-type is spread across space (SI Theorem 3), competing with mutants for space even in regions where the wild-type cannot grow. The fitness of mutants is decreased in the overlap region and leveled across the region where mutants can divide (Fig. 1d). Therefore, mutant cells grow slowly and outcompete the wild-type everywhere. In these conditions, the population is made of a single strain that can grow in $x \leq R$ (Fig. 1d) and the comet tail is now horizontal (Fig. 1c).

B. Cell motility accelerates and decelerates bacterial adaptive evolution.

We have shown that bacterial adaptation in antibiotic gradients is qualitatively different at low and high motility regimes. To gain quantitative understanding about this difference, we now study how motility affects the adaptation rate a_R (Fig. 1b, 2a), which is defined by two waiting times,

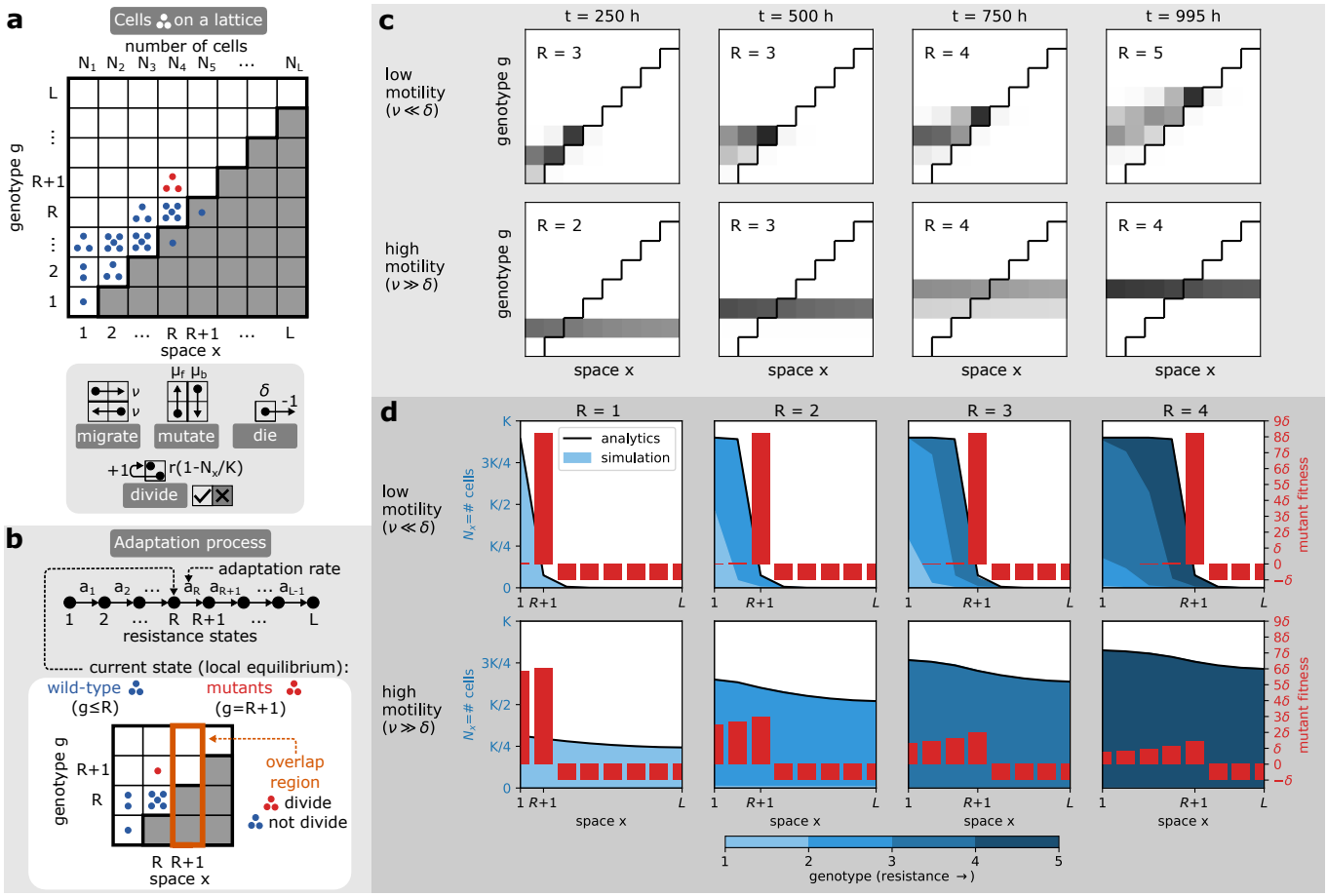


FIG. 1. **Model overview and basic evolutionary dynamics a)** Model description. In the staircase model, bacterial cells (black dots) have a specified position x (of increasing antibiotic concentration) and a genotype g (of increasing resistance). The population adapts to a stable spatial antibiotic gradient as cells migrate, mutate, die and divide at specified stochastic rates. Antibiotics prevent cell division of susceptible genotypes $g < x$ (shaded region under the staircase). **b)** Adaptation process. Bacterial adaptation in the staircase model can be conceptualised as a series of random jumps between locally stable resistance states R that happen at an adaptation rate a_R. At a given time, R denotes the genotype of highest resistance in the population, as well as the highest spatial compartment where this genotype can divide. Genotypes $g \leq R$ are defined as wild-type and genotype $g = R+1$ as mutant. Importantly, in the overlap region $x = R+1$, mutant cells can divide but wild-type cells cannot. **c)** Snapshots of the population evolving antibiotic resistance for different motility rates, ν. Shades of grey represent density of cells, where black represents highest density. The adaptive dynamics of bacteria with low (top) and high (bottom) motility is qualitatively different. **d)** Population profiles and fitness along the antibiotic gradient. Different motility regimes affect the distribution of wild-type bacteria along the gradient, which shapes the fitness of resistant mutants.

1. The evolutionary time T_{evo}, defined as the earliest time when the wild-type produces a first mutant in the overlap region (Fig. 2a).

2. The ecological time T_{eco}, defined as the earliest time when the first mutant establishes a mutant population that becomes larger than the wild-type in the overlap region (Fig. 2a).

These waiting times are random variables and their expected value contribute to the adaptation rate as follows:

$$a_R = \frac{1}{\mathbb{E}T_{evo} + \mathbb{E}T_{eco}}.$$

To determine waiting times and adaptation rates, we used computer simulations (Fig. 2b), and a combination of analytical and numerical techniques that confirm and extend our simulations (Fig. 2b, Fig. 2c and Methods).

How does the adaptation rate a_R vary in the two motility regimes we identified earlier? Our simulations show that low motility ($\nu < \delta$) accelerates the adaptation rate of bacteria (Fig. 2b), in accordance with previous theory [25, 26]. Our analysis shows that in these conditions the evolutionary time is much bigger than the ecological time (Fig. 2c), because mutants have high fitness in the overlap region and can grow quickly (Fig. 2d). For this reason, the ecological time has been neglected in previous works [26] and the adaptation rate has been defined by the evolutionary time alone as $a = 1/\mathbb{E}T_{evo}$, irrespective

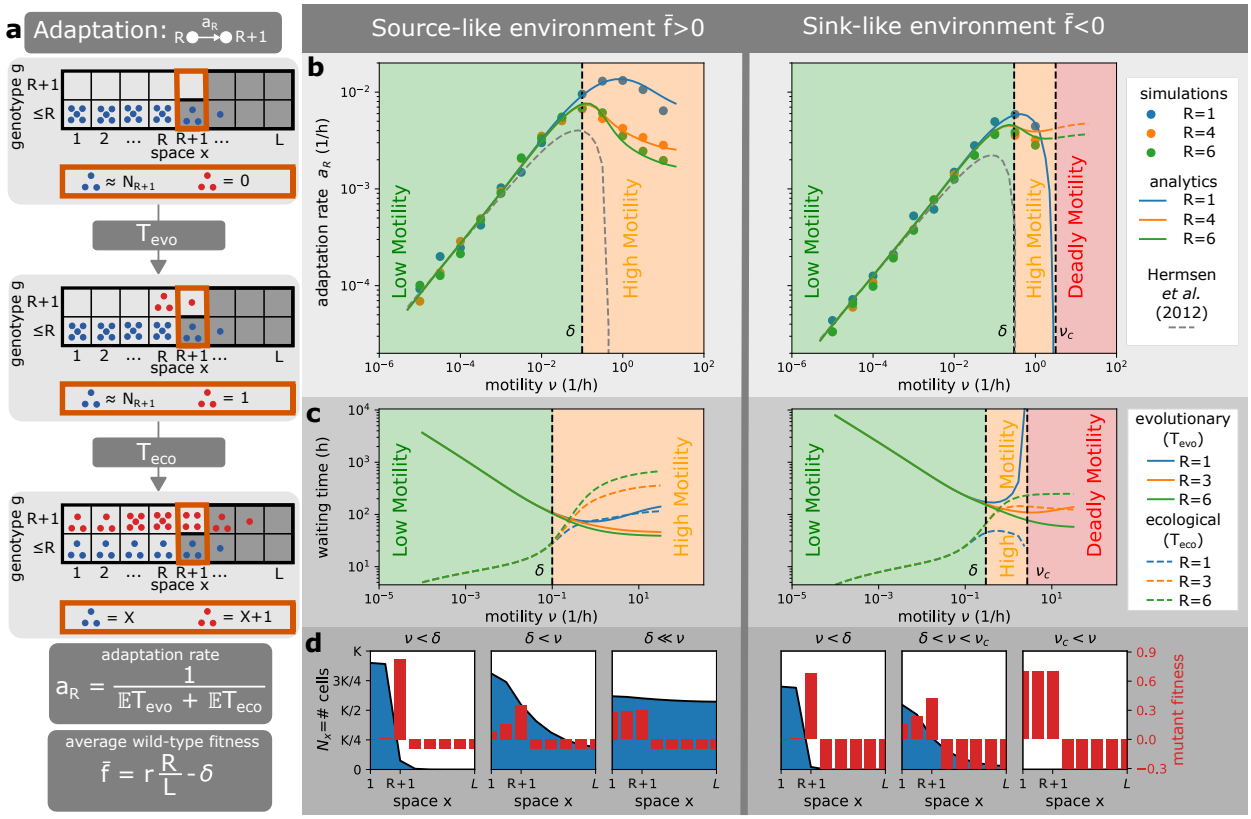


FIG. 2. **Motility can accelerate and decelerate the rate of bacterial adaptation to antibiotics. a)** Definitions. The adaptation process in the staircase model corresponds to jumps between resistance states $R \to R+1$ (Fig. 1b) that happen at the rate a_R. This adaptation rate is defined by two waiting times, the evolutionary time T_{evo} (the time it takes for a first mutant to appear in the overlap region) and the ecological time T_{eco} (the time it takes for mutants to have larger population than the wild-type in the overlap region). The average wild-type fitness is defined as $\overline{f} = rR/L - \delta$ with r being the birth rate, δ death rate, and R/L the proportion of space where wild-type cells can divide. **b)** Adaptation rate as a function of motility rate. Motility accelerates adaptation in the low motility regime, but decelerates adaptation in the high motility regime. If the wild-type has a negative average fitness and motility above a critical motility ν_c, it cannot survive and adapt (deadly motility regime). The current resistance state R does not affect the adaptation rate when motility is low, but decreases the adaptation rate when motility is high. We plot the predicted dynamics from the analytical theory of [26] (dashed line), which matches our results for lower motility only, but note that this theory was developed under the assumption of low motility rates and the interpolation for higher motility shown for completeness is not valid. **c)** Relative importance of evolutionary and ecological times. At low motility, bacterial adaptation is largely dependent on the evolutionary time T_{evo}. At high motility, however, the ecological time T_{eco} becomes equally important due to increased competition between strains. Near critical motility ν_c, bacterial adaptation is again mainly limited by the evolutionary time T_{evo}. At high motility, the number of compartments where the wild-type can divide R increases the ecological time T_{eco} and decreases the evolutionary time T_{evo}. **d)** The importance of wild-type profiles. At low motility regimes, motility promotes the movement of mutants from the population front into the overlap region (T_{evo} reduced), where mutants have larger fitness than the wild-type and grow quickly ($T_{evo} \gg T_{eco}$). At high motility regimes, motility promotes the spatial spread of the wild-type population with genotypes $g \leq R$, so that the first mutant is produced quickly but mutant fitness is decreased by increased competition with the wild-type ($T_{evo} < T_{eco}$).

of the resistance state R. Indeed, [26] that considered low motility only noticed that the rate at which the population front advances becomes constant (i.e., it is independent of R) and our work supports this claim as the adaptation rate a_R is identical for all R when motility is low (Fig. 2b) and the shape of the population front is independent of R (Fig. 1d, SI Theorem 3). Under these conditions, motility helps the first mutant to move from the population front into the overlap region and decreases the dominant waiting time, the evolutionary time T_{evo}. Therefore, in this low motility regime, cell motility accelerates bacterial adaptation. In contrast, high motility ($\nu > \delta$) decelerates the adaptation rate (Fig. 2b). In this regime, the ecological time becomes bigger than the evolutionary time (Fig. 2c) because mutant fitness is significantly decreased, which results from the increased competition for space with a large wild-type population in the overlap region. This competition is amplified with increasing motility ν and resistance state R, which promote the spread and size of the population (Fig. 2c).

The reduction in adaptation rate with higher motility becomes even more important if the wild-type has negative spatially-averaged fitness, which is defined as $\overline{f} = rR/L - \delta$ with r being the birth rate, δ death rate, and R/L the proportion of space where wild-type cells can divide. By definition, environments with this property, $\overline{f} < 0$, cannot sustain the wild-type if the system is spatially homogeneous, and we call them sink-like (as opposed to source-like with $\overline{f} > 0$). This terminology is inspired by the source-sink theory, which defines a source (resp. sink) as an environment with positive (resp. negative) local fitness [32], as opposed to the average fitness in our model that considers fitness across all compartments. Notably, we find that cell motility creates sufficient mixing in sink-like environments above a critical motility threshold ($\nu > \nu_c$), and that in these conditions the wild-type cannot survive and adapt (Fig. 2b, SI Theorem 1, SI Theorem 4, SI Corollary 1). Therefore, we call it a deadly motility regime. We note that the deadly motility regime would not occur if resistant cells were present in the population from the beginning, which would increase the average wild-type fitness from $\overline{f} < 0$ to $\overline{f} > 0$ due to higher R. In this case, the same environment would become source-like, with survival and adaptation possible for any motility (dashed lines in Fig. 2b). We further note that the existence of the deadly motility regime requires the system to be closed, with no external import of wild-type cells.

In summary, our analysis identifies three distinct regimes of bacterial adaptation in antibiotic gradients that depend on the motility rate of cells, which we call low motility, high motility and deadly motility regimes (Fig. 2a). Notably, we find the same regimes when antibiotic resistance carries fitness costs (Fig. S1, SI Text S4), and when instead of bacteriostatic antibiotics we consider bactericidal antibiotics (Fig. S2, SI Text S4). Unexpectedly, we find the same regimes even when cell motility is biased (Fig. S3, SI Text S4, SI Movies S3 and S4). It has recently been shown that antibiotics can trigger both positive [22] and negative [24] chemotaxis in bacterial cells, but our results show that such biased motility does affect our key findings (Fig. S3).

C. Evolutionary dynamics of phenotypically heterogeneous populations.

So far, we have considered evolving populations that are phenotypically homogenous and bacterial cells always move at the same rate. The same simplifying assumption has been made in previous models of bacterial evolution in antibiotic gradients [25–31]. However, in reality, bacteria can change their motility rate, and their communities are phenotypically heterogeneous in terms of cell motility. For example, one of the most important transitions in a bacterial life-cycle from free-living (plankton) to surface-attached communities (biofilms) is associated with major changes in cell motility. Planktonic bacteria can move at $45\mu m/s$ [2], while biofilm bacteria can move at $1\mu m/min$ or less [33]. The plankton-biofilm switch can occur both stochastically [34] or as a response to multiple factors, such as cell density [35] or the presence of other strains [36]. Another well-known transition in bacterial populations that is associated with changes in cell movement is the swarming phenotype, where bacteria move collectively on semi-solid surfaces when a threshold of cell density is reached [37]. We found it important to to understand if this phenotypic heterogeneity could affect the evolutionary dynamics we presented earlier.

To address this issue, we now consider populations harbouring two motility phenotypes where cells can switch between phenotypes either stochastically (Fig. 3) or as a response to local cell density (Fig. 4). In practice, for stochastic switching, we add a new dimension of motility phenotypes to our position-genotype lattice and allow cells to switch between phenotypes at a stochastic rate s, where each phenotype is characterized by different motility rates, ν_1 and ν_2 (Fig. 3a). Density-dependent motility $\nu(N_x)$ is modeled differently, and the switch between motility phenotypes is modeled by a step-function that is controlled by the local number of cells N_x and a switching density threshold S (Fig. 4a). If cell density N_x at position x is below (resp. above) the switching threshold S, all cells at position x move at low-density (resp. high-density) motility ν_L (resp. ν_H).

Interestingly, we find that in both models, phenotypically heterogeneous populations of bacteria generically behave as if all cells had a single effective motility phenotype whose motility is controlled by the switching rate s in stochastic switching, or the switching threshold S if switching is density-dependent. Therefore, the previously described adaptation regimes exist in these models and are governed by an effective motility that we characterize next. For stochastic switching, the effective motility matches the motility of the slower phenotype $\nu_- = \min(\nu_1, \nu_2)$ when the switching rate is low ($s \ll \delta$, SI Movie S5) and it matches the average motility $\nu_+ = (\nu_1 + \nu_2)/2$ when the switching rate is high ($s \gg \delta$, SI Movie S6), see Fig. 3b and SI Theorems 6, 7, 8 and SI Corollary 2, 3. This result stems from the fact that the slower phenotype is naturally selected for in the absence of phenotypic switching, given that the slower phenotype is exposed to antibiotics less often and maximises the local fitness of the wild-type, and stochastic switching levels out the numbers of both phenotypes (Fig. 3c). For density-dependent motility, the effective motility matches the low-density motility phenotype $\nu_- = \nu_L$ at high switching threshold ($S > K(1 - \delta/r)$) and the high-density motility phenotype $\nu_+ = \nu_H$ at low switching thresholds ($S < K(1-\delta/r)$), see Fig. 4b. This result stems from the fact that the population cannot reach a sufficient density for the high-density motility phenotype when the threshold is high, while a lower threshold switches the dense bulk of the population into the high-density motility phenotype.

Accordingly, if the population harbours phenotypic

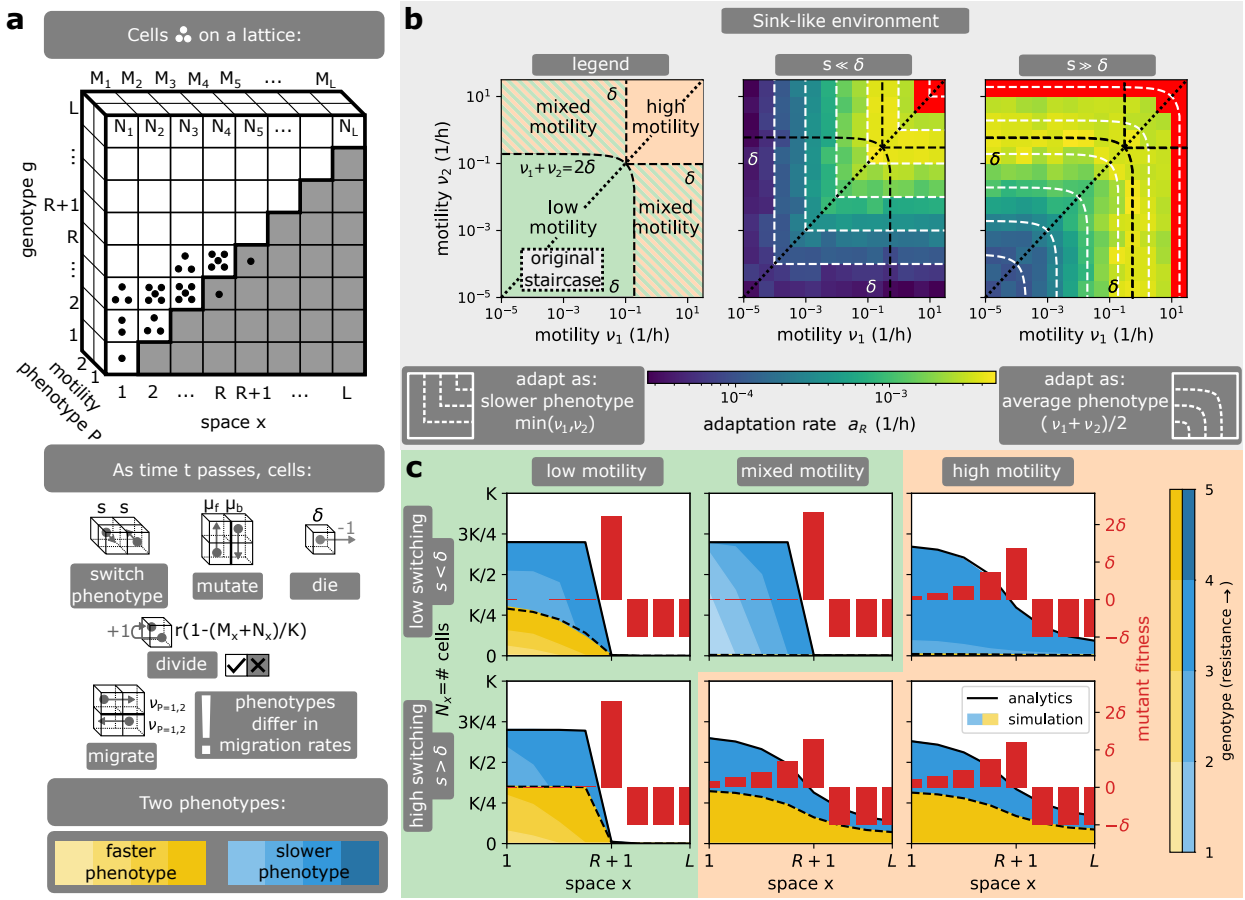


FIG. 3. **Effect of switching between motility phenotypes stochastically. a)** Model description. Phenotypes of motility $\nu_{1,2}$ are added as an extra dimension to the staircase model, with individual cells allowed to switch between these phenotypes at rate s. **b)** Adaptation rate heatmap on the (ν_1, ν_2) plane for different switching rates s. The (ν_1, ν_2) plane can be partitioned into different combinations of adaptation regimes and its diagonal corresponds to a population of a single motility. In this model, bacteria adapt as if all cells had the same effective motility, which corresponds to the intersection of the level sets (white dashed lines) with the diagonal. At low switching rate s, the effective motility matches the slower motility present in the population $\min(\nu_1, \nu_2)$. At high switching rate s, the effective motility matches the average motility $(\nu_1 + \nu_2)/2$. When the environment is sink-like, a deadly motility regime (red) exists in the high motility combination, and also appears in the mixed motility combination when the switching rate is high. **c)** Wild-type profiles for different switching rates s. At low switching rate s, the wild-type profile is dominated by the slower phenotype. At high switching rate s, the wild-type profile coincides with the profile of a single average-motility phenotype. Therefore, the adaptation in low (resp. high) motility combinations follows the low (resp. high) motility regime irrespective of switching rate s, but a change of s in the mixed motility combination can change the adaptation regime. In short, stochastic motility switching shapes the evolution of antibiotic resistance by determining the effective motility of bacterial populations.

heterogeneity of the same motility regime, the adaptation regime corresponding to the effective motility is not changed by the stochastic rate s and the density threshold S. However, for populations that harbour motility phenotypes from different motility regimes (mixed motility in Fig. 3b, Fig. 4b), the level of switching shapes the adaptation regime of bacteria (Fig. 3c, Fig. 4c). For stochastic switching, such heterogeneous populations have an effective motility in the low motility regime at low switching ($\nu_- < \delta$) and in the high (resp. deadly) motility regime at high switching ($\nu_+ > \delta$, resp. $\nu_+ > \nu_c$). Put differently, a gradual increase in switching rate changes the adaptation regime between the low, high and deadly motility regimes (Fig. 3c, S4a, SI Theorems 6, 8). Similarly to phenotypically homogeneous populations, the deadly motility regime can only occur if the average wild-type fitness is negative (sink-like environment, Fig. 3b, S4c, SI Theorem 8, SI Corollary 3).

For highly heterogenous populations with density-dependent motility, the gradual change in switching threshold S also dictates the effective motility and the corresponding adaptation regime, but there are three exceptions when the density threshold S is low (Fig. 4b). First, when cells switch from fast to slow above the

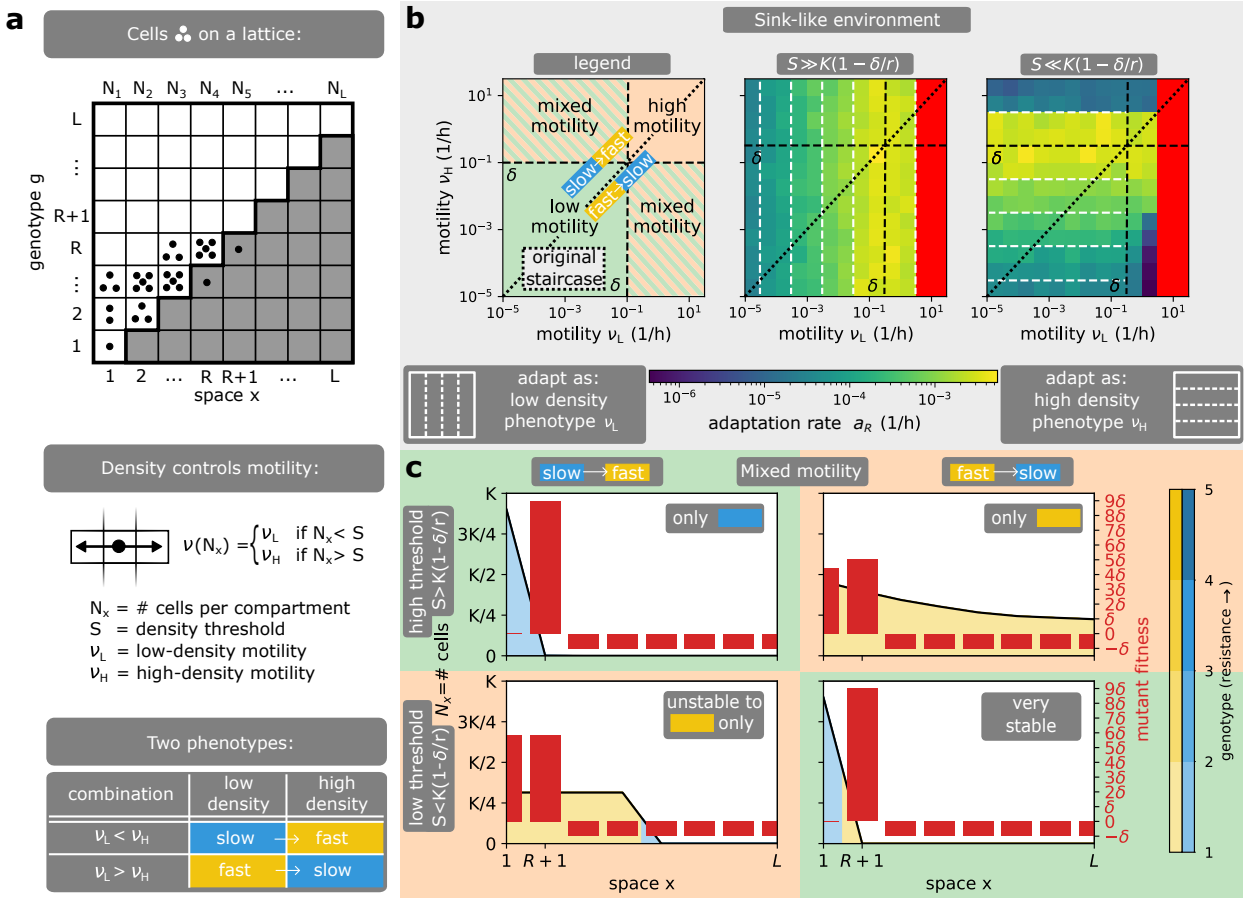


FIG. 4. **Effect of density-dependent motility. a)** Model description. In this model, bacterial motility $\nu(N_x)$ depends on the local number of cells N_x. If the cell density N_x at position x is below (resp. above) the switching threshold S, all cells at position x move at low-density (resp. high-density) motility ν_L (ν_H). **b)** Adaptation rate heatmap on the (ν_L, ν_H) plane for different switching thresholds S. The (ν_L, ν_H) plane can be partitioned into different combinations of adaptation regimes and its diagonal corresponds to a population of a single motility. The diagonal separates slow-to-fast and fast-to-slow switching combinations. Bacteria generically adapt as if all cells had the same effective motility, which corresponds to the intersection of the level sets (white dashed lines) with the diagonal. This effective motility generically matches the low-density (resp. high-density) motility at high (resp. low) threshold S. At low threshold S, this generic rule has three exceptions. First, the adaptation rate is reduced in the mixed motility combination of fast-to-slow switching. Second, the adaptation rate is reduced for slow-to-fast switching when the environment is sink-like and the high-density motility phenotype moves above the critical motility. Third, a deadly motility regime can only occur if the initial population is of low density and the low-density motility phenotype moves above the critical motility. **c)** Wild-type profiles in the mixed motility combination. Wild-type profiles match the low-density motility phenotype at high threshold S. At low threshold S, the low-density motility phenotype appears only at the front while the bulk of the population is in the high-density motility phenotype. This composition of phenotypes can reduce the antibiotic exposure of the bacterial population: in slow-to-fast switching, the spatial expansion of the fast (yellow) high-density motility phenotype is reduced in sink-like environments; while in fast-to-slow switching, a very slow (blue) and stable population occupies low antibiotic concentrations. In short, density-dependent motility shapes bacterial adaptation by determining the effective motility of bacterial populations and affecting their antibiotic exposure.

threshold S, the population becomes very stable as the fast cells at the front quickly return to the slow population bulk (Fig. 4c). As a result, the population has decreased antibiotic exposure and adaptation rate (Fig. 4b). Second, a deadly motility regime occurs if the low-density motility (not the effective motility) is above the critical motility (Fig. 4b). Third, if the high-density motility is above the critical motility, the population does not go extinct but instead decreases its range and antibiotic exposure (Fig. 4c, SI Movie S7), leading to a decreased adaptation rate (Fig. 4b). Since the last two exceptions are related to critical motility, they are only important in sink-like environments where the wild-type has negative average fitness (Fig. S5).

With the exception of the highly heterogeneous populations with density-dependent motility, we showed that bacterial populations with different motility phenotypes evolve antibiotic resistance as if all cells had the same

effective motility. Therefore, in these conditions, bacterial adaptation is characterised by the same adaptation regimes of phenotypically homogeneous populations. Switching between motility phenotypes can shape the adaptation regime under which bacteria evolve only if the motility of the different phenotypes is very different. But even in these cases, we find that our general conclusions hold, low effective motility accelerates bacterial adaptation while high effective motility decelerates it. For very high effective motility, bacterial populations can perish.

D. Cell motility as a factor that limits bacterial survival in sink-like environments.

We have seen how cell motility shapes the evolution of antibiotic resistance in a variety of models. We have shown that the wild-type profiles form the basis of bacterial adaptation by creating mutants and dictating their fitness in antibiotic landscapes (Fig. 1d). Moreover, all our models admit a deadly motility regime, defined by a critical motility (Fig. 2b, Fig. 3b, Fig. 4b). Above the critical motility, the wild-type goes extinct and adaptation is not possible. The goal of this last section is to understand the emergence of such critical motility regime.

We start by studying analytically how the wild-type profile forms at a fixed resistance state R (Fig. 1b, 1d). A wild-type profile $(N_1, \ldots, N_L)$ can be conceptualised as a point in the profile space of cell numbers (Fig. 5a) and its time-evolution can be described by the mean-field equations that are derived from the stochastic dynamics. We illustrate the general approach (SI Theorems 2, 3) using the staircase model with only two compartments $(L = 2, R = 1)$, which is known as source-sink model [25]. In this model, the mean-field equations for the wild-type profile formation are the following

$$\begin{aligned} \dot{N}_1 &= r(1 - N_1/K)N_1 - (\delta + \nu)N_1 + \nu N_2, \\ \dot{N}_2 &= -(\delta + \nu)N_2 + \nu N_1. \end{aligned} \tag{1}$$

This system can be analysed with dynamical systems theory, which studies qualitative features of systems of ordinary differential equations (ODE). More precisely, the time-evolution of this system on $\mathbb{R}^2$ can be characterised by phase portraits (Fig. 5b), which in this case have two fixed-points: a trivial fixed point $N_1 = N_2 = 0$ where the population goes extinct, and a non-trivial fixed-point where the population survives, which is defined as

$$\begin{aligned} N_1 &= K\left(1 - \frac{\delta(\delta + 2\nu)}{r(\delta + \nu)}\right), \\ N_2 &= \frac{\nu}{\nu + \delta} N_1. \end{aligned} \tag{2}$$

These phase portraits admit only two distinct types of trajectory flows (Fig. 5b). Either the non-trivial fixed point has positive physical densities $N_x > 0$ and globally attracts all possible wild-type profiles with $N_x > 0$ (i.e., the bacterial population survives); or the non-trivial fixed point is non-physical $N_x < 0$ and the trivial fixed point attracts all possible wild-type profiles with $N_x > 0$ (i.e., the bacterial population does not survive). These two types of phase portraits appear at different motility rates ν. Furthermore, we can understand this dynamics with bifurcation theory, and study how the non-trivial fixed point moves through the profile space as motility ν increases from zero (Fig. 5c). At low motility, the non-trivial fixed point is stable and located in the non-negative quadrant $N_{1,2} \geq 0$. As motility increases, the dynamics is different in source-like $(\overline{f} > 0)$ and sink-like $(\overline{f} < 0)$ environments, where the average wild-type fitness of the source-sink model is $\overline{f} = r/2 - \delta$. If the environment is source-like, the non-trivial fixed point never leaves the quadrant $N_{1,2} > 0$ and remains stable. However, in sink-like environment, the non-trivial fixed point collides with the trivial fixed point through a transcritical bifurcation, which exchanges their stability. In these cases, there is a critical motility,

$$\nu_c = \frac{\delta(r - \delta)}{2\delta - r} > 0 \tag{3}$$

which corresponds to this bifurcation. The emergence of a bifurcation is necessary, since a) the trivial fixed point must be stable as $\nu \to \infty$ when cells experience an average sink-like environment, and b) the non-trivial fixed point must be stable as $\nu \to 0$ when cells grow locally in the source at $x = 1$ (SI Theorem 3). Therefore, these fixed points must collide and exchange their stability through a transcritical bifurcation. This bifurcation theory argument is general and provides a robust mechanism for the emergence of the critical motility regime in all our models (SI Text S2).

In short, and put differently in a more biological realm, a bacterial population can be driven to extinction by increasing the rate at which cells move in sink-like environments (i.e., in spatially heterogeneous environments where the average wild-type fitness is negative) because highly motile cells experience an average environment.

II. DISCUSSION

In this work, we have used mathematical analysis to study how bacterial motility affects the evolution of antibiotic resistance in spatially heterogeneous environments that have different antibiotic concentrations. By doing so, our study contributes to the large body of literature that has recognized spatial heterogeneity as an important factor for bacterial adaptive evolution but where cell motility has largely been overlooked [9–11, 25–31].

We have identified three regimes of bacterial adaptation, which are defined by the degree of cell motility in evolving populations. Accordingly, we called them low motility, high motility and deadly motility regimes (Fig. 2b). The theory of the low motility regime has already been discussed [25, 26], and has been compared with experimental works that studied how antibiotic gradients

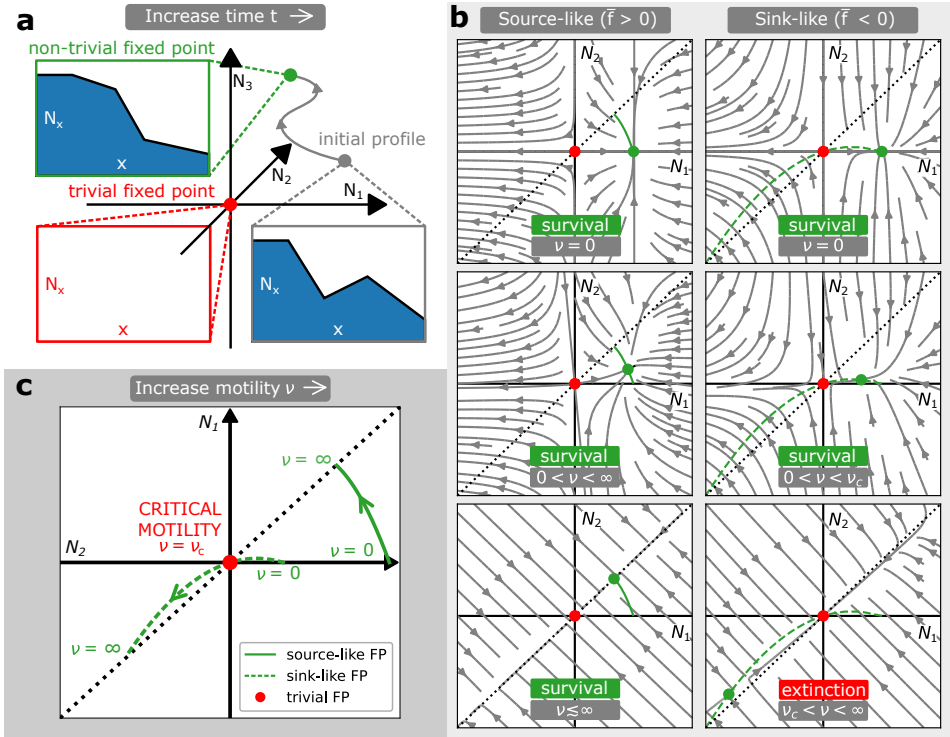


FIG. 5. **Cell motility as a factor that limits bacterial survival in sink-like environments. a)** The formation of the wild-type profile. The formation of a wild-type profile corresponding to a fixed resistance state R (Fig. 1b) is described by mean-field equations, which explain the flow of the initial profile in the profile space of cell numbers $(N_1, \ldots, N_L)$. This flow admits two fixed points: a trivial fixed point $N_x = 0$ (red, population extinction if stable) and a non-trivial fixed-point (green, population survival if stable). **b)** Phase portraits of the wild-type profile formation in the staircase model with $L = 2$, $R = 1$ (the so-called source-sink model [25]). Phase portraits are shown for different motility rates ν and different environment types (source-like/sink-like), depending on the average wild-type fitness $\overline{f} = r/2 - \delta$. Only two topologically distinct types of flow are possible: the non-trivial fixed point globally attracts all possible wild-type profiles with $N_x > 0$ (population survival), or the trivial fixed point globally attracts all possible wild-type profiles with $N_x > 0$ (population extinction). Extinction occurs in sink-like environments ($\overline{f} < 0$) with motility above the critical motility $\nu > \nu_c$. **c)** Critical motility ν_c corresponds to a bifurcation of this dynamical system. When motility ν is varied, the non-trivial fixed point moves through the space of possible profiles. At low motility, the non-trivial fixed point is stable and corresponds to a wild-type profile that predominantly occupies spatial positions $x \leq R$ where it can divide. As motility increases, the wild-type profile gets increasingly levelled across all spatial compartments, and its dynamics becomes governed by the average fitness $\overline{f}$. Precisely when the environment is sink-like $\overline{f} < 0$, the stable non-trivial fixed point (surviving wild-type) collides with the unstable trivial fixed point (extinct wild-type), and they exchange their stability at the critical motility ν_c. In short, highly motile populations experience an average environment, which can drive their extinction if the environment is sink-like.

affect bacterial adaptation rate [9, 10]. Interestingly, while [9] reported an increased adaptation rate in their antibiotic gradient compared to well-mixed conditions, [10] did not find such an acceleration in theirs. We note that while the same bacterial species was used, the antibiotic landscape that cells experience is very different in these two experimental systems, which would be sufficient to affect the adaptation rates according to our work (Fig. 2b). Such difference can emerge when cells move at different rates (Fig. 2b), or for different gradient steepness [28, 31]. Indeed, these effects are interchangeable because increasing gradient steepness brings spatial points of fixed antibiotic concentrations closer, which resembles the effect of increasing effective motility.

As a corollary of our analysis, we find that there is an optimal level of cell motility for bacterial adaptation in spatially heterogeneous environments (Fig. 2b). Similar optimum has been identified in models of HIV [38] and cancer [39] resistance evolution. These models are analogous to our setting in that they consider the movement of viruses or cancer cells between compartments of different drug concentrations. However, the motion of viruses considered in the HIV model is passive as opposed to the active movement of bacteria we consider, and the deadly motility regime was not identified in either of the models. It is not clear if these models do not allow for a deadly motility regime or if instead the authors simply did not explore their model fully. Since our models rest on general features of living cells, they may be useful to understand the evolutionary dynamics of cell types other

than bacteria, namely cancer cells during chemotherapy, where the role of cell motility and drug gradients remains poorly explored [40, 41].

Some of our findings are closely related to those from ecology works that study biological adaptation at range edges where species expand their range in an environmental gradient by dispersal and mutation [42–45]. In this literature, local adaptation is known to be prevented by high motility, the so-called motility load [43], where less fit wild-type genotypes migrate enough to increase competition with mutant genotypes, which decreases selection for the latter. As a result, genetic diversity is lower in their high motility regime when compared to the low motility regime, as we find in our work (Fig. 1d). This loss of genetic diversity in the ecological models is often associated with critical motility [43, 46, 47], which can be compared with our critical motility. Notably, the emergence of a critical motility in these works can be explained by the same bifurcation mechanism that we identified (SI Text S2), which highlights the robustness of our mechanism across modelling frameworks. Moreover, due to its simplicity, the staircase model can be used to provide analytical insights into biological adaptation at range edges as in [42, 43, 48], while considering the important stochasticity of natural processes as in [45].

Our models make multiple simplifying assumptions regarding the biology of bacteria. While we extended our initial staircase model to account for some of this biology, such as the effect of resistance costs (Fig. S1), bactericidal antibiotics (Fig. S2) and biased motility (Fig. S3), these processes could be modelled differently. For example, we modelled resistance costs as a reduction in division rate (Fig. S1), while one could study resistance costs that decrease motility rate. Moreover, we assumed implicitly that bacterial adaptation results from the vertical transmission of resistance genes only, and horizontal gene transfer (HGT) is known to affect the evolution of antibiotic resistance [49]. However, the staircase model is a closed system and it is known that for these systems, HGT does not have an important impact [50–54]. While considering the effect of immigration of cells was beyond the scope of this paper, we confirmed the predictions for closed systems by implementing HGT in the original staircase model. As predicted, we find that HGT does not affect our conclusions (Fig. S6, SI Text S4).

In addition to considering the effect of HGT more thoroughly, one may extend our models to account for other relevant bacterial biology such as the existence of persister phenotypes, which can affect the stability of microbial communities [55], or the effect of phage infection, which can affect bacterial motility [24]. Regardless, our current results lead us to conclude that cell motility limits bacterial adaptive evolution in spatially heterogeneous environments. This realization then suggests that manipulating motility may prove a useful approach for controlling bacterial systems and their impacts on us.

III. METHODS

A. Simulations.

A major component of our methods is computer simulations of the staircase model and its extensions. Hereby, we describe how to simulate the basic staircase model, and note that the same algorithm is used to simulate all the remaining models. The staircase model treats bacterial evolution in drug gradients as a continuous-time Markov process, which is a stochastic process whose future state depends only on the present state and not on the past states. A state of the staircase model is specified by the number of bacterial cells $N_{x,g}$ at each lattice point $(x, g) \in \{1, \ldots, L\}^2$ of the position-genotype lattice. The initial state of the staircase model at time $t = 0$ is chosen as $N_{1,1} = 10^{-3}K$ and $N_{x,g} = 0$ for other lattice points $(x, g) \neq (1, 1)$. The time-evolution of the initial state of the staircase model $N_{x,g}$ is determined by the following processes $i \in I$:

- death of a cell of genotype g at position x ($N_{x,g} \to N_{x,g} - 1$): rate $b_i = \delta N_{x,g}$,
- division of a cell of genotype g at position $x \leq g$ when $\sum_g N_{x,g} \leq K$ ($N_{x,g} \to N_{x,g} + 1$): rate $b_i = r(1 - \sum_g N_{x,g})$,
- movement of a cell of genotype g from position x to position $x + 1$ ($N_{x,g} \to N_{x,g} - 1$, $N_{x+1,g} \to N_{x+1,g} + 1$): rate $b_i = \nu N_{x,g}$,
- movement of a cell of genotype g from position x to position $x - 1$ ($N_{x,g} \to N_{x,g} - 1$, $N_{x-1,g} \to N_{x-1,g} + 1$): rate $b_i = \nu N_{x,g}$,
- mutation of a cell at position x from genotype g to genotype $g + 1$ ($N_{x,g} \to N_{x,g} - 1$, $N_{x,g+1} \to N_{x,g+1} + 1$): rate $b_i = \mu_f N_{x,g}$,
- mutation of a cell at position x from genotype g to genotype $g - 1$ ($N_{x,g} \to N_{x,g} - 1$, $N_{x,g-1} \to N_{x,g-1} + 1$): rate $b_i = \mu_g N_{x,g}$.

To simulate this time-evolution, we use the exact sampling of stochastic trajectories of the staircase model via a computationally efficient Next Reaction Method [56], which samples the next process that takes place and the time when it happens. Effectively, this method combines computationally efficient data structures and the Gillespie's algorithm [57]:

1. sample the next process i, which happens with probability $b_i / \sum_{j \in I} b_j$,
2. sample the waiting time Δt, which has distribution $\mathrm{Exp}(\sum_{i \in I} b_i)$,
3. update the time $t \to t + \Delta t$ when the next process i happens and execute the process i.

Using this sampling technique, the evolutionary (resp. ecological) times T_{evo} (resp. T_{eco}) at a resistance state R can be sampled by measuring when mutants appear (resp. outcompete the wild-type) in the overlap region $x = R+1$. By averaging over many independent samples, the expected evolutionary and ecological times can be estimated and the adaptation rate $a_R = 1/(\mathbb{E}T_{evo} + \mathbb{E}T_{eco})$ computed. As in [26], we have used 50 independent samplings. The same algorithm is used to simulate the extensions of the staircase model and the precise parameters used in simulations are detailed in the SI Text S8.

B. Analytical and numerical techniques.

In addition to computer simulations, we developed a combination of analytical and numerical techniques, which are used to compute the evolutionary time, ecological time and the adaptation rate. In contrast to simulations, these techniques are based on closed-form solutions of Master equations, numerical methods for solving algebraic equations (such as Newton-Raphson method) and solving ODEs (such as Runge-Kutta method). To illustrate these techniques, we show how to compute the expected waiting times and the adaptation rate a_R in the original staircase model:

1. **Find the wild-type profile N_x at a resistance state R.** Use Newton-Raphson method to solve the algebraic equations for the non-trivial fixed point of the mean-field theory

$$\begin{aligned} 0 &= r(1 - N_x/K)N_x \mathbb{1}_{x \le R} - \delta N_x \\ &+ \nu(N_{x-1} - N_x)\mathbb{1}_{1<x} + \nu(N_{x+1} - N_x)\mathbb{1}_{x<L}, \end{aligned} \tag{4}$$

   where $\mathbb{1}$ is the indicator function.

2. **Compute the evolutionary time $\mathbb{E}T_{evo}$.** A closed-form solution exists and follows from a modification of the theory in [25, 26]. We consider three independent adaptation paths along which wild-type can produce a first mutant in the overlap region $x = R + 1$. On the position-genotype lattice (x, g), they are the division path D: $(R, R) \to (R, R+1) \to (R+1, R+1)$ (mutation at $x = R$ followed by migration to the right); the overlap path O: $(R+1, R) \to (R+1, R+1)$ (mutation at $x = R+1$); and the no-division path N: $(R+2, R) \to (R+2, R+1) \to (R+1, R+1)$ (mutation at $x = R+2$ followed by migration to the left). As the original theory [26] assumes low motility, it neglects the no-division path since there are no wild-type cells to mutate at position $x = R+2$ in this regime (SI Theorem 3). Moreover, we assume that the flux of mutants between neighbouring compartments in the division region $x \le R$ and the no-divison region $x \ge R+2$ is equilibrated, so that we can ignore the net probability flux of mutants that enter or exit the compartments $x = R-1, R, R+1$ and also ignore longer paths such as $(R-1, R) \to (R-1, R+1) \to (R, R+1) \to (R+1, R+1)$ (mutation at $x = R-1$ followed by double migration to the right). This is a reasonable assumption as the flux of mutant-producing wild-type is equilibrated far from the population front. The computation of $\mathbb{E}T_{evo}$ has the following structure.

   (a) The waiting times till a mutant is produced in the overlap region along path $i \in \{D, O, N\}$ are denoted by T_i (Fig. 2a).

   (b) The probability that the waiting will be longer than t is $S_i(t) = \mathbb{P}(T_i > t)$ and the probability density function of T_i is given by $F_i(t) = -\mathrm{d}S_i/\mathrm{d}t$.

   (c) Notice that $T_{evo} = \min_i T_i$ and that T_i are independent. Thus, $S(t) = \mathbb{P}(T_{evo} > t) = S_D(t)S_O(t)S_N(t)$.

   (d) The probability density function of T_{evo} is

$$\begin{aligned} F(t) = -\frac{\mathrm{d}S}{\mathrm{d}t} &= F_D(t)S_O(t)S_N(t) \\ &+ S_D(t)F_O(t)S_N(t) \\ &+ S_D(t)S_O(t)F_N(t). \end{aligned} \tag{5}$$

   (e) The expected time $\mathbb{E}T_{evo} = \int_0^\infty tF(t)\mathrm{d}t$ is found by numerical integration.

Therefore, the problem is solved by finding the functions $S_i(t)$ and differentiating them to produce $F_i(t)$. According to [25, 26], these functions are:

(a) **For path D:**

$$S_D(t) = \left(\frac{c_D e^{b_D t}}{c_D \cosh(c_D t) + b_D \sinh(c_D t)} \right)^{a_D} \tag{6}$$

where

$$\begin{aligned} a_D &= \frac{\mu_f N_R}{r' - \mu_f}, \\ b_D &= \frac{1}{2}(\nu + \delta - r' + \mu_f + \mu_b), \\ c_D &= \frac{1}{2}[(\nu + r' - \mu_f + \delta + \mu_b)^2 \\ &\quad - 4(r' - \mu_f)(\delta + \mu_b)]^{1/2}, \end{aligned} \tag{7}$$

and

$$F_D(t) = \nu S_D(t)\langle n(t)\rangle_D, \tag{8}$$

where $\langle n(t)\rangle_D$ is the average mutant population in compartment $x = R$ at time t

$$\langle n(t)\rangle_D = \frac{\mu_f N_R \tanh(c_D t)}{c_D + b_D \tanh(c_D t)}. \tag{9}$$

(b) **For path O:**

$$S_O(t) = e^{-\mu_f N_{R+1} t}, \tag{10}$$

$$F_O(t) = \mu_f N_{R+1} e^{-\mu_f N_{R+1} t}. \tag{11}$$

(c) **For path N:**

$$S_N(t) = \\ = \left(\frac{c_N e^{b_N t}}{c_N \cosh(c_N t) + b_N \sinh(c_N t)} \right)^{a_N} \tag{12}$$

where

$$\begin{aligned} a_N &= -N_{R+2}, \\ b_N &= \frac{1}{2}(\nu + \delta + \mu_f + \mu_b), \\ c_N &= \frac{1}{2}[(\nu - \mu_f + \delta + \mu_b)^2 \\ &\quad + 4\mu_f(\delta + \mu_b)]^{1/2}, \end{aligned} \tag{13}$$

and

$$F_N(t) = \nu S_N(t) \langle n(t) \rangle_N, \tag{14}$$

where $\langle n(t) \rangle_N$ is the average mutant population in compartment $x = R + 2$ at time t

$$\langle n(t) \rangle_N = \frac{\mu_f N_{R+2} \tanh(c_N t)}{c_N + b_N \tanh(c_N t)}. \tag{15}$$

3. **Compute the ecological time $\mathbb{E}T_{eco}$.** The competition between wild-type and mutants can be described well by the mean-field equations:

$$\begin{aligned} \frac{\mathrm{d}w_x}{\mathrm{d}t} &= -\mu_f w_x + \mu_b m_x \\ &\quad + r\left(1 - \frac{w_x + m_x}{K}\right) w_x \mathbb{1}_{x \le R} - \delta w_x \\ &\quad + \nu(w_{x-1} - w_x)\mathbb{1}_{1<x} + \nu(w_{x+1} - w_x)\mathbb{1}_{x<L} \\ \frac{\mathrm{d}m_x}{\mathrm{d}t} &= +\mu_f w_x - \mu_b m_x \\ &\quad + r\left(1 - \frac{w_x + m_x}{K}\right) m_x \mathbb{1}_{x \le R} - \delta m_x \\ &\quad + \nu(m_{x-1} - m_x)\mathbb{1}_{1<x} + \nu(m_{x+1} - m_x)\mathbb{1}_{x<L} \end{aligned} \tag{16}$$

with initial conditions at time $t = 0$:

$$\begin{aligned} w_x &= N_x - \langle n(\mathbb{E}T_{evo}) \rangle_D && , x = 1, \ldots, R \\ w_x &= N_{R+1} - 1 && , x = R + 1 \\ w_x &= N_x - \langle n(\mathbb{E}T_{evo}) \rangle_N && , x = R + 2, \ldots, L \\ m_x &= \langle n(\mathbb{E}T_{evo}) \rangle_D && , x = 1, \ldots, R \\ m_x &= 1 && , x = R + 1 \\ m_x &= \langle n(\mathbb{E}T_{evo}) \rangle_N && , x = R + 2, \ldots, L \end{aligned} \tag{17}$$

In these equations w_x and m_x stand for the number of wild-type and mutant cells in spatial compartment x, respectively. The system is simulated from the end state of the previous stochastic process when a first mutant appears in the overlap region. This system can be simulated efficiently by a Runge-Kutta (RK4) method. The time $t = \mathbb{E}T_{evo}$ is reached when $w_{R+1} < m_{R+1}$ for the first time, i.e., the mutants outcompete the wild-type in the overlap region.

4. **Find the adaptation rate.** The adaptation rate is found from

$$a_R = \frac{1}{\mathbb{E}T_{evo} + \mathbb{E}T_{eco}}. \tag{18}$$

For convenience, we refer to this set of techniques as analytical techniques, even though they implement many numerical procedures. The benefit of the analytical techniques, as opposed to simulations, comes from their computational efficacy. In addition to these form of analytical techniques, we developed a purely analytical approach, which describes the formation of the wild-type profile and this approach is detailed in the Supplementary Text.

REFERENCES

[1] Adler, J. Chemotaxis in bacteria. *Science* **153**, 708–716 (1966).

[2] Berg, H. C. *E. coli in Motion* (Springer-Verlag New York, 2004).

[3] Gude, S. *et al.* Bacterial coexistence driven by motility and spatial competition. *Nature* **578**, 588—592 (2020).

[4] Josenhans, C. & Suerbaum, S. The role of motility as a virulence factor in bacteria. *International Journal of Medical Microbiology* **291**, 605–614 (2002).

[5] Ottemann, K. M. & Miller, J. F. Roles for motility in bacterial–host interactions. *Molecular Microbiology* **24**, 1109–1117 (1997).

[6] Stocker, R., Seymour, J., Samadani, A., Hunt, D. & Polz, M. Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4209–14 (2008).

[7] Yawata, Y. *et al.* Competition–dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations. *Proceedings of the National Academy of Sciences* **111**, 5622–5627 (2014).

[8] Davies, S. C., Fowler, T., Watson, J., Livermore, D. M. & Walker, D. Annual report of the chief medical officer: infection and the rise of antimicrobial resistance. *The Lancet* **381**, 1606–1609 (2013).

[9] Zhang, Q. *et al.* Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Science* **333**, 1764–1767 (2011).

[10] Baym, M., Stone, L. K. & Kishony, R. Multidrug evolutionary strategies to reverse antibiotic resistance. *Science* **351**, aad3292 (2016).

[11] Hol, F. J., Hubert, B., Dekker, C. & Keymer, J. E. Density-dependent adaptive resistance allows swimming

bacteria to colonize an antibiotic gradient. *The ISME journal* **10**, 30–38 (2016).

[12] Kearns, D. B. A field guide to bacterial swarming motility. *Nature Reviews Microbiology* **8**, 634–644 (2010).

[13] Lai, S., Tremblay, J. & Déziel, E. Lai s, tremblay j, deziel e.. swarming motility: a multicellular behaviour conferring antimicrobial resistance. environ microbiol 11: 126-136. *Environmental microbiology* **11**, 126–36 (2008).

[14] Butler, M. T., Wang, Q. & Harshey, R. M. Cell density and mobility protect swarming bacteria against antibiotics. *Proceedings of the National Academy of Sciences* **107**, 3776–3781 (2010).

[15] Bhattacharyya, S., Walker, D. M. & Harshey, R. M. Dead cells release a 'necrosignal'that activates antibiotic survival pathways in bacterial swarms. *Nature communications* **11**, 1–12 (2020).

[16] Coleman, S. R., Blimkie, T., Falsafi, R. & Hancock, R. E. Multidrug adaptive resistance of pseudomonas aeruginosa swarming cells. *Antimicrobial agents and chemotherapy* **64**, e01999–19 (2020).

[17] Kim, W., Killam, T., Sood, V. & Surette, M. G. Swarm-cell differentiation in salmonella enterica serovar typhimurium results in elevated resistance to multiple antibiotics. *Journal of bacteriology* **185**, 3111–3117 (2003).

[18] Partridge, J. D., Ariel, G., Schvartz, O., Harshey, R. M. & Be'er, A. The 3d architecture of a bacterial swarm has implications for antibiotic tolerance. *Scientific reports* **8**, 1–11 (2018).

[19] Overhage, J., Bains, M., Brazas, M. D. & Hancock, R. E. Swarming of pseudomonas aeruginosa is a complex adaptation leading to increased production of virulence factors and antibiotic resistance. *Journal of bacteriology* **190**, 2671–9 (2008).

[20] Roth, D. *et al.* Identification and characterization of a highly motile and antibiotic refractory subpopulation involved in the expansion of swarming colonies of p aeni-bacillus vortex. *Environmental microbiology* **15**, 2532–2544 (2013).

[21] Turnbull, A. L. & Surette, M. G. L-cysteine is required for induced antibiotic resistance in actively swarming salmonella enterica serovar typhimurium. *Microbiology* **154**, 3410–3419 (2008).

[22] Oliveira, N. M. *et al.* Suicidal chemotaxis in bacteria. *Nature Communications.* Accepted (2022).

[23] Liu, Y., Kyle, S. & Straight, P. D. Antibiotic stimulation of a bacillus subtilis migratory response. *Msphere* **3**, e00586–17 (2018).

[24] Bru, J.-L. *et al.* Pqs produced by the pseudomonas aeruginosa stress response repels swarms away from bacteriophage and antibiotics. *Journal of bacteriology* **201**, e00383–19 (2019).

[25] Hermsen, R. & Hwa, T. Sources and sinks: a stochastic model of evolution in heterogeneous environments. *Physical review letters* **105**, 248104 (2010).

[26] Hermsen, R., Deris, J. & Hwa, T. On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient. *Proceedings of the National Academy of Sciences* **109**, 10775 – 10780 (2012).

[27] Greulich, P., Waclaw, B. & Allen, R. J. Mutational pathway determines whether drug gradients accelerate evolution of drug-resistant cells. *Physical review letters* **109**, 088101 (2012).

[28] Hermsen, R. The adaptation rate of a quantitative trait in an environmental gradient. *Physical Biology* **13**, 065003 (2016).

[29] Gralka, M., Fusco, D., Martis, S. & Hallatschek, O. Convection shapes the trade-off between antibiotic efficacy and the selection for resistance in spatial gradients. *Physical Biology* **14**, 045011 (2017).

[30] De Jong, M. G. & Wood, K. B. Tuning spatial profiles of selection pressure to modulate the evolution of drug resistance. *Phys. Rev. Lett.* **120**, 238102 (2018).

[31] Steel, H. & Papachristodoulou, A. The effect of spatiotemporal antibiotic inhomogeneities on the evolution of resistance. *Journal of theoretical biology* **486**, 110077 (2019).

[32] Pulliam, H. R. Sources, sinks, and population regulation. *The American Naturalist* **132**, 652–661 (1988).

[33] Oliveira, N. M., Foster, K. R. & Durham, W. M. Single-cell twitching chemotaxis in developing biofilms. *Proceedings of the National Academy of Sciences* **113**, 6532–6537 (2016).

[34] Sousa, A. M., Machado, I. & Pereira, M. O. Phenotypic switching: an opportunity to bacteria thrive. *Science Against Microbial Pathogens: Communicating Current Research and Technological Advances* **1** (2012).

[35] Miller, M. B. & Bassler, B. L. Quorum sensing in bacteria. *Annual Reviews in Microbiology* **55**, 165–199 (2001).

[36] Oliveira, N. M. *et al.* Biofilm formation as a response to ecological competition. *PLoS biology* **13**, e1002191 (2015).

[37] Kearns, D. B. & Losick, R. Swarming motility in undomesticated bacillus subtilis. *Molecular microbiology* **49**, 581–590 (2003).

[38] Kepler, T. & Perelson, A. Drug concentration heterogeneity facilitates the evolution of drug resistance. *Proceedings of the National Academy of Sciences of the United States of America* **95 20**, 11514–9 (1998).

[39] Fu, F., Nowak, M. A. & Bonhoeffer, S. Spatial heterogeneity in drug concentrations can facilitate the emergence of resistance to cancer therapy. *PLOS Computational Biology* **11**, 1–22 (2015).

[40] Lambert, G. *et al.* An analogy between the evolution of drug resistance in bacterial communities and malignant tissues. *Nature Reviews Cancer* **11**, 375–382 (2011).

[41] Wu, A. *et al.* Cell motility and drug gradients in the emergence of resistance to chemotherapy. *Proceedings of the National Academy of Sciences* **110**, 16103–16108 (2013).

[42] Kirkpatrick, M. & Barton, N. Evolution of a species' range. *The American Naturalist* **150**, 1 – 23 (1997).

[43] Lenormand, T. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* **17**, 183–189 (2002).

[44] Angert, A., Bontrager, M. & Ågren, J. What do we really know about adaptation at range edges? *Annual Review of Ecology, Evolution, and Systematics* **51**, 341–361 (2020).

[45] Polechová, J. & Barton, N. H. Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences* **112**, 6401–6406 (2015).

[46] Bulmer, M. Multiple niche polymorphism. *The American Naturalist* **106**, 254–257 (1972).

[47] Holt, R. D. & Gomulkiewicz, R. How does immigration influence local adaptation? a reexamination of a familiar paradigm. *The American Naturalist* **149**, 563–572 (1997).

[48] Nagylaki, T. Conditions for the existence of clines. *Genetics* **80**, 595–615 (1975).

[49] Niehus, R., Mitri, S., Fletcher, A. G. & Foster, K. R. Migration and horizontal gene transfer divide microbial genomes into multiple niches. *Nature communications* **6**, 1–9 (2015).

[50] Cohan, F. M. The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. *The American Naturalist* **143**, 965–986 (1994).

[51] Cohan, F. M. Does recombination constrain neutral divergence among bacterial taxa? *Evolution* **49**, 164–175 (1995).

[52] Shapiro, B. J., David, L. A., Friedman, J. & Alm, E. J. Looking for darwin's footprints in the microbial world. *Trends in microbiology* **17**, 196–204 (2009).

[53] Gevers, D. *et al.* Re-evaluating prokaryotic species. *Nature Reviews Microbiology* **3**, 733–739 (2005).

[54] Levin, B. R. Periodic selection, infectious gene exchange and the genetic structure of e. coli populations. *Genetics* **99**, 1–23 (1981).

[55] Haas, P. A., Oliveira, N. M. & Goldstein, R. E. Subpopulations and stability in microbial communities. *Physical Review Research* **2**, 022036 (2020).

[56] Gibson, M. A. & Bruck, J. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The journal of physical chemistry A* **104**, 1876–1889 (2000).

[57] Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry* **81**, 2340–2361 (1977).

IV. ACKNOWLEDGEMENTS.

We thank Kacper Kornet and Deryck Thake from the IT team of the Department of Applied Mathematics and Theoretical Physics of the University of Cambridge for their help with the managing of the supercomputing cluster. We thank Julian Parkhill for discussions about the evolution of antibiotic resistance and for suggesting the study of HGT. We thank Kevin Foster, Oscar Despard, George Fortune, Pierre Haas, Julian Parkhill, Robert Austin, Mike Cates and Philip Maini for their comments on earlier versions of this manuscript. V.P. was supported by a CMS Summer Research Grant, and had additional support from Gonville & Caius College. N.M.O. was supported by a BBSRC Discovery Fellowship No. BB/T009098/1.

SUPPLEMENTARY TEXT

The main purpose of the supplementary text (SI text) is to present the analytical study of the formation of wild-type profiles in our models. In the main text, we claim that the wild-type profile is key for bacterial adaptation in an antibiotic gradient as it shapes the fitness landscape of mutants (see Results). The quantitative analysis of the wild-type profile provides the first step to the computation of the evolutionary time, the ecological time and the adaptation rate (see Methods). Therefore, understanding the formation of the wild-type profile is key for both intuitive and quantitative analysis.

This SI text is organised as separate sections for convenience and clarity. We start by restating the analytical results for the wild-type profile of the simple source-sink model presented in the final subsection of Results. Starting from the end of the main text is deliberate because the source-sink model is a toy model, which provides an intuition for the formation of the wild-type profile in other more complex models considered in this work. However, the sections of the SI text are relatively self-contained and can be read in any order. We provide next a short guide to explain the content of each section and its links to other parts of our work:

- Section S1: summary of results about the wild-type profile in the source-sink model presented in Fig. 5 (fourth subsection of Results).
- Section S2: utilization of Section S1 to explain why the bifurcation argument for the emergence of critical motilities applies to a large variety of models (fourth subsection of Results).
- Section S3: wild-type profile in the staircase model presented in Fig. 1 and Fig. 2 (first and second subsections of Results).
- Section S4: modified staircase models and their wild-type profile, including the effect of resistance costs in Fig. S1, bactericidal antibiotics in Fig. S2, bacterial chemotaxis in Fig. S3 (second subsection of Results) and horizontal gene transfer in S6 (Discussion).
- Section S5: wild-type profile in the model of stochastic phenotypic switching presented in Fig. 3 and Fig. S4 (third subsection of Results).
- Section S6: wild-type profile in the model of density-dependent motility presented in Fig. 4 and Fig. S5 (third subsection of Results).
- Section S7: mathematical proofs of all theorems presented in all the SI sections mentioned above.
- Section S8: supplementary methods detailing the exact parameter choices for each figure of our work, including information about main and supplementary figures.

S1. Source-sink model.

In the main text, we studied the wild-type profile of the source-sink model (staircase with $L = 2, R = 1$), described by the dynamical system (1). The phase portraits of the system are shown in Fig. 5b and the results are summarised in Theorem 1.

Theorem 1. *System* (1) *admits exactly two fixed points*

- *trivial fixed point:* $N_2 = N_1 = 0$,
- *non-trivial fixed point:*

$$\begin{aligned} N_1 &= K\left(1 - \frac{\delta(\delta+2\nu)}{r(\delta+\nu)}\right), \\ N_2 &= \frac{\nu}{\nu+\delta} N_1. \end{aligned} \tag{s1}$$

Exactly one of these fixed points is stable and corresponds to a stable wild-type profile, with non-negative cell numbers $N_{1,2} \geq 0$. *Let* $\overline{f} = (r - 2\delta)/2$ *be the average wild-type fitness. The stable wild-type profile corresponds to the non-trivial fixed point iff*

- *the environment is source-like* $\overline{f} > 0$ *and cell motility* ν *is arbitrary, or*

- *the environment is sink-like* $\overline{f} < 0$ *and cell motility is constrained to* $\nu \in [0, \nu_c]$*, where the critical motility is*

$$\nu_c = \frac{\delta(r-\delta)}{2\delta - r} > 0. \tag{s2}$$

There are two implications of Theorem 1. First, this theorem provides an explanation for the existence of a critical motility as explained in the fourth subsection of Results and further discussed in the next SI section S2. Second, it explains the shape of the wild-type profile N_x in Fig. 1d. The population in compartment $x = 2$ is significantly suppressed at low motility ($\nu \ll \delta$) and the population is homogeneously distributed across space at high motility ($\nu \gg \delta$). We elaborate on this phenomenon in the SI section S3.

S2. Robust mechanism for the emergence of the critical motility.

In the main text, we claimed that the critical motility can be understood as a bifurcation point of a dynamical system for the wild-type profile formation and that this mechanism is applicable to a large variety of models. Below we first provide a heuristic argument for why critical motilities shall arise for all models in our paper, and then we make a comparison with critical motilities in ecological models [S1–S3].

Theorems with analogous structure to Theorem 1 shall exist for more complicated models of our paper and shall predict the existence of a critical motility. Heuristically, we can argue as follows. Vary motility $\nu \geq 0$ as the bifurcation parameter. When there is no motility $\nu = 0$, the dynamics between spatial compartments decouples and the non-trivial fixed point is stable while the trivial fixed point is not. As cell motility increases, the non-negative quadrant $N_x \geq 0$ becomes forward invariant and its boundary contains a single fixed point, the trivial fixed point. Assuming that the underlying Markov process is ergodic, a unique stable fixed point must exist within the non-negative quadrant $N_x \geq 0$. This implies that the non-trivial fixed point either stays in the positive quadrant and remains stable, or it leaves the positive quadrant through a transcritical bifurcation with the trivial fixed point, see Fig. 5b. Populations of high motility are generically governed by the average fitness [S4], explaining the stability of the trivial fixed point at high motility in sink-like environments. Therefore, there must exist a critical motility associated with the transcritical bifurcation if the environment is sink-like.

In the main text, we claimed that this bifurcation mechanism also applies to the critical motility in ecological models that study range expansion [S1–S3]. We elaborate on this claim with a migration-selection model from [S3] and relate this model to our source-sink model. Consider two alleles a and A and two habitats, such that allele A is selected for in habitat 1 with strength s and allele a is selected for in habitat 2 with strength αs ($\alpha, s > 0$). Moreover, these alleles can move between habitats at rate ν. The relative frequencies of allele a in habitat $i \in \{1, 2\}$ are denoted by $a_i \in [0, 1]$, and their dynamics can be described by a replicator equation with constant selection:

$$\begin{aligned} \dot{a}_1 &= s a_1(1-a_1) - \nu a_1 + \nu a_2 \\ \dot{a}_2 &= -\alpha s a_2(1-a_2) + \nu a_1 - \nu a_2, \end{aligned} \tag{s3}$$

This is a dynamical system on a square $(a_1, a_2) \in [0, 1]^2$. When $\nu = 0$, there is a globally attracting non-trivial fixed point $(a_1, a_2) = (1, 0)$. When $\nu > 0$, the square $[0, 1]^2$ becomes forward invariant and the boundary supports two trivial fixed points: $(a_1, a_2) = (0, 0)$ (allele A wins) and $(a_1, a_2) = (1, 1)$ (allele a wins). Moreover, as ν increases from 0, the globally attracting fixed point is pushed from the boundary of the square into the interior. Similarly to the source-sink model, the non-trivial fixed point can leave the square only through a transcritical bifurcation with one of the trivial fixed points, giving rise to a critical motility ν_c. A straightforward calculation for the non-trivial fixed point shows that it collides with the trivial fixed point $(a_1, a_2) = (0, 0)$ when $\alpha > 1$ and with the trivial fixed point $(a_1, a_2) = (1, 1)$ when $0 < \alpha < 1$ at

$$\nu_c = \frac{s\alpha}{|1-\alpha|}. \tag{s4}$$

Therefore, in accordance with [S3], both alleles coexist if $\nu < \nu_c$ and the overall more fit allele takes over if $\nu > \nu_c$.

The similarity between this model and the source-sink model is not a coincidence. The source-sink system (1) can be manipulated and rearranged into

$$\begin{aligned} \dot{a_1} &= (r-\delta) a_1(1-a_1) - \nu a_1 + \nu a_2 - \delta a_1^2, \\ \dot{a_2} &= -\delta a_2(1-a_2) + \nu a_1 - \nu a_2 - \delta a_2^2, \end{aligned} \tag{s5}$$

by relabelling $N_i \equiv a_i \in [0, \infty)$. This is similar to (s3) with $s = r - \delta$, $\alpha = \delta/(r-\delta)$ near the trivial fixed point $(a_1, a_2) = (0, 0)$. However, the point $(a_1, a_2) = (1, 1)$ is no longer a fixed point due to the corrections of $-\delta a_i^2$.

Therefore, when $\alpha > 1$, the environment is sink-like and a critical motility exists, with (s4) matching (s2) upon substitution of s and α. When $0 < \alpha < 1$, the environment is source-like and no critical motility exists, since the corresponding trivial fixed point is removed from the dynamics by the correction terms $-\delta a_i^2$.

S3. Staircase model.

In the main text, we studied the shape of the wild-type profile in the staircase model and its impact on mutant fitness (Fig. 1d). Hereby, we study this process analytically by using the same framework as for the source-sink model (SI section S1, fourth subsection of Results). The wild-type profile N_x at state R of the adaptation process (Fig. 1b) is governed by a dynamical system (s6) on $\mathbb{R}^L$, where N_x is the total number of cells in spatial compartment x:

$$\begin{aligned}
\dot{N}_1 &= r\left(1-\frac{N_1}{K}\right)N_1 && +\nu N_2 && -(\nu+\delta)N_1, \\
\dot{N}_x &= r\left(1-\frac{N_x}{K}\right)N_x && +\nu N_{x-1}+\nu N_{x+1} && -(2\nu+\delta)N_x,\ x && =2,\dots,R \\
\dot{N}_x &= && \nu N_{x-1}+\nu N_{x+1} && -(2\nu+\delta)N_x,\ x && =R+1,\dots,L-1 \\
\dot{N}_L &= && \nu N_{L-1} && -(\nu+\delta)N_L., &&
\end{aligned} \tag{s6}$$

Similarly to the source-sink model, the non-negative quadrant $N_x \geq 0$ can only contain the trivial and the non-trivial fixed point[1]. Let us concentrate on the non-trivial fixed point N_x. While a closed form expression for the fixed point does not exist for $R > 1$, important properties of the non-trivial fixed point N_x can be derived analytically. To study the curvature of the wild-type profile, we define the convexity of the profile [2].

Definition 1. *N_x is convex, resp. concave on $[a, b]$ if for any $x \in [a, b] \cap \{1, \dots, L\}$:*

$$\frac{N_{x+1}+N_{x-1}}{2} > N_x, \ \textit{resp.}\ \frac{N_{x+1}+N_{x-1}}{2} < N_x.$$

Using this definition, important qualitative properties of the wild-type profile can be derived.

Theorem 2. *Let N_x be any fixed point of* (s6) *that lies in the strictly positive quadrant $N_x > 0$. Note that N_x coincides with the non-trivial fixed point N_x if it is stable. This wild-type profile N_x of resistance R satisfies:*

1. *N_x is a decreasing function of x.*
2. *In any compartment above the staircase, the effective wild-type birth rate $r(1 - N_x/K)$ is bigger than the death rate δ.*
3. *N_x is concave above the staircase, i.e., on $[1, R + 1]$.*

---

[1] Other fixed points are at play in the general staircase model but are irrelevant for the dynamics. The system of equations for the fixed points $\dot{N}_x = 0$ can be manipulated into a polynomial equation of degree 2^R by repeated substitution and elimination of variables. Therefore, there are at most 2^R fixed points by the Fundamental Theorem of Algebra. Indeed, when $\nu = 0$, the equations decouple and there are two choices for $\dot{N}_x = 0$ in $x \leq R$ ($N_x = 0$ or $N_x = K(1 - \delta/r)$) and a single choice of $\dot{N}_x = 0$ in $x > R$ ($N_x = 0$), leading to 2^R fixed points. Importantly, this includes the unstable trivial fixed point ($N_x = 0$), the stable non-trivial fixed point ($N_x = K(1 - \delta/r)$ in $x \leq R$, $N_x = 0$ in $x > R$), and other $2^R - 2$ unstable irrelevant fixed points. As ν increases, the non-negative ($N_x \geq 0$) and non-positive ($N_x \leq 0$) quadrants become forward invariant with the trivial fixed being the only allowed fixed point on the boundary of these quadrants. The other quadrants are called irrelevant. Indeed, the irrelevant fixed points are unstable in directions perpendicular to the boundary of the positive quadrant and are therefore pushed into irrelevant quadrants as ν increases from 0. Moreover, the irrelevant fixed points move on smooth trajectories as ν increases and cannot smoothly enter into either the positive or negative quadrants, as this would require a bifurcation with the trivial fixed point and discontinuous $\mathrm{d}N_x/\mathrm{d}\nu$. Numerical simulations suggest that the irrelevant fixed points bifurcate to annihilate each other, and never bifurcate to appear again in some of the quadrants. One can define new coordinates: $N = (N_1 + \cdots + N_L)/L$ along the diagonal $N_1 = \cdots = N_L$ and other coordinates for directions perpendicular to this diagonal. As $\nu \to \infty$, the dynamics in directions perpendicular to the diagonal quickly pushes the system close to the diagonal. On the diagonal a slower dynamics occurs, governed by $\dot{N} = rR(1 - N/K)N - \delta LN$. Therefore, only the trivial and non-trivial fixed points are preserved as $\nu \to \infty$, while the other fixed points must annihilate through bifurcations with each other. This explanation justifies why we have restricted our attention to the trivial and non-trivial fixed points only, which are the only fixed points to affect the dynamics in the physically relevant non-negative quadrant.

[2] This definition coincides with the definition of convexity for functions $\mathbb{R} \to \mathbb{R}$ when the points on the graph, $(1, N_1), \dots, (L, N_L)$, are joined by line segments.

4. *N_x is convex below the staircase, i.e., on $[R, L]$.*

5. *The curvature of the wild-type population profile at position x is proportional to wild-type fitness $r(1 - N_x/K)N_x\mathbb{1}_{x\leq R} - \delta N_x$ at this position and inversely proportional to motility.*

6. *The mutant fitness $r(1 - N_x/K)N_x\mathbb{1}_{x\leq R+1} - \delta N_x$ is a positive, convex and increasing function on $[1, R + 1]$, and a constant $-\delta$ function on $[R + 2, L]$. In particular, mutant fitness is maximised in the overlap region and matches the wild-type fitness characterised above in compartments $x \leq R$.*

7. *In any compartment $N_x < K(1 - \delta/r)$.*

8. *$r > \delta$ is a necessary condition for the existence of this non-trivial fixed point in the strictly positive quadrant.*

Theorem 2 explains the generic shape of a wild-type profile for any motility ν (Fig. 1d) and the shape of the mutant fitness landscape. However, it does not explain the difference between the low and high motility regimes discussed in the main text. In order to understand the behaviour of the wild-type profile at low and high motility, asymptotic expansion for the non-trivial fixed point are needed.

Theorem 3. *In the low motility regime $\nu \ll \delta$, the non-trivial fixed point is given by:*

$$N_x = \begin{cases} K\left(1 - \frac{\delta}{r}\right) + O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x = 1, \dots, R-1 \\ K\left(1 - \frac{\delta+\nu}{r}\right) + O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x = R \\ K\frac{\nu}{\delta}\left(1 - \frac{\delta}{r}\right) + O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x = R+1 \\ 0 + O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x = R+2, \dots, L \end{cases} \tag{s7}$$

In the high motility regime $\nu \gg \delta$, the non-trivial fixed point is given by:

$$N_x = K\left(1 - \frac{\delta L}{rR}\right) + b_x\frac{\delta}{\nu} + O\left(\left(\frac{\delta}{\nu}\right)^2\right), \tag{s8}$$

where the coefficients b_x satisfy

$$\begin{aligned} b_1 &= \frac{\delta K}{6rR}\left(3L^3 - 3L^2R + 3LR^2 + (R-L)(R^2-1)\left(\frac{2L}{R} - \frac{r}{\delta}\right)\right), \\ b_x &= b_1 + K\left(1 - \frac{\delta L}{rR}\right)\frac{x(x-1)}{2}\left(1 - \frac{L}{R}\right), \text{ if } x = 2, \dots, R+1, \\ b_x &= b_1 + K\left(1 - \frac{\delta L}{rR}\right)\left(\frac{x(x-1) + L(1+R-2x)}{2}\right), \text{ if } x = R, \dots, L. \end{aligned} \tag{s9}$$

Using Theorem 3, we can see that the shape of the population front is always the same in the low motility regime and overlap region $x = R+1$ provides a refuge for mutants (Fig. 1d). The high motility regime leads to a homogeneous wild-type population N_x, which increases with resistance state R and competes with mutants for space (Fig. 1d). Moreover, this fixed point can be in the negative quadrant to first order if the environment is sink-like. In this case, the non-trivial fixed point leaves the positive quadrant $N_x > 0$ at a critical motility ν_c and exchanges its stability with the trivial fixed point via a transcritical bifurcation.

Corollary 1. *Let $\overline{f} = (rR - \delta L)/L$ be the average wild-type fitness. Then, the stable wild-type profile is non-vanishing iff*

- *the environment is source-like $\overline{f} > 0$ and motility ν is arbitrary, or*
- *the environment is sink-like $\overline{f} < 0$ and motility is constrained to $\nu \in [0, \nu_c]$, where to first order*

$$\nu_c = \frac{\delta b_L}{K\left(\frac{\delta L}{rR} - 1\right)}. \tag{s10}$$

In the main text, we find a deadly motility regime characterised by extinction and no adaptation when $\nu > \nu_c$. We can prove that the adaptation rate is vanishing by deriving its upper bound.

Theorem 4. *Fix a resistance state R and define $a_{evo} = 1/\mathbb{E}T_{evo}$. Then,*

$$a_R < a_{evo} < \mu_f \sum_x N_x,$$

In particular, when critical motility ν_c exists, the adaptation rate a_R decreases to 0 faster than $\mu_f \sum_x N_x$ as $\nu \to \nu_c^-$. Moreover, no adaptation can occur if the wild-type population is extinct.

S4. Modified staircase model.

In the main text, we noted that the staircase model can be modified to consider the effect of fitness costs of resistance (Fig. S1), bactericidal antibiotics (Fig. S2), bacterial chemotaxis (Fig. S3) and horizontal gene transfer (Fig. S6). In this section, we explain each of these modifications of the original staircase model in turn and present the results for the modified models, with the key conclusion that motility governs the bacterial adaptation by different mechanisms in the low motility, high motility and deadly motility regimes as identified in the original staircase model (Fig. 2a). We finish this section with a mathematical analysis of the wild-type profile.

a. Resistance costs. The original staircase model assumes that there are no fitness costs associated to resistance. Yet, adaptations that confer antibiotic resistance often have a cost that is expressed in the absence of antibiotic pressure, and in these cases resistant genotypes often grow slower than susceptible ones [S5, S6]. We model such fitness costs of resistance by decreasing the division rate of cells above the staircase $g \geq x$ by a factor of $(1-c)^{g-x}$, where $c \in [0, 1)$ is the resistance cost [S7]. We find that sufficiently high resistance costs can decrease the adaptation rate a_R when cell motility is sufficiently low (Fig. S1). This agrees with [S8] that shows that the resistance cost c affects the adaptation rate precisely when cell motility is very low ($\nu < c^2\delta$). The rationale for this result is simple. When cell motility is low, resistant genotypes that harbour a resistance cost are less able to avoid competition with the wild-type and cannot easily move to where they have higher growth advantage. Importantly, we find that resistance costs do not affect our key conclusion: low motility accelerates adaptation, high motility decelerates adaptation and very high motility leads to extinction if the average wild-type fitness $\overline{f} = rR/L - \delta$ is negative. As a result, the deadly motility regime is not influenced by resistance costs (Fig. S1).

b. Bactericidal antibiotics. The original staircase model studies the effect of antibiotics that inhibit bacterial growth (i.e., we considered bacteriostatic antibiotics). However, many natural and clinical antibiotics kill bacteria (bactericidal antibiotics). We model bactericidal antibiotics by increasing the death rate under the staircase σ-times as in [S7]. Importantly, our results show that bacterial adaptation to bactericidal antibiotics again has the same regimes dictated by cell motility (Fig. S2). Interestingly, while bacterial adaptation is identical to the one found for bacteriostatic antibiotics in the low motility regime, a deadly motility regime can be induced at high motility if the bactericidal effect is sufficiently strong $\sigma \gg 1$. In this case, bacteria experience deadly antibiotic concentrations more often and the average wild-type fitness $\overline{f}$ is decreased by σ, as derived later.

c. Bacterial chemotaxis. The original staircase model assumes that bacteria move randomly. Yet, it is well-known that bacteria can bias their motion in chemical gradients (chemotaxis). In particular, it has been shown that antibiotics can trigger both positive [S9] and negative [S10] chemotaxis. We model the effect of biased motility by changing the probability p of bacterial movement up or down the antibiotic gradient. Notably, the usual adaptation regimes can be identified in this model (Fig. S3). Our results show that chemotaxis has little impact on the adaptation rate when cell motility is low (Fig. S3), mimicking what we found for bactericidal antibiotics described above. However, when motility is high, positive chemotaxis $p > 0.5$ increases antibiotic exposure and promotes the deadly motility regime (SI Movie S3), while negative chemotaxis $p < 0.5$ decreases antibiotic exposure and prevents the deadly motility regime (SI Movie S4). Put differently, chemotaxis changes the average wild-type fitness $\overline{f}$, decreasing it for positive chemotaxis and increasing it for negative chemotaxis, as derived later. For negative chemotaxis, the wild-type has higher average fitness but resides far from higher antibiotic concentrations. Therefore, mutants are unlikely to reach the overlap region and the adaptation rate is decreased (Fig. S3).

d. Horizontal gene transfer. Bacteria can exchange resistance genes through horizontal gene transfer (HGT), which is known to accelerate bacterial adaptation [S11]. In particular, it has been shown that the immigration of susceptible cells from a detached habitat can promote adaptation in a focal habitat via HGT [S12]. However, for closed systems without immigration of susceptible genotypes, it has been argued that HGT does not affect bacterial adaptation substantially [S13–S17]. In an attempt to recapitulate these results, we model HGT in our system by allowing less resistant genotypes g to acquire resistance from $N_{x,g'}$ more resistant cells of genotype $g' > g$ located at the same position x at a rate $h \times N_{x,g'}$. For realistic values of HGT rate h (< 1 transfer per generation [S12]), we find that the adaptation rate is slightly increased, but our key conclusion still holds: low motility accelerates bacterial adaptation, while high motility decelerates it. Moreover, we find that bacteria with very high motility experience deadly motility regimes if they have negative average fitness $\overline{f} = rR/L - \delta$, irrespective of HGT rate (Fig. S6).

e. Wild-type profile. So far, we considered various modifications of the staircase model. We now aim to understand how the conditions for the deadly motility regime change in the modified models. We see that there is no change to the deadly motility regime in the case of resistance costs and HGT (Fig. S1, S6). Therefore, we focus on the bactericidal antibiotics and bacterial chemotaxis that impact the deadly motility regime by modulating the average wild-type fitness $\overline{f}$ (Fig. S1, S6). We find the formulae for the average wild-type fitness $\overline{f}$ in these two modified models by studying the wild-type profiles in general modified staircase models with spatially variable division rate r_x, death rate δ_x and motility rate ν_x. In this setting, chemotaxis corresponds to the choice $\nu_{R,x} = 2p\nu$ and $\nu_{L,x} = 2(1-p)\nu$, where p is the probability to move up the gradient; the effect of bactericidal antibiotics corresponds to stress-induced death rate $\delta_x = \sigma\delta$ below the staircase ($x > R$) and the usual death rate $\delta_x = \delta$ above the staircase ($x \leq R$). The wild-type profile formation at state R of the adaptation process (Fig. 1b) is governed by the mean-field equations:

$$\begin{aligned}
\dot{N}_1 &= r_1\left(1-\frac{N_1}{K}\right)N_1 && +\nu_{L,1}N_2 && -(\nu_{R,1}+\delta_x)N_1, \\
\dot{N}_x &= r_x\left(1-\frac{N_x}{K}\right)N_x && +\nu_{R,x-1}N_{x-1}+\nu_{L,x}N_{x+1} && -(\nu_{R,x}+\nu_{L,x-1}+\delta_x)N_x,\ x = 2,\dots,L-1 \\
\dot{N}_L &= r_L\left(1-\frac{N_L}{K}\right)N_L && +\nu_{R,L-1}N_{L-1} && -(\nu_{L,L-1}+\delta_L)N_L.
\end{aligned} \tag{s11}$$

This dynamical system has two fixed points of interest, the trivial fixed point $N_x = 0$ and the non-trivial fixed point, which correspond to possible wild-type populations. Next, we compute the non-trivial fixed point at low and high motility and derive results for the existence of a critical motility. To do this, we assume that $\nu_{L,x}, \nu_{R,x} \sim O(\nu)$, $\delta_x \sim O(\delta)$, which is true for the modifications mentioned above.

Theorem 5. *In the low motility regime $\nu \ll \delta$, the wild-type profile is given by:*

$$N_x = \begin{cases} K\left(1-\frac{\delta_x}{r_x}\right) + K\left(\frac{r_{x+1}-\delta_{x+1}}{r_x-\delta_x}\frac{\nu_{L,x}}{r_{x+1}} - \frac{\nu_{R,x}}{r_x}\right) + O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x = 1 \\ K\left(1-\frac{\delta_x}{r_x}\right) + K\left(\frac{r_{x+1}-\delta_{x+1}}{r_x-\delta_x}\frac{\nu_{L,x}}{r_{x+1}} + \frac{r_{x-1}-\delta_{x-1}}{r_x-\delta_x}\frac{\nu_{R,x-1}}{r_{x-1}} - \frac{\nu_{R,x}+\nu_{L,x-1}}{r_x}\right) + O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x = 2,\dots,L-1 \\ K\left(1-\frac{\delta_x}{r_x}\right) + K\left(\frac{r_{x-1}-\delta_{x-1}}{r_x-\delta_x}\frac{\nu_{R,x-1}}{r_{x-1}} - \frac{\nu_{L,x-1}}{r_x}\right) + O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x = L \end{cases} \tag{s12}$$

In the high motility regime $\nu \gg \delta$, the wild-type profile is given by:

$$N_x = \left[\prod_{y=1}^{x-1}\frac{\nu_{R,y}}{\nu_{L,y}}\right] K \left[\frac{\sum_{z=1}^{L}(r_z-\delta_z)\prod_{y=1}^{z-1}\frac{\nu_{R,y}}{\nu_{L,y}}}{\sum_{z=1}^{L} r_z \prod_{y=1}^{z-1}\frac{\nu_{R,y}}{\nu_{L,y}}}\right] + O\left(\frac{\delta}{\nu}\right). \tag{s13}$$

The average wild-type fitness therefore is $\overline{f} = \sum_{z=1}^{L}(r_z-\delta_z)\prod_{y=1}^{z-1}\frac{\nu_{R,y}}{\nu_{L,y}} / \sum_{z=1}^{L}\prod_{y=1}^{z-1}\frac{\nu_{R,y}}{\nu_{L,y}}$. The stable non-vanishing wild-type profile exists iff

1. *the environment is source-like $\overline{f} > 0$ and motility ν is arbitrary, or*
2. *the environment is sink-like $\overline{f} < 0$ and motility is constrained to $\nu \in [0, \nu_c]$ for a critical motility ν_c.*

We note that this result coincides with the one from the basic staircase model (SI Theorem 3, SI Corollary 1) if $r_x = r$ above the staircase and 0 otherwise, $\delta_x = \delta$ and $\nu_{R,x} = \nu_{L,x} = \nu$. Furthermore, this result explains the forms of the average wild-type fitness $\overline{f}$ that were presented above for bactericidal antibiotics and chemotaxis.

For bactericidal antibiotics, the antibiotics-induced death increases σ-times, which decreases the average wild-type fitness $\overline{f}$ below zero if σ is sufficiently high, as

$$\overline{f} = (r-\delta)\frac{R}{L} - \sigma\delta\frac{L-R}{L}. \tag{s14}$$

Therefore, Theorem 5 explains why the bactericidal effect σ converts the high motility regime into the deadly motility regime, while leaving adaptation in the low motility regime almost unchanged (Fig. S2). Intuitively, at low motility regime the population is not present in stressful regions below the staircase, making the adaptation invariant of antibiotics-induced death. In the high motility regime, the population experiences the average environment with increased overall death rate, which decreases the average wild-type fitness.

For chemotaxis, probability p that a cell moves up the antibiotic gradient decreases (resp. increases) the average wild-type fitness $\overline{f}$ below (resp. above) zero when p is sufficiently high (resp. low), as

$$\overline{f} = r(1-p)^{L-R}\frac{p^R + p^{R-1}(1-p) + \dots + p(1-p)^{R-1} + (1-p)^R}{p^L + p^{L-1}(1-p) + \dots + p(1-p)^{L-1} + (1-p)^L} - \delta. \tag{s15}$$

This result means that positive chemotaxis can always create a deadly motility regime (irrespective of R), and negative chemotaxis can always prevent a deadly motility regime (irrespective of R), as we confirm with simulations (Fig. S3). Theorem 5 also shows that, at high motility, the wild-type profile of a population with negative chemotaxis becomes concentrated away from the overlap region. This result explains the reduction in the adaptation rate in Fig. S3 as a consequence of reduced antibiotic exposure. Moreover, Theorem 5 also predicts that chemotaxis does not affect the wild-type at low motility, explaining why the adaptation rate is not affected at low motility in Fig. S3. We note that it is easier to introduce the deadly motility regime by controlling chemotaxis (power law dependence on p) rather than controlling bactericidal effects of antibiotics (linear dependence on σ).

S5. Stochastic switching of motility phenotypes.

In the main text, we studied the wild-type profiles of phenotypically heterogeneous populations with bacteria switching between different motility rates stochastically. Hereby, we study this process analytically. For convenience, we modify the notation slightly. The phenotypes have cell numbers N_x and M_x; motilities ν_N and ν_M; and phenotypes switch stochastically at rate s. Without loss of generality, we assume that $\nu_M \geq \nu_N$. At ecological timescales, the wild-type profile is governed by a dynamical system (s16):

$$\begin{aligned}
\dot{N}_1 &= r\left(1-\frac{N_1+M_1}{K}\right)N_1 && +\nu_N N_2 && +sM_1-(\nu_N+\delta+s)N_1,\\
\dot{N}_x &= r\left(1-\frac{N_x+M_x}{K}\right)N_x && +\nu_N N_{x-1}+\nu_N N_{x+1} && +sM_x-(2\nu_N+\delta+s)N_x,\ x=2,\dots,R\\
\dot{N}_x &= && +\nu_N N_{x-1}+\nu_N N_{x+1} && +sM_x-(2\nu_N+\delta+s)N_x,\ x=R+1,\dots,L-1\\
\dot{N}_L &= && +\nu_N N_{L-1} && +sM_L-(\nu_N+\delta+s)N_L,\\
\dot{M}_1 &= r\left(1-\frac{N_1+M_1}{K}\right)M_1 && +\nu_M M_2 && +sN_1-(\nu_M+\delta+s)M_1,\\
\dot{M}_x &= r\left(1-\frac{N_x+M_x}{K}\right)M_x && +\nu_M M_{x-1}+\nu_M M_{x+1} && +sN_x-(2\nu_M+\delta+s)M_x,\ x=2,\dots,R\\
\dot{M}_x &= && +\nu_M M_{x-1}+\nu_M M_{x+1} && +sN_x-(2\nu_M+\delta+s)M_x,\ x=R+1,\dots,L-1\\
\dot{M}_L &= && +\nu_M M_{L-1} && +sN_L-(\nu_M+\delta+s)M_L,
\end{aligned} \tag{s16}$$

As before, the wild-type profile approaches the unique stable fixed point in the non-negative quadrant, which is either the trivial or the non-trivial fixed point. We next discuss the dynamics in the source-sink model ($L=2, R=1$) and state the results for the staircase model (any L, R).

In the source-sink model, there is a closed form expression for the wild-type profile.

Theorem 6. *The dynamical system* (s16) *for the source-sink model ($L=2, R=1$) has two fixed points:*

- *the trivial fixed point:* $N_{1,2}=M_{1,2}=0$,

- *the non-trivial fixed point:*

$$\begin{aligned}
N_1 &= \frac{K}{1+B(s)}\left(1-\frac{c_N(s)-d(s)B(s)}{r}\right),\\
M_1 &= \frac{K}{1+B(s)^{-1}}\left(1-\frac{c_M(s)-d(s)B(s)^{-1}}{r}\right),\\
N_2 &= a_N(s)M_1+b_N(s)N_1,\\
M_2 &= a_M(s)N_1+b_M(s)M_2,
\end{aligned} \tag{s17}$$

with coefficients

$$
\begin{aligned}
a_N(s) &= \frac{s\nu_M}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)},\\
b_N(s) &= \frac{(\delta+s+\nu_M)\nu_N}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)},\\
c_N(s) &= s+\delta+\nu_N\frac{\delta(\delta+\nu_M)+s(2\delta+\nu_M)}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)},\\
d(s) &= s\left(1+\frac{\nu_M\nu_N}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)}\right),\\
A(s) &= \frac{(\nu_M-\nu_N)\delta(\delta+2s)}{s[(\nu_N\nu_M+(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)]}\geq 0,\\
B(s) &= \frac{-A+\sqrt{A^2+4}}{2}\leq 1,
\end{aligned}
\tag{s18}
$$

and other follow from $N\leftrightarrow M$ symmetry. Importantly, $M_1=B(s)N_1$.

In the main text, we claimed that the diagonal of Fig. 3b corresponds to the population of uniform motility $\nu_N=\nu_M\equiv\nu$. A dynamical system for $(N+M)_{1,2}$ behaves as the basic source-sink model of Theorem 1, leading to a non-trivial fixed point

$$
\begin{aligned}
N_1=M_1&=\frac{K}{2}\left(1-\frac{\delta(\delta+2\nu)}{r(\delta+\nu)}\right),\\
N_2=M_2&=\frac{\nu}{\nu+\delta}(N_1+N_2),
\end{aligned}
$$

which is unaffected by switching rate s. In the main text, we also showed that phenotypic switching can modulate the adaptation regime. This effect is controlled by the dimensionless number $A(s)$. As $A(s)$ contains only the δ, s rates (besides motilities), the low switching ($A\gg 1$, $s\ll\delta$) and high switching ($A\ll 1$, $s\gg\delta$) regimes can be distinguished and described.

Corollary 2. *At low switching ($A\gg 1$, $s\ll\delta$), the non-trivial fixed point of the source-sink model ($L=2, R=1$) is:*

$$
\begin{aligned}
N_1w &= K\left(1-\frac{\delta(\delta+2\nu_N)}{r(\delta+\nu_N)}\right)-\frac{s}{\delta}\frac{\nu_N\nu_M+(\delta+\nu_N)(\delta+\nu_M)}{\delta(\nu_M-\nu_N)}K\left(1-\frac{\delta(\delta+2\nu_N)}{r(\delta+\nu_N)}\right)+\dots,\\
M_1 &= \frac{s}{\delta}K\left(1-\frac{\delta(\delta+2\nu_N)}{r(\delta+\nu_N)}\right)\frac{\nu_N\nu_M+(\delta+\nu_N)(\delta+\nu_M)}{\delta(\nu_M-\nu_N)}+\dots,\\
N_2 &= \frac{\nu_N}{\delta+\nu_N}K\left(1-\frac{\delta(\delta+2\nu_N)}{r(\delta+\nu_N)}\right)\left(1-\frac{s}{\delta}\frac{(\delta+\nu_N)(\nu_N\nu_M+(\delta+\nu_N)(\delta+\nu_M))+\delta^2(\nu_M-\nu_N)}{\delta(\nu_M-\nu_N)(\delta+\nu_N)}+\dots\right),\\
M_2 &= \frac{s}{\delta}\frac{\nu_M}{\delta+\nu_M}K\left(1-\frac{\delta(\delta+2\nu_N)}{r(\delta+\nu_N)}\right)\left(\frac{\nu_N\nu_M+(\delta+\nu_N)(\delta+\nu_M)}{\delta(\nu_M-\nu_N)}+\frac{\delta}{\delta+\nu_N}\right)+\dots,
\end{aligned}
\tag{s19}
$$

with the dimensionless coefficients

$$
\begin{aligned}
A(s) &= \frac{\delta}{s}\frac{\delta(\nu_M-\nu_N)}{\nu_N\nu_M+(\delta+\nu_N)(\delta+\nu_M)}+\dots,\\
B(s) &= \frac{s}{\delta}\frac{\nu_N\nu_M+(\delta+\nu_N)(\delta+\nu_M)}{\delta(\nu_M-\nu_N)}+\dots.
\end{aligned}
\tag{s20}
$$

In this case, the slower phenotype dominates the faster phenotype ($N_{1,2}\gg M_{1,2}$) and the system behaves as if only the slower phenotype was present ($(N+M)_{1,2}$ matches Theorem 1 with $\nu=\nu_N$).

At high switching ($A \ll 1$, $s \gg \delta$), the non-trivial fixed point of the source-sink model ($L = 2, R = 1$) is:

$$\begin{aligned} N_1 &= \frac{K}{2}\left(1 - \frac{2\delta(\delta + \nu_M + \nu_N)}{r(2\delta + \nu_N + \nu_M)}\right) + \dots, \\ M_1 &= \frac{K}{2}\left(1 - \frac{2\delta(\delta + \nu_M + \nu_N)}{r(2\delta + \nu_N + \nu_M)}\right) + \dots, \\ N_2 &= \frac{\nu_N + \nu_M}{2\delta + \nu_N + \nu_M}\frac{K}{2}\left(1 - \frac{2\delta(\delta + \nu_M + \nu_N)}{r(2\delta + \nu_N + \nu_M)}\right) + \dots, \\ M_2 &= \frac{\nu_N + \nu_M}{2\delta + \nu_N + \nu_M}\frac{K}{2}\left(1 - \frac{2\delta(\delta + \nu_M + \nu_N)}{r(2\delta + \nu_N + \nu_M)}\right) + \dots, \end{aligned} \tag{s21}$$

with dimensionless coefficients

$$\begin{aligned} A(s) &= \frac{\delta}{s}\frac{2(\nu_M - \nu_N)}{2\delta + \nu_N + \nu_M} + \dots, \\ B(s) &= 1 - \frac{\delta}{s}\frac{\nu_M - \nu_N}{2\delta + \nu_N + \nu_M} + \dots. \end{aligned} \tag{s22}$$

In this case, both phenotypes are similarly important ($N_{1,2} \approx M_{1,2}$) and the system behaves as if only an average phenotype was present ($(N + M)_{1,2}$ matches Theorem 1 with $\nu = (\nu_M + \nu_N)/2$).

This result proves the following claim we made in the main text that at low switching ($A \ll 1$, $s \gg \delta$), the population adapts as if only the slower phenotype was present, and that at high switching ($A \ll 1$, $s \gg \delta$), the population adapts as if it had a uniform motility that matches the average motility.

In the main text, we also discussed the deadly motility regime. In contrast to systems with uniform motility, the deadly motility regime now occupies the (ν_N, ν_M, s) space; and instead of critical motility, there is a critical surface in the (ν_N, ν_M, s) space.

Corollary 3. *The deadly motility regime for the source-sink model ($L = 2, R = 1$) is described as a subspace of (ν_N, ν_M, s) space such that:*

$$D(s) > r - \delta, \tag{s23}$$

with

$$D(s) = s + \nu_N(1 - b_N(s)) - d(s)B(s). \tag{s24}$$

$D(s)$ has the following properties:

- *$D(s)$ is a strictly increasing smooth function of $s \in \mathbb{R}^+$*
- *$D(0) = \frac{\delta\nu_N}{\delta+\nu_N}$*
- *$D(\infty) = \frac{2\delta}{2\delta+\nu_N+\nu_M}$ as $s \to \infty$*

The deadly motility regime exist whenever the environment is sink-like $\overline{f} = (r - 2\delta)/2 < 0$, and

- *the slower phenotype is in the deadly motility regime $\min(\nu_N, \nu_M) > \nu_c$, and switching s is arbitrary, or*
- *the average phenotype is in the deadly motility regime $(\nu_N + \nu_M)/2 > \nu_c$, and switching s is above some critical value $s > s_c(\nu_N, \nu_M)$.*

In particular, the critical surface asymptotes with $\min(\nu_N, \nu_M) = \nu_c$ at low switching ($s \ll \delta$) and with $(\nu_N + \nu_M)/2 = \nu_c$ at high switching ($s \gg \delta$).

For completeness, we state the results for the staircase model, which can be used to support the same claims as those made above for the simpler source-sink model. The asymptotic form of the wild-type profile can be derived.

Theorem 7. *In the low motility combination $\nu_N, \nu_M \ll \delta$, the non-trivial fixed point of the staircase model is given by:*

$$M_x = \begin{cases} \frac{K}{2}\left(1-\frac{\delta}{r}\right)+\dots & \text{if } x=1,\dots,R-1 \\ \frac{K}{2}\left(1-\frac{\delta+(\nu_N+\nu_M)/2}{r}\right)-\frac{K}{8s}\left(1-\frac{\delta}{r}\right)(\nu_M-\nu_N)+\dots & \text{if } x=R \\ \frac{K}{4}\left(1-\frac{\delta}{r}\right)\left(\frac{\nu_N+\nu_M}{\delta}+\frac{\nu_M-\nu_N}{\delta+2s}\right)++\dots & \text{if } x=R+1 \\ 0+\dots & \text{if } x=R+2,\dots,L \end{cases}$$
$$N_x = \begin{cases} \frac{K}{2}\left(1-\frac{\delta}{r}\right)+\dots & \text{if } x=1,\dots,R-1 \\ \frac{K}{2}\left(1-\frac{\delta+(\nu_N+\nu_M)/2}{r}\right)-\frac{K}{8s}\left(1-\frac{\delta}{r}\right)(\nu_N-\nu_M)+\dots & \text{if } x=R \\ \frac{K}{4}\left(1-\frac{\delta}{r}\right)\left(\frac{\nu_N+\nu_M}{\delta}+\frac{\nu_N-\nu_M}{\delta+2s}\right)++\dots & \text{if } x=R+1 \\ 0+\dots & \text{if } x=R+2,\dots,L \end{cases} \tag{s25}$$

In the high motility combination $\nu_N, \nu_M \gg \delta$, the non-trivial fixed point of the staircase model is given by:

$$N_x = M_x = \frac{K}{2}\left(1-\frac{L\delta}{Rr}\right)+\dots. \tag{s26}$$

In the mixed motility combination $\nu_N \ll \delta \ll \nu_M$, the non-trivial fixed point of the staircase model is given by

$$M_x = K\left(1-\frac{\delta+s(1-B(s))}{r}\right)\frac{B(s)}{1+B(s)}+\dots$$
$$N_x = \begin{cases} K\left(1-\frac{\delta+s(1-B(s))}{r}\right)\frac{1}{1+B(s)}+\dots & \text{if } x\le R \\ K\left(1-\frac{\delta+s(1-B(s))}{r}\right)\frac{sB(s)}{(1+B(s))(s+\delta)}+\dots & \text{if } x>R \end{cases} \tag{s27}$$

with

$$A(s) = \left(\frac{L}{R}-1\right)\frac{\delta^2+2\delta s}{s^2+\delta s} > 0,$$
$$B(s) = \frac{-A+\sqrt{A^2+4}}{2} < 1. \tag{s28}$$

Importantly, to lowest order $N_x/M_x = B(s)$ in the region above the staircase $x \ge R$.

A non-perturbative result characterising the behaviour of the staircase model at low and high switching can also be derived analytically.

Theorem 8. *At low switching $s \ll \delta$, the overall wild-type profile behaves as if it has a single motility of $\nu = \min(\nu_N, \nu_M)$. At high switching $s \gg \delta$, the overall wild-type profile behaves as if it has a single motility of $\nu = (\nu_N+\nu_M)/2$.*

In particular, this result affects the deadly motility regime in the (ν_N, ν_M, s) space. Let $\overline{f} = (rR-\delta L)/L$ be the average wild-type fitness. The wild-type profile is non-trivial iff:

- *the environment is source-like $\overline{f} > 0$, motility rates ν_N, ν_M are arbitrary, and switching rate s is arbitrary,*
- *the environment is sink-like $\overline{f} < 0$ and*
  - *switching is low $s \ll \delta$ and the slower phenotype is in the deadly motility regime $\min(\nu_N, \nu_M) > \nu_c$,*
  - *switching is high $s \gg \delta$ and the average-motility phenotype is in the deadly motility regime $(\nu_N+\nu_M)/2 > \nu_c$.*

S6. Density-dependent motility.

In the main text, we studied the wild-type profiles of populations with density-dependent motility. We briefly outline important features of these wild-type profiles. The wild-type profile can be described with the mean-field

equations:

$$\begin{aligned}
\dot{N}_1 &= r\left(1-\frac{N_1}{K}\right)N_1 && +\nu(N_2)N_2 && -(\nu(N_1)+\delta)N_1, \\
\dot{N}_x &= r\left(1-\frac{N_x}{K}\right)N_x && +\nu(N_{x-1})N_{x-1}+\nu(N_{x+1})N_{x+1} && -(2\nu(N_x)+\delta)N_x,\ x && =2,\dots,R \\
\dot{N}_x &= && \nu(N_{x-1})N_{x-1}+\nu(N_{x+1})N_{x+1} && -(2\nu(N_x)+\delta)N_x,\ x && =R+1,\dots,L-1 \\
\dot{N}_L &= && \nu(N_{L-1})N_{L-1} && -(\nu(N_x)+\delta)N_L., &&
\end{aligned} \tag{s29}$$

where

$$\nu(N) = \begin{cases} \nu_L & \text{if } N < S, \\ \nu_H & \text{if } N \geq S. \end{cases} \tag{s30}$$

Since the forcing of this dynamical system is discontinuous whenever $N_x = S$, the fixed-point analysis is more complicated and we only mention two important features of this system.

First, multiple stable fixed points can exist. For example, consider a situation when a low-density motility phenotype has motility above the critical motility $\nu_L > \nu_c$ and the high-density motility phenotype is slower $\nu_H < \nu_L$. If the system is initially started from a low density, the low-density motility phenotype cannot survive and a deadly motility regime occurs as in Fig. 4c (attraction to the trivial fixed-point). However, if the system is initially started from a high density, the high-density motility phenotype can survive and no deadly motility regime would have occured in Fig. 4c (attraction to a non-trivial fixed-point).

Second, the surfaces of discontinuous forcing ($N_x = S$ for some x) might attract the trajectories. This result often happens in the slow-to-fast switching case, see Fig. 4b at low threshold S. In this case, the bulk of the high-density population is driven very close to density $N_x = S$.

S7. Proofs.

Theorem 1. *System* (1) *admits exactly two fixed points*

- *trivial fixed point:* $N_2 = N_1 = 0$,
- *non-trivial fixed point:*

$$\begin{aligned}
N_1 &= K\left(1-\frac{\delta(\delta+2\nu)}{r(\delta+\nu)}\right), \\
N_2 &= \frac{\nu}{\nu+\delta}N_1.
\end{aligned} \tag{s1}$$

Exactly one of these fixed points is stable and corresponds to a stable wild-type profile, with non-negative cell numbers $N_{1,2} \geq 0$. *Let* $\overline{f} = (r-2\delta)/2$ *be the average wild-type fitness. The stable wild-type profile corresponds to the non-trivial fixed point iff*

- *the environment is source-like* $\overline{f} > 0$ *and cell motility* ν *is arbitrary, or*
- *the environment is sink-like* $\overline{f} < 0$ *and cell motility is constrained to* $\nu \in [0, \nu_c]$, *where the critical motility is*

$$\nu_c = \frac{\delta(r-\delta)}{2\delta - r} > 0. \tag{s2}$$

Proof. The two fixed points can be found by setting $\dot{N}_1 = \dot{N}_2 = 0$. The stability of these fixed points can be checked by considering the signs of trace and determinant of the Jacobian matrix:

$$J = \begin{pmatrix} r - 2rN_1/K - \delta - \nu & \nu \\ \nu & -\delta - \nu \end{pmatrix}. \tag{s31}$$

For example, the trivial fixed point has tr $J = r - 2\delta - 2\nu$ and det $J = -\delta(r-\delta) + \nu(2\delta - r)$. If $\overline{f} > 0$, or if $\overline{f} < 0$ and $\nu < \nu_c$, then det $J < 0$ and the trivial fixed point is a saddle node. Otherwise, det $J > 0$, tr $J < 0$ and the trivial fixed point is stable. Similar analysis applies to the non-trivial fixed point. □

Theorem 2. *Let N_x be any fixed point of* (s6) *that lies in the strictly positive quadrant $N_x > 0$. Note that N_x coincides with the non-trivial fixed point N_x if it is stable. This wild-type profile N_x of resistance R satisfies:*

1. *N_x is a decreasing function of x.*
2. *In any compartment above the staircase, the effective wild-type birth rate $r(1 - N_x/K)$ is bigger than the death rate δ.*
3. *N_x is concave above the staircase, i.e., on $[1, R+1]$.*
4. *N_x is convex below the staircase, i.e., on $[R, L]$.*
5. *The curvature of the wild-type population profile at position x is proportional to wild-type fitness $r(1 - N_x/K)N_x\mathbb{1}_{x\leq R} - \delta N_x$ at this position and inversely proportional to motility.*
6. *The mutant fitness $r(1 - N_x/K)N_x\mathbb{1}_{x\leq R+1} - \delta N_x$ is a positive, convex and increasing function on $[1, R+1]$, and a constant $-\delta$ function on $[R+2, L]$. In particular, mutant fitness is maximised in the overlap region and matches the wild-type fitness characterised above in compartments $x \leq R$.*
7. *In any compartment $N_x < K(1 - \delta/r)$.*
8. *$r > \delta$ is a necessary condition for the existence of this non-trivial fixed point in the strictly positive quadrant.*

Proof. Part 1 and 2: Rewrite the fixed point equation for (s6) as

$$\begin{aligned}
N_1 - N_2 &= \frac{N_1}{\nu}\left[r\left(1 - \frac{N_1}{K}\right) - \delta\right] \\
N_x - N_{x+1} &= \frac{N_x}{\nu}\left[r\left(1 - \frac{N_x}{K}\right) - \delta\right] + N_{x-1} - N_x, \quad x = 2, \ldots, R \\
N_x - N_{x+1} &= -\frac{\delta N_x}{\nu} + N_{x-1} - N_x, \quad x = R+1, \ldots, L-1 \\
N_{L-1} - N_L &= \frac{\delta N_L}{\nu}.
\end{aligned}$$

Firstly, assume that $N_1 \leq N_2$. The first equation implies that $r(1-N_1/K)-\delta \leq 0$. As $N_1 \leq N_2$, also $r(1-N_2/K)-\delta \leq r(1 - N_1/K) - \delta \leq 0$. Then, the second equation implies that $N_2 \leq N_3$. Inductively, the first $L-1$ equations imply that

$$N_1 \leq \cdots \leq N_x \leq N_{x+1} \leq \cdots \leq N_L.$$

However, the last equations then implies that $N_L \leq 0$, contradicting our restriction to $N_x > 0$. Therefore, $N_1 > N_2$. The first equation implies that $r(1 - N_1/K) - \delta > 0$. As $N_1 > N_2$, $r(1 - N_2/K) - \delta > r(1 - N_1/K) - \delta > 0$. The second equation implies that $N_2 - N_3 > 0$. Inductively, the first $L-1$ equations imply the conclusion of part 1 that

$$N_1 > \cdots > N_x > N_{x+1} > \cdots > N_L.$$

Notice that along the way, part 2 has also been proven.

Part 3: Notice that above the staircase,

$$\frac{N_{x+1} + N_{x-1}}{2} - N_x = -\frac{N_x}{2\nu}\left[r\left(1 - \frac{N_x}{K}\right) - \delta\right] < 0.$$

Part 4: Notice that below the staircase,

$$\frac{N_{x+1} + N_{x-1}}{2} - N_x = \frac{N_x\delta}{2\nu} > 0.$$

Part 5: Write the fixed point equation in the form:

$$\begin{aligned}
-\nu(N_2 - N_1) &= \left[r\left(1 - \frac{N_1}{K}\right) - \delta\right] N_1 \\
\nu\left[N_{x+1} + N_{x-1} - 2N_x\right] &= \left[r\left(1 - \frac{N_x}{K}\right) - \delta\right] N_x, \quad x = 2, \ldots, R \\
\nu\left[N_{x+1} + N_{x-1} - 2N_x\right] &= -\delta N_x, \quad x = R+1, \ldots, L \\
-\nu(N_{L-1} - N_L) &= -\delta N_L.
\end{aligned}$$

Part 6: The mutant fitness is given by the difference between the effective birth-rate (resulting from competition with wild-type) and death rate. Therefore, on $[1, R+1]$, this mutant fitness is

$$r\left(1-\frac{N_x}{K}\right)-\delta.$$

Therefore, it is increasing (by part 1), positive (by part 2) and convex (by part 3). On $[R+2, L]$, mutants cannot divide and their fitness is $-\delta$.

Part 7: Follows directly from part 2.

Part 8: If $r \leq \delta$, part 6 implies that $N_x \leq 0$. □

Theorem 3. *In the low motility regime $\nu \ll \delta$, the non-trivial fixed point is given by:*

$$N_x = \begin{cases} K\left(1-\frac{\delta}{r}\right)+O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x=1,\dots,R-1 \\ K\left(1-\frac{\delta+\nu}{r}\right)+O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x=R \\ K\frac{\nu}{\delta}\left(1-\frac{\delta}{r}\right)+O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x=R+1 \\ 0+O\left(\left(\frac{\nu}{\delta}\right)^2\right) & \text{if } x=R+2,\dots,L \end{cases} \tag{s7}$$

In the high motility regime $\nu \gg \delta$, the non-trivial fixed point is given by:

$$N_x = K\left(1-\frac{\delta L}{rR}\right)+b_x\frac{\delta}{\nu}+O\left(\left(\frac{\delta}{\nu}\right)^2\right), \tag{s8}$$

where the coefficients b_x satisfy

$$\begin{aligned} b_1 &= \frac{\delta K}{6rR}\left(3L^3-3L^2R+3LR^2+(R-L)(R^2-1)\left(\frac{2L}{R}-\frac{r}{\delta}\right)\right), \\ b_x &= b_1+K\left(1-\frac{\delta L}{rR}\right)\frac{x(x-1)}{2}\left(1-\frac{L}{R}\right), \text{ if } x=2,\dots,R+1, \\ b_x &= b_1+K\left(1-\frac{\delta L}{rR}\right)\left(\frac{x(x-1)+L(1+R-2x)}{2}\right), \text{ if } x=R,\dots,L. \end{aligned} \tag{s9}$$

Proof. For the low motility regime, assume that $N_x = a_x + \frac{\nu}{\delta}b_x + \dots$ Then, at $O(1)$ of (s6), we recover the equilibrium of local birth/death dynamics:

$$\begin{aligned} r\left(1-\frac{a_x}{K}\right)a_x &= \delta a_x, \quad x=1,2,\dots,R \\ 0 &= \delta a_x, \quad x=R+1,\dots,L \end{aligned}$$

implying that $a_x = K(1-\delta/r)$ when $x \leq R$, and $a_x = 0$ otherwise. At $O(\nu/\delta)$, we recover the first order corrections to balance this local dynamics with curvature of the population curve:

$$\begin{aligned} \frac{r}{\delta}\left(1-\frac{2a_1}{K}\right)b_1+a_2 \qquad\qquad &= b_1+a_1 \\ \frac{r}{\delta}\left(1-\frac{2a_x}{K}\right)b_x+a_{x+1} \quad +a_{x-1} &= b_x+2a_x, \quad x=2,\dots,R \\ +a_{x+1} \quad +a_{x-1} &= b_x+2a_x, \quad x=R+1,\dots,L-1 \\ +a_{L-1} &= b_L+a_L. \end{aligned}$$

Using the previous formulae for a_x this can be solved for b_x. The result follows.

For the high motility regime, assume that $N_x = a_x + \frac{\delta}{\nu}b_x + \left(\frac{\delta}{\nu}\right)^2 c_x + \dots$ Then, at $O(\left(\frac{\delta}{\nu}\right)^{-1})$, we only get the spatial diffusion with neglected local birth-death processes:

$$\begin{aligned} a_1 &= a_2 \\ a_{x+1}+a_{x-1} &= a_x, \quad x=2,\dots,L-1 \\ a_{L-1} &= a_L. \end{aligned}$$

This result implies that $a_x = a$ for some constant a. Also notice that this system is under-determined and constant a must be fixed from the equations at the next order. At $O(1)$, the first order correction to curvature due to local birth/death processes is:

$$\begin{aligned}
\frac{r}{\delta}\left(1-\frac{a}{K}\right)a + b_2 \qquad\qquad &= b_1 + a \\
\frac{r}{\delta}\left(1-\frac{a}{K}\right)a + b_{x+1} \quad +b_{x-1} &= 2b_x + a, \qquad x = 2,\dots,R \\
+\, b_{x+1} \quad +b_{x-1} &= 2b_x + a, \qquad x = R+1,\dots,L-1 \\
+b_{L-1} &= b_L + a.
\end{aligned}$$

Summing the equations yields

$$R\frac{r}{\delta}\left(1-\frac{a}{K}\right)a = La,$$

and thus

$$a = K\left(1-\frac{\delta L}{rR}\right).$$

With this choice of a, the system of equations for b_x takes a form of under-determined system of difference equations. By solving this system, b_x's can can be expressed in terms of b_1 as in (s9). In order to fix b_1, the equations for c_x at order $O(\frac{\delta}{\nu})$ need to be found:

$$\begin{aligned}
\frac{r}{\delta}\left(1-\frac{2a}{K}\right)b_1 + c_2 \qquad\qquad &= b_1 + c_1 \\
\frac{r}{\delta}\left(1-\frac{2a}{K}\right)b_x + c_{x+1} \quad +c_{x-1} &= b_x + 2c_x, \qquad x = 2,\dots,R \\
+\, c_{x+1} \quad +c_{x-1} &= b_x + 2c_x, \qquad x = R+1,\dots,L-1 \\
+c_{L-1} &= b_L + c_L.
\end{aligned}$$

Summation cancels the c_x coefficients and gives a single equation for b_x:

$$\frac{r}{\delta}\left(1-\frac{2a}{K}\right)(b_1 + \cdots + b_R) = b_1 + \cdots + b_L.$$

Using the relationships for b_x:

$$b_1 + \cdots + b_R = Rb_1 + \frac{a}{6}(R^2-1)(R-L),$$

and

$$b_1 + \cdots + b_L = Lb_1 + \frac{a}{6}L(-3L^2 - R^2 + 3LR),$$

the coefficient b_1 is obtained as in (s9). □

Corollary 1. *Let $\overline{f} = (rR - \delta L)/L$ be the average wild-type fitness. Then, the stable wild-type profile is non-vanishing iff*

- *the environment is source-like $\overline{f} > 0$ and motility ν is arbitrary, or*
- *the environment is sink-like $\overline{f} < 0$ and motility is constrained to $\nu \in [0, \nu_c]$, where to first order*

$$\nu_c = \frac{\delta b_L}{K\left(\frac{\delta L}{rR} - 1\right)}. \tag{s10}$$

Proof. Notice that a stable wild-type profile N_x is expected to be decreasing in x by Theorem 2. Hence, a non-trivial fixed point stays within the positive quadrant and remains stable iff $N_L > 0$. To the lowest order, this condition is equivalent to $\overline{f} > 0$. If $\overline{f} > 0$, the wild-type profile exists for any ν. If $\overline{f} < 0$, the first order approximation dictates that $N_L > 0$ iff $\nu < \nu_c$ with ν_c given by (s10). At $\nu = \nu_c$, the non-trivial wild-type fixed point collides with the trivial wild-type fixed point and exchanges stability through a transcritical bifurcation. □

Theorem 4. *Fix a resistance state R and define $a_{evo} = 1/\mathbb{E}T_{evo}$. Then,*

$$a_R < a_{evo} < \mu_f \sum_x N_x,$$

In particular, when critical motility ν_c exists, the adaptation rate a_R decreases to 0 faster than $\mu_f \sum_x N_x$ as $\nu \to \nu_c^-$. Moreover, no adaptation can occur if the wild-type population is extinct.

Proof. To find the upper bound on a_{evo}, we need to find the lower bound on the waiting time T_{evo} till the stable wild-type population produces a mutant in the overlap region (mutants migrating from other compartments count as well). This waiting time T_{evo} is bounded below by the waiting time till a mutant is produced anywhere T_a, i.e., $T_{evo} > T_a$. In each spatial compartment x, only the most resistant wild-type cells can mutate. Call their number N'_x. Then, the waiting time T_x till the compartment x produces a mutant is $\sim \mathrm{Exp}(\mu_f N'_x)$. The waiting times T_x are independent and $T_a = \min_x T_x$. Therefore, $T_a \sim \mathrm{Exp}(\mu_f \sum_x N'_x)$. Finally,

$$\mathbb{E}T_{evo} > \mathbb{E}T_a = \frac{1}{\mu_f \sum_x N'_x}.$$

Thus,

$$a_{evo} = \frac{1}{\mathbb{E}T_{evo}} < \mu_f \sum_x N'_x \leq \mu_f \sum_x N_x = \mu_f N.$$

Result follows. □

Theorem 6. *The dynamical system* (s16) *for the source-sink model ($L = 2, R = 1$) has two fixed points:*

- *the trivial fixed point: $N_{1,2} = M_{1,2} = 0$,*
- *the non-trivial fixed point:*

$$\begin{aligned}
N_1 &= \frac{K}{1+B(s)}\left(1 - \frac{c_N(s) - d(s)B(s)}{r}\right),\\
M_1 &= \frac{K}{1+B(s)^{-1}}\left(1 - \frac{c_M(s) - d(s)B(s)^{-1}}{r}\right),\\
N_2 &= a_N(s)M_1 + b_N(s)N_1,\\
M_2 &= a_M(s)N_1 + b_M(s)M_2,
\end{aligned} \tag{s17}$$

with coefficients

$$\begin{aligned}
a_N(s) &= \frac{s\nu_M}{(\delta+\nu_M)(\delta+\nu_N) + s(2\delta+\nu_N+\nu_M)},\\
b_N(s) &= \frac{(\delta+s+\nu_M)\nu_N}{(\delta+\nu_M)(\delta+\nu_N) + s(2\delta+\nu_N+\nu_M)},\\
c_N(s) &= s+\delta+\nu_N \frac{\delta(\delta+\nu_M) + s(2\delta+\nu_M)}{(\delta+\nu_M)(\delta+\nu_N) + s(2\delta+\nu_N+\nu_M)},\\
d(s) &= s\left(1 + \frac{\nu_M\nu_N}{(\delta+\nu_M)(\delta+\nu_N) + s(2\delta+\nu_N+\nu_M)}\right),\\
A(s) &= \frac{(\nu_M-\nu_N)\delta(\delta+2s)}{s[(\nu_N\nu_M + (\delta+\nu_M)(\delta+\nu_N) + s(2\delta+\nu_N+\nu_M)]} \geq 0,\\
B(s) &= \frac{-A+\sqrt{A^2+4}}{2} \leq 1,
\end{aligned} \tag{s18}$$

and other follow from $N \leftrightarrow M$ symmetry. Importantly, $M_1 = B(s)N_1$.

Proof. Notice that the system (s16) is symmetric under $M \leftrightarrow N$. Now, express M_2, N_2 in terms of N_1, M_1 using the equations $\dot{N}_2 = 0$ and $\dot{M}_2 = 0$:

$$\begin{aligned}
N_2 &= a_N(s)M_1 + b_N(s)N_1,\\
M_2 &= a_M(s)N_1 + b_M(s)M_1,
\end{aligned}$$

with the coefficients

$$a_N(s) = \frac{s\nu_M}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)},$$
$$b_N(s) = \frac{(\delta+s+\nu_M)\nu_N}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)},$$

and $a_M(s), b_M(s)$ follow by $N \leftrightarrow M$. This changes the the equations $\dot{N}_1 = 0$ and $\dot{M}_1 = 0$ into

$$0 = r(1-(N_1+M_1)/K)N_1 - c_N(s)N_1 + d(s)M_1,$$
$$0 = r(1-(N_1+M_1)/K)M_1 - c_M(s)M_1 + d(s)N_1,$$

with coefficients

$$c_N(s) = s+\delta+\nu_N(1-b_N(s)) = s+\delta+\nu_N\frac{\delta(\delta+\nu_M)+s(2\delta+\nu_M)}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)},$$
$$d(s) = s+\nu_N a_N(s) = s\left(1+\frac{\nu_M\nu_N}{(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)}\right),$$

and $c_M(s)$ follows by $N \leftrightarrow M$. Multiplying the first equation by M_1, the second equation by N_2 and subtracting, we obtain

$$(c_M(s)-c_N(s))M_1N_1 = d(s)(N_1^2-M_1^2).$$

Therefore, either $N_1 = 0$ or $N_1 \neq 0$. In the first case, we obtain the trivial fixed point by following back through the equations. In the latter case,

$$\frac{N_1}{M_1} - \frac{M_1}{N_1} = \frac{c_M(s)-c_N(s)}{d(s)} \equiv A(s),$$

with

$$A(s) = \frac{(\nu_M-\nu_N)\delta(\delta+2s)}{s[(\nu_N\nu_M+(\delta+\nu_M)(\delta+\nu_N)+s(2\delta+\nu_N+\nu_M)]} \geq 0$$

Therefore,

$$\frac{M_1}{N_1} = \frac{-A+\sqrt{A^2+4}}{2} \equiv B(s) \leq 1.$$

Finally, this result can be used to derive:

$$N_1 = \frac{K}{1+B(s)}\left(1-\frac{c_N(s)-d(s)B(s)}{r}\right).$$

We also note that the symmetry $N \leftrightarrow M$ is preserved as

$$M_1 = B(s)N_1 = \frac{K}{1+B(s)^{-1}}\left(1-\frac{c_M(s)-d(s)B(s)^{-1}}{r}\right).$$

□

Corollary 3. *The deadly motility regime for the source-sink model ($L = 2, R = 1$) is described as a subspace of (ν_N, ν_M, s) space such that:*

$$D(s) > r-\delta, \tag{s23}$$

with

$$D(s) = s+\nu_N(1-b_N(s)) - d(s)B(s). \tag{s24}$$

$D(s)$ has the following properties:

- *$D(s)$ is a strictly increasing smooth function of $s \in \mathbb{R}^+$*
- $D(0) = \frac{\delta\nu_N}{\delta+\nu_N}$
- $D(\infty) = \frac{2\delta}{2\delta+\nu_N+\nu_M}$ *as* $s \to \infty$

The deadly motility regime exist whenever the environment is sink-like $\overline{f} = (r-2\delta)/2 < 0$, and

- *the slower phenotype is in the deadly motility regime $\min(\nu_N, \nu_M) > \nu_c$, and switching s is arbitrary, or*
- *the average phenotype is in the deadly motility regime $(\nu_N + \nu_M)/2 > \nu_c$, and switching s is above some critical value $s > s_c(\nu_N, \nu_M)$.*

In particular, the critical surface asymptotes with $\min(\nu_N, \nu_M) = \nu_c$ at low switching ($s \ll \delta$) and with $(\nu_N + \nu_M)/2 = \nu_c$ at high switching ($s \gg \delta$).

Proof. From Theorem 6, the deadly motility regime is implicitly described $0 > r - c_N(s) + d(s)B(s)$. Using $c_N(s) = s + \delta + \nu_N(1 - b_N(s)$, we get the condition in terms of $D(s)$. Notice that $b_N(s)$, $d(s)$ and $B(s)$ are all smooth strictly decreasing functions of $s \in \mathbb{R}^+$, hence $D(s)$ is smooth strictly increasing. The limits follow using the formulae in the Theorem 6. Now, fix ν_N, ν_M and vary s.

First, we are in the deadly motility regime irrespective of switching s if $D(0) > r - \delta$. Notice that this condition is equivalent to $\nu_N > \nu_c = \frac{\delta(r-\delta)}{2\delta-r}$, where we recognise the critical motility ν_c in Theorem 1.

Second, we can reach the deadly motility regime for sufficiently high switching rates $s > s_c$, where s_c is the unique critical switching $s_c > 0$ given by $D(s_c) = r - \delta$. Such a critical switching rate exists precisely when $D(\infty) > r - \delta$. This happens precisely when $\frac{\nu_N+\nu_M}{2} > \nu_c = \frac{\delta(r-\delta)}{2\delta-r}$. □

Theorem 7. *In the low motility combination $\nu_N, \nu_M \ll \delta$, the non-trivial fixed point of the staircase model is given by:*

$$M_x = \begin{cases} \frac{K}{2}\left(1-\frac{\delta}{r}\right) + \dots & \text{if } x = 1, \dots, R-1 \\ \frac{K}{2}\left(1 - \frac{\delta+(\nu_N+\nu_M)/2}{r}\right) - \frac{K}{8s}\left(1-\frac{\delta}{r}\right)(\nu_M - \nu_N) + \dots & \text{if } x = R \\ \frac{K}{4}\left(1-\frac{\delta}{r}\right)\left(\frac{\nu_N+\nu_M}{\delta} + \frac{\nu_M-\nu_N}{\delta+2s}\right) + + \dots & \text{if } x = R+1 \\ 0 + \dots & \text{if } x = R+2, \dots, L \end{cases}$$
$$N_x = \begin{cases} \frac{K}{2}\left(1-\frac{\delta}{r}\right) + \dots & \text{if } x = 1, \dots, R-1 \\ \frac{K}{2}\left(1 - \frac{\delta+(\nu_N+\nu_M)/2}{r}\right) - \frac{K}{8s}\left(1-\frac{\delta}{r}\right)(\nu_N - \nu_M) + \dots & \text{if } x = R \\ \frac{K}{4}\left(1-\frac{\delta}{r}\right)\left(\frac{\nu_N+\nu_M}{\delta} + \frac{\nu_N-\nu_M}{\delta+2s}\right) + + \dots & \text{if } x = R+1 \\ 0 + \dots & \text{if } x = R+2, \dots, L \end{cases} \tag{s25}$$

In the high motility combination $\nu_N, \nu_M \gg \delta$, the non-trivial fixed point of the staircase model is given by:

$$N_x = M_x = \frac{K}{2}\left(1 - \frac{L\delta}{Rr}\right) + \dots. \tag{s26}$$

In the mixed motility combination $\nu_N \ll \delta \ll \nu_M$, the non-trivial fixed point of the staircase model is given by

$$M_x = K\left(1 - \frac{\delta + s(1-B(s))}{r}\right)\frac{B(s)}{1+B(s)} + \dots$$
$$N_x = \begin{cases} K\left(1 - \frac{\delta+s(1-B(s))}{r}\right)\frac{1}{1+B(s)} + \dots & \text{if } x \le R \\ K\left(1 - \frac{\delta+s(1-B(s))}{r}\right)\frac{sB(s)}{(1+B(s))(s+\delta)} + \dots & \text{if } x > R \end{cases} \tag{s27}$$

with

$$A(s) = \left(\frac{L}{R} - 1\right)\frac{\delta^2 + 2\delta s}{s^2 + \delta s} > 0,$$
$$B(s) = \frac{-A + \sqrt{A^2+4}}{2} < 1. \tag{s28}$$

Importantly, to lowest order $N_x/M_x = B(s)$ in the region above the staircase $x \ge R$.

Proof. **Low motility combination.** Assume that ν_N, ν_M are two independent small parameters, $N_x = a_x + \nu_N b_{Nx} + \nu_M b_{Mx} \ldots$, and $M_x = c_x + \nu_M d_{Mx} + \nu_N d_{Nx} \ldots$. As before, expand in dimensional powers of $O(\nu_N^k, \nu_M^l)$ to simplify the algebra.

At $O(1)$, the equations describe the local compartmental equilibrium:

$$\begin{aligned} r\left(1 - \frac{a_x + c_x}{K}\right) a_x + sc_x &= (\delta + s)a_x,\ x = 1, \ldots, R \\ r\left(1 - \frac{a_x + c_x}{K}\right) c_x + sa_x &= (\delta + s)c_x,\ x = 1, \ldots, R \\ + sc_x &= (\delta + s)a_x,\ x = R+1, \ldots, L \\ + sa_x &= (\delta + s)c_x,\ x = R+1, \ldots, L \end{aligned}$$

Subtracting the first two and the latter two equations, we can see that

$$\begin{aligned} (a_x - c_x)\left[r\left(1 - \frac{a_x + c_x}{K}\right) - 2s - \delta\right] = 0 \quad &, x = 1, \ldots, R \\ (a_x - c_x)(2s + \delta) = 0 \quad &, x = R+1, \ldots, L \end{aligned}$$

For $x > R$, this implies that $a_x = c_x$, and in turn that $a_x = 0 = c_x$. For $x \leq R$, one of the brackets must vanish. If the latter bracket was to vanishes, then the original equations imply that $a_x + c_x = 0$. This is impossible as $a_x, c_x > 0$. Therefore, the first bracket vanishes, $a_x = c_x$ and the original equations imply that

$$a_x = c_x = \frac{K}{2}\left(1 - \frac{\delta}{r}\right).$$

For further use, call this constant $A = \frac{K}{2}\left(1 - \frac{\delta}{r}\right)$. To summarise,

$$\begin{aligned} a_x = c_x &= \frac{K}{2}\left(1 - \frac{\delta}{r}\right),\ x = 1, \ldots, R \\ a_x = c_x &= 0,\ x = R+1, \ldots, L. \end{aligned}$$

At $O(\nu_N)$, the equations give the first order corrections to curvature as

$$\begin{aligned} r\left[\left(1 - \frac{3A}{K}\right) b_{Nx} - \frac{A}{K} d_{Nx}\right] \quad &+ sd_{Nx} = (\delta + s)b_{Nx}, && x = 1, \ldots, R-1 \\ r\left[\left(1 - \frac{3A}{K}\right) d_{Nx} - \frac{A}{K} b_{Nx}\right] \quad &+ sb_{Nx} = (\delta + s)d_{Nx}, && x = 1, \ldots, R-1 \\ r\left[\left(1 - \frac{3A}{K}\right) b_{Nx} - \frac{A}{K} d_{Nx}\right] - A \quad &+ sd_{Nx} = (\delta + s)b_{Nx}, && x = R \\ r\left[\left(1 - \frac{3A}{K}\right) d_{Nx} - \frac{A}{K} b_{Nx}\right] \quad &+ sb_{Nx} = (\delta + s)d_{Nx}, && x = R \\ + A \quad &+ sd_{Nx} = (\delta + s)b_{Nx}, && i = R+1 \\ &+ sb_{Nx} = (\delta + s)d_{Nx}, && x = R+1 \\ &+ sd_{Nx} = (\delta + s)b_{Nx}, && x = R+2, \ldots, L \\ &+ sb_{Nx} = (\delta + s)d_{Nx}, && x = R+2, \ldots, L \end{aligned}$$

Using the subtraction as before, it can be seen that $b_{Nx} = d_{Nx} = 0$ unless $x = R$ or $R+1$. By solving the appropriate system of linear equations for $x = R, R+1$, we also see that

$$\begin{aligned} b_{NR} &= -\frac{A}{2}\left(\frac{1}{r - \delta} + \frac{1}{2s}\right) \\ d_{NR} &= -\frac{A}{2}\left(\frac{1}{r - \delta} - \frac{1}{2s}\right) \\ b_{NR+1} &= \frac{A}{2}\left(\frac{1}{\delta} + \frac{1}{\delta + 2s}\right) \\ d_{NR+1} &= \frac{A}{2}\left(\frac{1}{\delta} - \frac{1}{\delta + 2s}\right) \end{aligned}$$

At $O(\nu_M)$, the equations are the same up to swapping M with N, and b with d.

High motility combination. To simplify the algebra, assume a large single parameter ν such that $\nu_N, \nu_M \sim O(\nu)$, denote the orders ν^k as $O(k)$ and expand $N_x = a_x + b_x/\nu + \dots$, and $M_x = c_x + d_x/\nu \dots$

At $O(-1)$, the equations describe the motility equilibrium by

$$\begin{aligned} a_1 &= a_2, \\ a_{x+1} + a_{x-1} &= 2a_x, \quad x = 1, \dots, L-1 \\ a_{L-1} &= a_L, \\ c_1 &= c_2, \\ c_{x+1} + c_{x-1} &= 2c_x, \quad x = 1, \dots, L-1 \\ c_{L-1} &= c_L, \end{aligned}$$

which imply that $a_x = a$ and $c_x = c$ for some a, c constants.

At $O(0)$, the equations describe the first order correction to curvature due to local compartmental effects by

$$\begin{aligned} r\left(1 - \frac{a+c}{K}\right)a & & + \frac{\nu_N}{\nu} b_2 + sc &= (\delta + s)a + \frac{\nu_N}{\nu} b_1, & \\ r\left(1 - \frac{a+c}{K}\right)a + \frac{\nu_N}{\nu} b_{x-1} & & + \frac{\nu_N}{\nu} b_{x+1} + sc &= (\delta + s)a + 2\frac{\nu_N}{\nu} b_x, & x = 2, \dots, R \\ + \frac{\nu_N}{\nu} b_{x-1} & & + \frac{\nu_N}{\nu} b_{x+1} + sc &= (\delta + s)a + 2\frac{\nu_N}{\nu} b_x, & x = R+1, \dots, L-1 \\ + \frac{\nu_N}{\nu} b_{L-1} & & + sc &= (\delta + s)a + \frac{\nu_N}{\nu} b_L, & \\ r\left(1 - \frac{a+c}{K}\right)c & & + \frac{\nu_M}{\nu} d_2 + sa &= (\delta + s)c + \frac{\nu_M}{\nu} d_1, & \\ r\left(1 - \frac{a+c}{K}\right)c + \frac{\nu_M}{\nu} d_{x-1} & & + \frac{\nu_M}{\nu} d_{x+1} + sa &= (\delta + s)c + 2\frac{\nu_M}{\nu} d_x, & x = 2, \dots, R \\ + \frac{\nu_M}{\nu} d_{x-1} & & + \frac{\nu_M}{\nu} d_{x+1} + sa &= (\delta + s)c + 2\frac{\nu_M}{\nu} d_x, & x = R+1, \dots, L-1 \\ + \frac{\nu_M}{\nu} d_{L-1} & & + sa &= (\delta + s)c + \frac{\nu_M}{\nu} d_L, & \end{aligned}$$

Summing the first and later four equations, we get

$$\begin{aligned} Rr\left(1 - \frac{a+c}{K}\right)a + Lsc &= L(\delta + s)a, \\ Rr\left(1 - \frac{a+c}{K}\right)c + Lsa &= L(\delta + s)c, \end{aligned}$$

Multiplying the first equation by c and the latter by a, and subtracting, we get $Ls(c^2 - a^2) = 0$. Therefore, $c = a$. Then, we get

$$a = c = \frac{K}{2}\left(1 - \frac{L\delta}{Rr}\right).$$

Mixed motility combination. Assume that there is a small ν such that $\nu_N \sim O(\nu)$ and $\nu_M \sim O(1/\nu)$. Then, $N_x = a_x + b_x\nu + \dots$, and $M_x = c_x + d_x\nu \dots$.

At $O(-1)$, the equations for high motility equilibrium are:

$$\begin{aligned} c_1 &= c_2, \\ c_{x+1} + c_{x-1} &= 2c_x, \quad x = 1, \dots, L-1 \\ c_{L-1} &= c_L. \end{aligned}$$

This result implies that there is a constant c such that $c_x = c$.

At $O(0)$, the local dynamics of the slow population involves the faster population:

$$\begin{aligned}
r\left(1-\frac{a_x+c}{K}\right)a_x \qquad\qquad\qquad\qquad &+sc &&= (\delta+s)a_x, && x=1,\dots,R\\
&+sc &&= (\delta+s)a_x, && x=R+1,\dots,L\\
r\left(1-\frac{a_1+c}{K}\right)c \qquad\qquad &+\nu_M\nu d_2+sa_1 &&= (\delta+s)c+\nu_M\nu d_1, &&\\
r\left(1-\frac{a_x+c}{K}\right)c+\nu_M\nu d_{x-1} &+\nu_M\nu d_{x+1}+sa_x &&= (\delta+s)c+2\nu_M\nu d_x, && x=2,\dots,R\\
+\nu_M\nu d_{x-1} &+\nu_M\nu d_{x+1}+sa_x &&= (\delta+s)c+2\nu_M\nu d_x, && x=R+1,\dots,L-1\\
+\nu_M\nu d_{L-1} &+sa_L &&= (\delta+s)c+\nu_M\nu d_L, &&
\end{aligned}$$

Notice that the first equation can be thought of as a quadratic equation for a_x when c is thought of as a parameter:

$$0=f(a_x)=-\frac{r}{K}a_x^2+\left[\left(1-\frac{c}{K}\right)-(\delta+s)\right]a_x+sc.$$

Since $-r/K<0$ and $f(0)=sc>0$, the parabola has exactly one root with $a_x>0$. Therefore, $a_x=a$ when $x\le R$. The second equation similarly gives that $a_x=b$ when $x>R$.

By summing the other equations, we get the system for variables a,b,c:

$$r\left(1-\frac{a+c}{K}\right)a+sc = (\delta+s)a, \tag{s32}$$

$$+sc = (\delta+s)b, \tag{s33}$$

$$Rr\left(1-\frac{a+c}{K}\right)c+s(Ra+(L-R)b) = L(\delta+s)c. \tag{s34}$$

Firstly, eliminate b from the equation (s34) to get

$$Rr\left(1-\frac{a+c}{K}\right)c+Rsa+(L-R)\frac{s^2}{\delta+s}c=L(\delta+s)c \tag{s35}$$

Multiplying equation (s35) by a and subtracting it from the equation (s32) multiplied by Rc, we can obtain

$$\frac{a}{c}-\frac{c}{a}=\left(\frac{L}{R}-1\right)\frac{\delta^2+2\delta s}{s^2+\delta s}:=A(s), \tag{s36}$$

which defines a dimensionless number $A(s)>0$. The equation for c/a can be solved by requiring that $c/a>0$:

$$\frac{c}{a}=\frac{-A+\sqrt{A^2+4}}{2}:=B(s)<1, \tag{s37}$$

where $B(s)$ is the proportion of high motility phenotype to low motility phenotype above the staircase. Combining (s32) and (s37), it is straightforward to find the solution

$$\begin{aligned}
a&=K\left(1-\frac{\delta+s(1-B(s))}{r}\right)\frac{1}{1+B(s)},\\
b&=K\left(1-\frac{\delta+s(1-B(s))}{r}\right)\frac{sB(s)}{(1+B(s))(s+\delta)},\\
c&=K\left(1-\frac{\delta+s(1-B(s))}{r}\right)\frac{B(s)}{1+B(s)}.
\end{aligned} \tag{s38}$$

□

Theorem 8. *At low switching $s\ll\delta$, the overall wild-type profile behaves as if it has a single motility of $\nu=\min(\nu_N,\nu_M)$. At high switching $s\gg\delta$, the overall wild-type profile behaves as if it has a single motility of $\nu=(\nu_N+\nu_M)/2$.*

In particular, this result affects the deadly motility regime in the (ν_N,ν_M,s) space. Let $\overline{f}=(rR-\delta L)/L$ be the average wild-type fitness. The wild-type profile is non-trivial iff:

- *the environment is source-like* $\overline{f} > 0$*, motility rates* ν_N, ν_M *are arbitrary, and switching rate* s *is arbitrary,*
- *the environment is sink-like* $\overline{f} < 0$ *and*
  - *switching is low* $s \ll \delta$ *and the slower phenotype is in the deadly motility regime* $\min(\nu_N, \nu_M) > \nu_c$*,*
  - *switching is high* $s \gg \delta$ *and the average-motility phenotype is in the deadly motility regime* $(\nu_N + \nu_M)/2 > \nu_c$*.*

Proof. As for the source-sink model, the deadly motility regime corresponds to region in the (ν_N, ν_M, s) space. Now, we will fix the motility rates ν_N, ν_M and vary s.

At high switching ($s \gg \delta$), the dominant dynamics $\dot{N}_x \approx (M_x - N_x)$ (and $N_x \leftrightarrow M_x$) pushes the dynamics close to the surface $N_x = M_x$. On this surface, we can define coordinates $N'_x = N_x + M_x$ and derive the dynamics

$$\begin{aligned}
\dot{N}'_1 &= r\left(1 - \frac{N'_1}{K}\right) N'_1 \quad -\delta N'_1 + \frac{\nu_N + \nu_M}{2}(N'_2 - N'_1), \\
\dot{N}'_x &= r\left(1 - \frac{N'_x}{K}\right) N'_x \quad -\delta N'_x + \frac{\nu_N + \nu_M}{2}(N'_{x-1} + N'_{x+1} - 2N'_x),\ x = 2, \dots, R \\
\dot{N}'_x &= \qquad\qquad -\delta N'_x + \frac{\nu_N + \nu_M}{2}(N'_{x-1} + N'_{x+1} - 2N'_x),\ x = R+1, \dots, L-1 \\
\dot{N}'_L &= \qquad\qquad -\delta N'_L + \frac{\nu_N + \nu_M}{2}(N'_{L-1} - N'_L),
\end{aligned}$$

Therefore, the overall population N'_x behaves as a population with average motility. In particular, the overall population is in the deadly motility regime at high switching ($s \gg \delta$) iff the average-motility population is.

For low switching, we prove the stability of the non-trivial fixed point at $s = 0$. This fixed point is given by $M_x = 0$ and N_x, which coincides with the non-trivial fixed point of the staircase model for a single phenotype of motility $\nu = \nu_N$ (s6). The corresponding Jacobian can be derived from (s16) and has the form

$$J = \begin{pmatrix} J_1 & \star \\ 0 & J_2(\nu_M) \end{pmatrix},$$

where

$$J_1 = \begin{pmatrix} r(1 - 2N_1/K) - \nu_N - \delta & \nu_N & 0 & \dots \\ \nu_N & r(1 - 2N_2/K) - 2\nu_N - \delta & \nu_N & \dots \\ 0 & \nu_N & r(1 - 2N_3/K) - 2\nu_N - \delta & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and

$$J_2(\nu_M) = \begin{pmatrix} r(1 - N_1/K) - \nu_M - \delta & \nu_M & 0 & \dots \\ \nu_M & r(1 - N_2/K) - 2\nu_M - \delta & \nu_M & \dots \\ 0 & \nu_M & r(1 - N_3/K) - 2\nu_M - \delta & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

We prove the stability of the non-trivial fixed point if we show that the eigenvalues of J have negative real part. Notice that the set of eigenvalues of J is the union of the sets of eigenvalues of J_1 and $J_2(\nu_M)$. Therefore, it suffices to show that the eigenvalues of J_1 and $J_2(\nu_M)$ are negative. The eigenvalues of J_1 must be negative since J_1 is the Jacobian corresponding to the non-trivial fixed point of the staircase model for a single phenotype with motility ν_N, which is stable from previous analysis. The proof that the eigenvalues of $J_2(\nu_M)$ are negative for $\nu_M > \nu_N$ is identical to the proof [S18], which consists of two steps. First, one shows that the largest eigenvalue of $J_2(\nu_N)$ is 0 because $(N_1, \dots, N_L)$ is the eigenvector and only the largest eigenvalue can have an eigenvector with positive entries. Second, one shows that this eigenvalue increases as ν_M increases from ν_N. To sum up, the overall population is dominated by the slower population N_x at low switching ($s \ll \delta$).

The result for the existence of the deadly motility regime follows, as the system behaves as a single phenotype of single motility which interpolates between the slower motility $\min(\nu_N, \nu_M)$ (at low switching $s \ll \delta$) and the average motility $(\nu_N + \nu_M)/2$ (at high switching $s \gg \delta$). □

S8. Figure Notes

This section aims to provide the parameter values used for producing all our figures as well as other specific guidance for reproducing our results fully. Unless stated otherwise, we follow previous work [S7] and use as lattice size $L = 8$, carrying capacity $K = 10^5$, mutation rates $\mu_f = 10^{-7}$ and $\mu_b = 10^{-4}$, division rate $r = 1$ and death rate $\delta = 0.1$ (resp. $\delta = 0.3$) for source-like (resp. sink-like) environments.

Specifically, in Fig. 1, the death rate is $\delta = 0.1$. Panel b was created from a simulation of the staircase model with $\nu = 1e-2$ for low motility and $\nu = 1$ for high motility. Panel d was created from the same simulations. Moreover, the black curve describing the total population numbers was plotted by finding the non-trivial fixed point of the mean-field theory (s6) via the Newton-Raphson method.

In Fig. 2, motility rate ν is varied in logarithmic steps between 10^{-5} and 10, and the appropriate death rate $\delta = 0.1$, 0.3 is chosen for the source-like and sink-like environments, respectively. The dots of panel b correspond to simulations while the lines correspond to the analytical techniques. Panel c was produced by the analytical techniques. Panel d was plotted by finding the non-trivial fixed point of the mean-field theory (s6) via the Newton-Raphson method.

In Fig. 3, the death rate is $\delta = 0.3$ and the switching rates s are $s = 0$ (low switching $s \ll \delta$) and $s = 5$ (high switching $s \gg \delta$). In panel b, motility rates $\nu_{1,2}$ are varied in logarithmic steps between 10^{-5} and 10 and the resulting heatmap plots of the adaptation rate are constructed for resistance state $R = 5$ (Fig. 1b) from simulations, where R is the genotype of highest resistance before the adaptation jump $R \to R + 1$. In panel c, motility combinations are represented by $(\nu_1, \nu_2) = (10^{-2}, 10^{-4})$ (low motility), $(\nu_1, \nu_2) = (3.16, 10^{-4})$ (mixed motility) and $(\nu_1, \nu_2) = (3.16, 1)$ (high motility). Wild-type profiles are plotted for resistance state $R = 4$ from simulations. Black curves describe the total number of cells in each phenotype and were plotted by finding the non-trivial fixed point of the mean-field theory s39 via the Newton-Raphson method,

$$\begin{aligned} 0 = &\quad -sN_x + sM_x + r\left(1 - \frac{N_x + M_x}{K}\right) N_x \mathbb{1}_{x \le R} - \delta N_x + \nu(N_{x-1} - N_x)\mathbb{1}_{1<x} + \nu(N_{x+1} - N_x)\mathbb{1}_{x<L} \\ 0 = &\quad +sN_x - sM_x + r\left(1 - \frac{N_x + M_x}{K}\right) M_x \mathbb{1}_{x \le R} - \delta M_x + \nu(M_{x-1} - M_x)\mathbb{1}_{1<x} + \nu(M_{x+1} - M_x)\mathbb{1}_{x<L} \end{aligned} \tag{s39}$$

In panel b of Fig. 4, the death rate is $\delta = 0.3$ and switching thresholds S are $S = 9.5 \times 10^4$ (high threshold $S \gg K(1 - \delta/r)$) and $S = 100$ (low threshold $S \ll K(1 - \delta/r)$). Motility rates $\nu_{L,H}$ are varied in logarithmic steps between 10^{-5} and 10 and the resulting heatmap plots of the adaptation rate are constructed for resistance state $R = 2$ from simulations. In panel c, the death rate is $\delta = 0.1$ and switching thresholds S are $S = 9.5 \times 10^4$ (high threshold $S \gg K(1 - \delta/r)$) and $S = 10^{4.5}$ (low threshold $S \ll K(1 - \delta/r)$). This choice is different from the one used in panel b but allows for better visualisation of the wild-type profiles when the scale of cell numbers is kept linear as in previous figures. The mixed motility combination is represented by $(\nu_L, \nu_H) = (3.16, 10^{-4})$ (fast-to-slow) and $(\nu_L, \nu_H) = (10^{-4}, 3.16)$ (slow-to-fast). The wild-type profiles (including the black curves) are plotted for resistance state $R = 1$ from simulations. The Newton-Raphson method is not reliable in this case because the step-function used for density-dependent motility is not smooth.

In Fig. 5, we consider $L = 2$ and $R = 1$. $K = 10^5$ and $r = 1$ are chosen as usual and mutation rates do not need not be chosen as they do not affect the dynamics on ecological time-scales. In panel b, the phase portraits are plotted with matplotlib [S19] for the choice of the death rate $\delta = 0.25$ and the motility rates $\nu = 0$, 0.1, 100 (source-like environment), resp. for the death rate $\delta = 0.75$ and the motility rates $\nu = 0$, 0.1, 5 (sink-like environment). In panel c, the bifurcation diagrams are plotted using the solution for the non-trivial fixed point (s1). The source-like environment is represented by $\delta = 0.25$ and the sink-like environment is represented by $\delta = 0.75$.

In Fig. S1, the death rate $\delta = 0.1$, 0.3 is chosen for source-like and sink-like environment, respectively. Motility rate ν is varied in logarithmic steps between 10^{-5} and 10 and resistance cost c is varied in logarithmic steps between $10^{-4.5}$ and $10^{-0.5}$. The resulting heatmap plots of adaptation rate are constructed for resistance state $R = 3$ from simulations.

In Fig. S2, the death rate is $\delta = 0.1$. Motility rate ν is varied in logarithmic steps between 10^{-5} and 10 and the multiplicative increase in death rate σ is varied linearly between 1 and 9. The resulting heatmap plots of the adaptation rate are constructed for resistance state $R = 5$ from simulations.

In Fig. S3, the death rate is $\delta = 0.1$. Motility rate ν is varied in logarithmic steps between 10^{-5} and 10 and the chemotactic probability to move up the antibiotic gradient p is varied linearly between 0.1 and 0.9. The resulting heatmap plots of the adaptation rate are constructed for resistance state $R = 5$ from simulations.

In panel a of Fig. S4 the switching rate is varied in logarithmic steps between 10^{-5} and 10. The mixed motility combination is chosen as $(\nu_1, \nu_2) = (10^{-3}, 10)$ and the appropriate death rate $\delta = 0.1$, 0.3 is chosen for the source-like and sink-like environment, respectively. The dots of panel b correspond to simulations while the lines are obtained by analytical techniques for a population of homogeneous motility $\min(\nu_1, \nu_2)$ at low switching and $(\nu_1 + \nu_2)/2$ at high

switching. In panel b, the death rate is $\delta = 0.1$ and the switching rate is varied $s = 0$, 0.001, 0.01, 0.1, 1. Motility rates $\nu_{1,2}$ are varied in logarithmic steps between 10^{-5} and 10 and the resulting heatmap plots of the adaptation rate are constructed for resistance state $R = 5$ from simulations. In panel c, the death rate is $\delta = 0.3$ and the switching rate is varied $s = 0.1$, 0.4, 0.67, 0.9, 1, 10. Motility rates $\nu_{1,2}$ are varied in logarithmic steps between $\nu = 10^{-5}$ and $\nu = 10$. The survival probability is computed as follows. We run the Newton-Raphson method to search for a fixed point of (s39) in the non-negative quadrant $N_x, M_x \geq 0$. If the search returns the trivial (resp. non-trivial) fixed point, we set the survival probability to 0 (resp. 1).

In Fig. S5, the death rate is $\delta = 0.1$ and switching thresholds S are $S = 9.5 \times 10^4$ (high threshold $S \gg K(1 - \delta/r)$) and $S = 100$ (low threshold $S \ll K(1 - \delta/r)$). Motility rates $\nu_{L,H}$ are varied in logarithmic steps between 10^{-5} and 10 and the resulting heatmap plots of the adaptation rate are constructed for resistance state $R = 2$ from simulations.

In Fig. S6, the death rate $\delta = 0.1$, 0.3 is chosen for the source-like and sink-like environment, respectively. Motility rate ν is varied in logarithmic steps between 10^{-5} and 10 and the horizontal gene transfer rate h is in logarithmic steps between 10^{-7} and 10^{-3}. The resulting heatmap plots of the adaptation rate are constructed for resistance state $R = 4$ from simulations.

SUPPLEMENTARY FIGURES

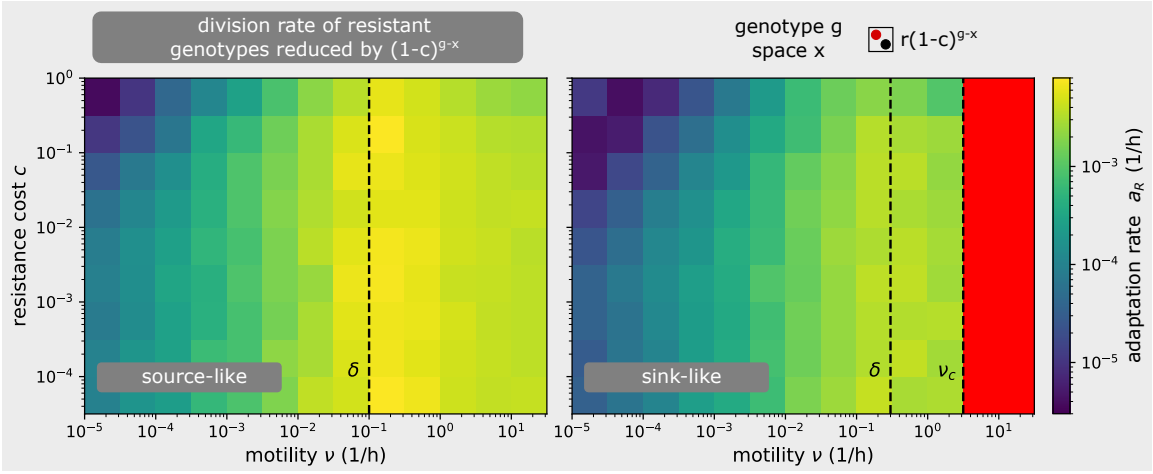


FIG. S1. **Effect of resistance costs.** Resistance cost c is modelled as a reduction in the division rate of genotypes g at position x by a factor of $(1-c)^{g-x}$ as in [S7]. The heatmap of adaptation rate on the (ν, c) plane explains how resistance cost c changes the relationship between motility rate ν and adaptation rate a_R. Generically, resistance cost does not affect the relationship between adaptation rate and motility: low motility accelerates adaptation while high motility decelerates adaptation. Only when the motility rate is very low $\nu \ll \delta$ and resistance cost is very high $c > \sqrt{\nu/\delta}$, the adaptation rate is decreased by c, as proved in [S7]. However, such resistance costs are considered high [S8]. Even if the motility is low as $\nu = 10^{-2}\delta$ (i.e., 99% of bacteria do not migrate during their lifetime), the resistance cost is important only if $c > 0.1$ [S8]. Moreover, resistance cost cannot influence the deadly motility regime (red) as in this regime the susceptible wild-type goes extinct. A detailed description of this model can be found in SI Text S4.

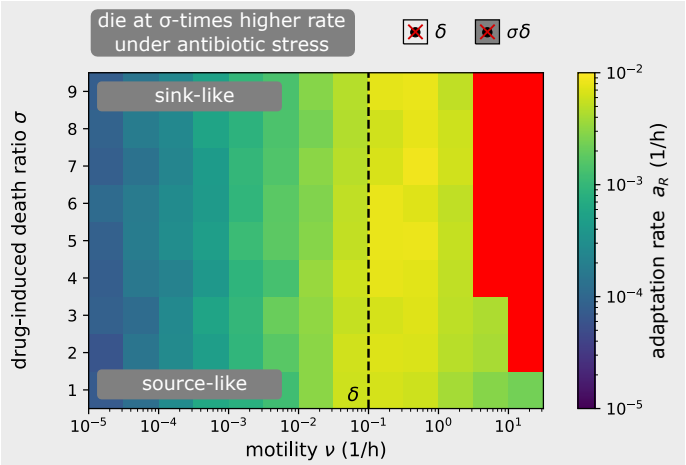


FIG. S2. **Effect of bactericidal antibiotics.** Bactericidal antibiotics are considered by increasing the death rate of susceptible genotypes $g < x$ (under the staircase) σ-times. The heatmap of adaptation rate on the (ν, σ) plane explains how σ changes the relationship between motility ν and adaptation rate a_R. In the low motility regime $\nu < \delta$, the adaptation rate increases with motility ν and is unaffected by σ. In the high motility regime $\nu > \delta$, the adaptation rate decreases with motility ν. Moreover, even if the environment is source-like for $\sigma = 1$, the high motility regime can transition into the deadly motility regime (red) provided the bactericidal effect is strong enough $\sigma \gg 1$. A detailed description of this model can be found in SI Text S4.

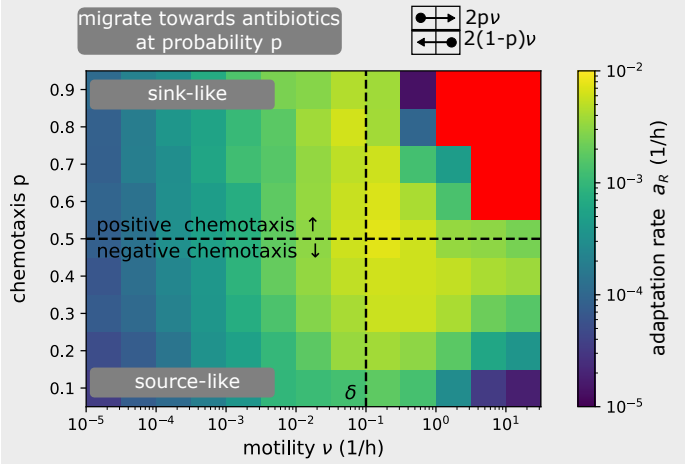


FIG. S3. **Effect of chemotaxis.** Motility bias in the chemical gradient (chemotaxis) is considered by varying the probability p that a cell moves up or down the antibiotic gradient. In our original model $p = 0.5$. Cell movement is biased up the antibiotic gradient if $p > 0.5$ (positive chemotaxis), and down the gradient $p < 0.5$ (negative chemotaxis). The heatmap of adaptation rate on the (ν, σ) plane explains how p changes the relationship between motility ν and adaptation rate a_R. In the low motility regime $\nu < \delta$, the adaptation rate increases with motility ν and is unaffected by chemotaxis p. In high motility regime $\nu > \delta$, positive chemotaxis $p > 0.5$ converts the high motility regime into the deadly motility regime even when the environment is source-like at $p = 0.5$. Negative chemotaxis $p < 0.5$ prevents the deadly motility regime and slows down the adaptation in the high motility regime, compared to unbiased motion $p = 0.5$. A detailed description of this model can be found in SI Text S4.

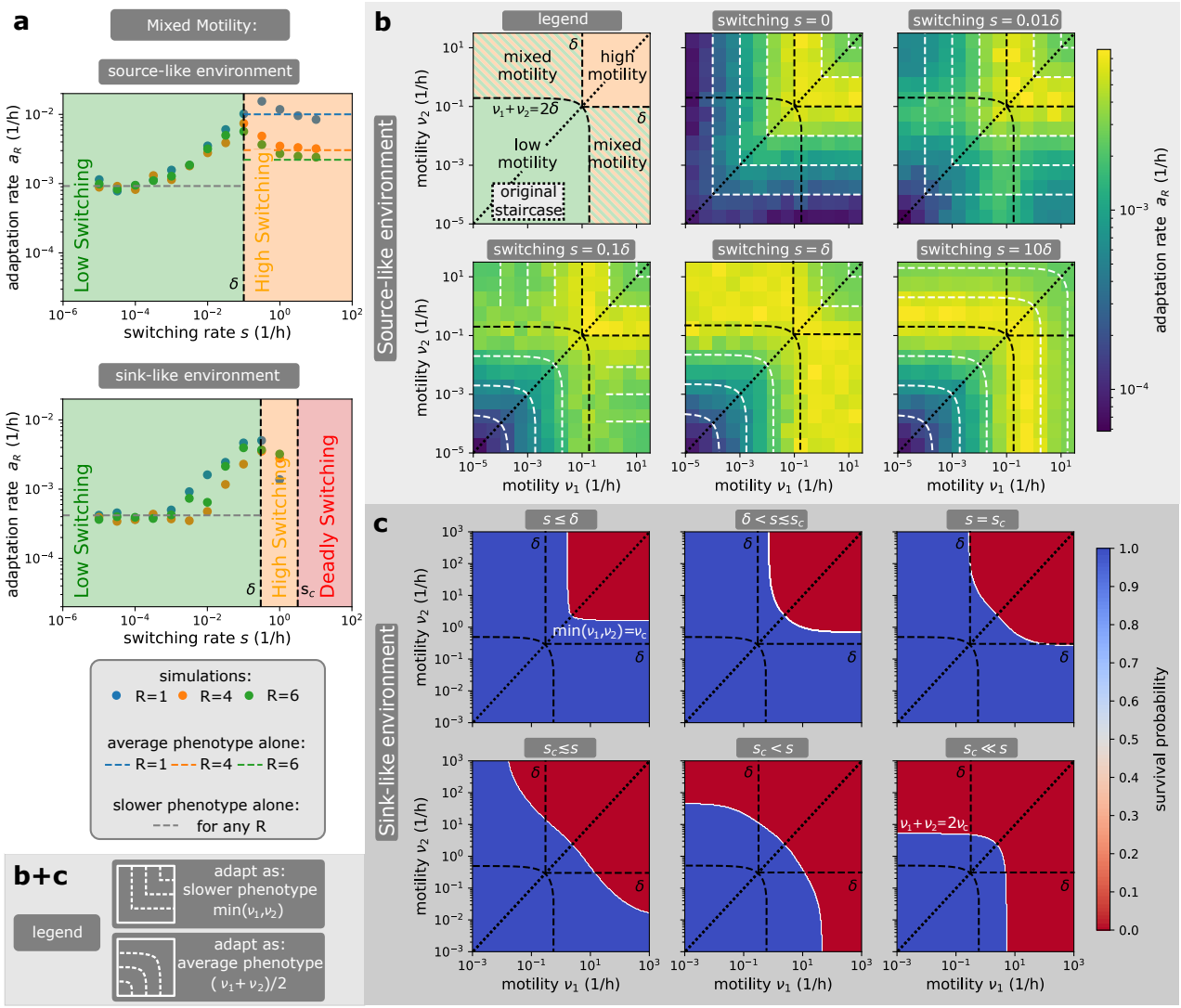


FIG. S4. **Further effects of switching between motility phenotypes stochastically. a)** Switching rate s modifies the adaptation rate a_R and the adaptation regime in the mixed motility combination $\nu_1 \ll \delta \ll \nu_2$. Dots correspond to simulations, while the asymptotic lines are derived with analytical techniques. At low switching ($s \ll \delta$), the population adapts as if a single slower phenotype is present, whose adaptation rate is independent of the resistance state R (grey dashed line). At high switching ($s \ll \delta$), the population behaves as if a single phenotype of average motility is present, whose adaptation rate depends on the number of compartments R where wild-type can divide (coloured dashed lines). Moreover, if the environment is sink-like and the average motility is above the critical motility, there is a critical switching rate s_c which limits bacterial survival. **b)** Adaptation rate heatmap on the (ν_1, ν_2) plane for different switching rates s. This is the same plot as in Fig. 3b but for a source-like environment instead of a sink-like environment. The (ν_1, ν_2) plane can be partitioned into different combinations of adaptation regimes and its diagonal corresponds to a population of a single motility. Bacteria adapt as if all cells had the same effective motility, which corresponds to the intersection of the level sets (white dashed lines) with the diagonal. At low switching rate s, the effective motility matches the slower motility phenotype present $\min(\nu_1, \nu_2)$. At high switching rate s, the effective motility matches the average motility $(\nu_1 + \nu_2)/2$. Unlike for the sink-like environment, there is no deadly motility regime for any motility $\nu_{1,2}$ and switching s in the source-like environment. **c)** Deadly motility regime and critical surface. In sink-like environments, there is an abrupt change in survival probability from 1 to 0 when the population is sufficiently well-mixed. The deadly motility regime corresponds to the space with zero survival probability (red) and is bounded by a critical surface (white). The critical surface asymptotes with $\min(\nu_1, \nu_2) = \nu_c$ at low switching ($s \ll \delta$) and with $(\nu_1 + \nu_2)/2 = \nu_c$ at high switching ($s \gg \delta$).

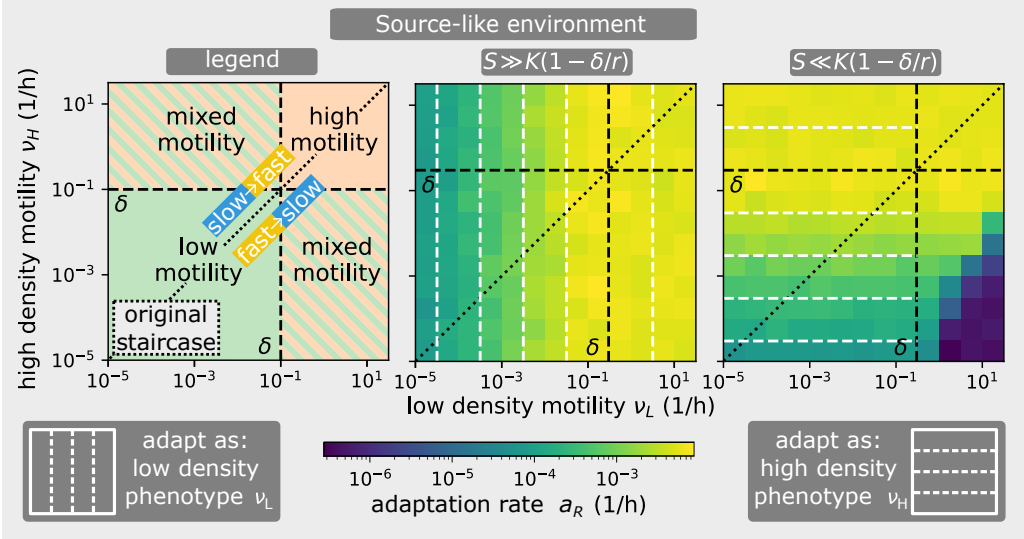


FIG. S5. **Further effects of density-dependent motility.** Adaptation rate heatmap on the (ν_L, ν_H) plane for different switching thresholds S in the source-like environment, where $\nu_{L,H}$ are the low-density/high-density motility rates (Fig. 4a). This is the same plot as in Fig. 4b but for a source-like environment instead of a sink-like environment. The (ν_L, ν_H) plane can be partitioned into different combinations of adaptation regimes and its diagonal corresponds to a population of a single motility. The diagonal separates slow-to-fast and fast-to-slow switching combinations, which differ in relative motility at low-to-high density. Bacteria generically adapt as if all cells had the same effective motility, which corresponds to the intersection of the level sets (white dashed lines) with the diagonal. This effective motility generically matches the low-density (resp. high-density) motility at high (resp. low) threshold S. At low threshold S, this generic rule has one exception. The adaptation rate is reduced in the mixed motility combination of the fast-to-slow switching case. The reduction in adaptation rate in slow-to-fast switching in Fig. 4b is caused by the critical motility, which does not appear in the source-like environment.

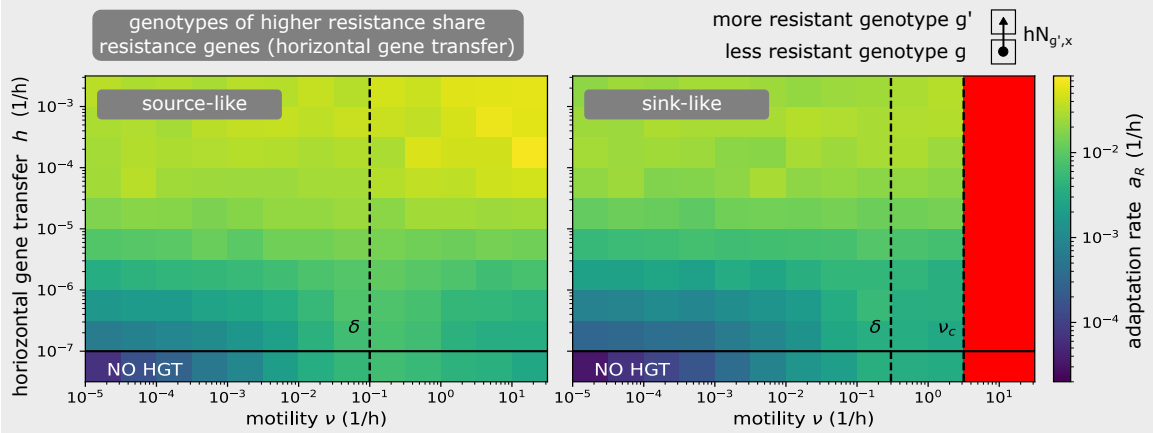
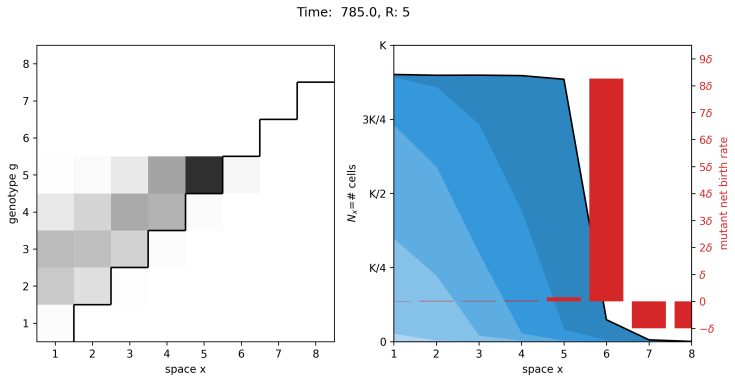
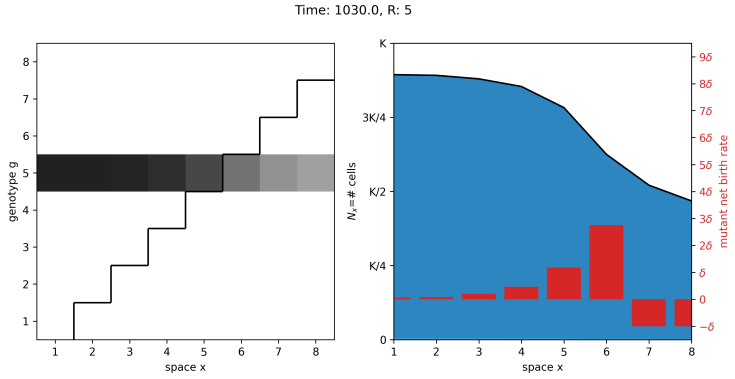


FIG. S6. **Effect of horizontal gene transfer (HGT).** The heatmap of adaptation rate on the (ν, h) plane explains how the HGT rate h changes the relationship between motility ν and adaptation rate a_R. HGT is modelled by allowing cells of less resistant genotypes g to acquire resistance from $N_{x,g'}$ more resistant cells of genotype $g' > g$ located at the same position x at a rate $h \times N_{x,g'}$. When HGT rate is low, the relationship between adaptation rate and motility is not significantly changed: low motility accelerates adaptation while high motility decelerates adaptation. When HGT rate is high, the adaptation rate is significantly increased and changes in motility do not play a significant role. However, the required HGT threshold $h \approx 10^{-5}$ corresponds to $hK/\delta \approx 1$ horizontal gene transfers per generation. Such HGT rates are considered very high, compared to the upper bound of 10^{-4} transfers per generation known from the literature [S12]. Moreover, HGT cannot influence the deadly motility regime (red) as the donors of resistance genes do not have enough time to form before the wild-type goes extinct. A detailed description of this model can be found in SI Text S4.

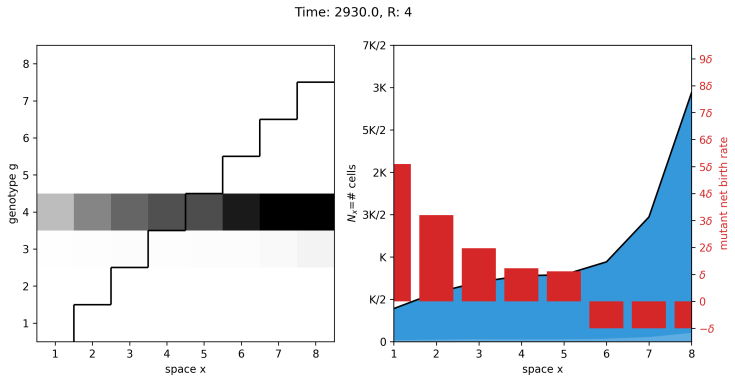
SUPPLEMENTARY MOVIES



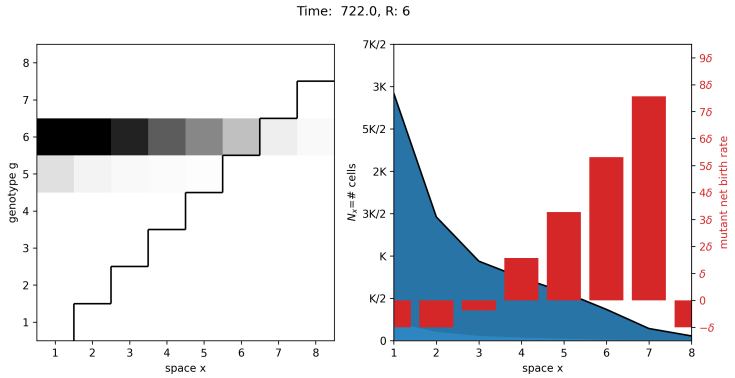
MOVIE S1. **Bacterial adaptation in the low motility regime ($\nu < \delta$).** Simulation of a bacterial population evolving antibiotic resistance in the staircase model in the low motility regime. It adapts and expands its range in a stepwise fashion, leaving behind an inclined comet tail (left) where shades of gray represent cell density in each compartment, which corresponds to a diversity of strains with different antibiotic susceptibility (right). Find a full description in the caption of Fig. 1. Parameters used: $L = 8$, $K = 10^5$, $r = 1$, $\delta = 0.1$, $\mu_f = 10^{-7}$, $\mu_b = 10^{-4}$, $\nu = 0.01$.



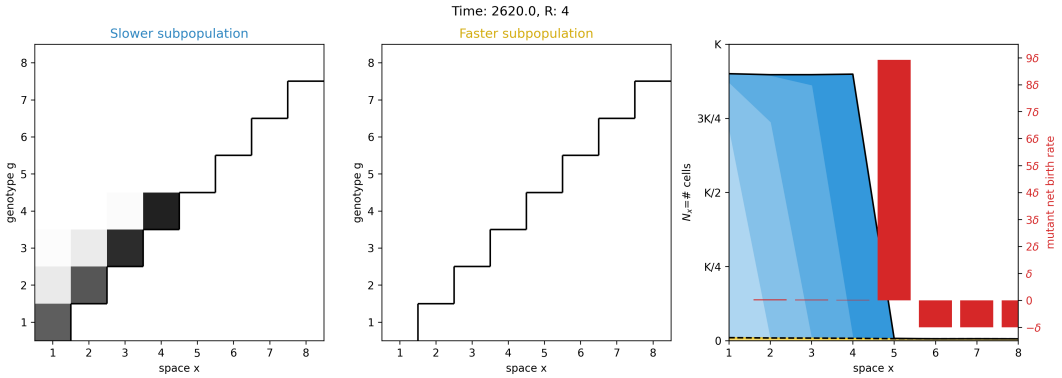
MOVIE S2. **Bacterial adaptation in the high motility regime ($\nu > \delta$).** Similar to Movie S1, but for the high motility regime. Note that the comet tail that was inclined in the low motility regime is now horizontal (left), and the population is made of a single strain that can grow in $x \leq R$ (right). Find a full description in the caption of Fig. 1. Parameters used: $L = 8$, $K = 10^5$, $r = 1$, $\delta = 0.1$, $\mu_f = 10^{-7}$, $\mu_b = 10^{-4}$, $\nu = 1$.



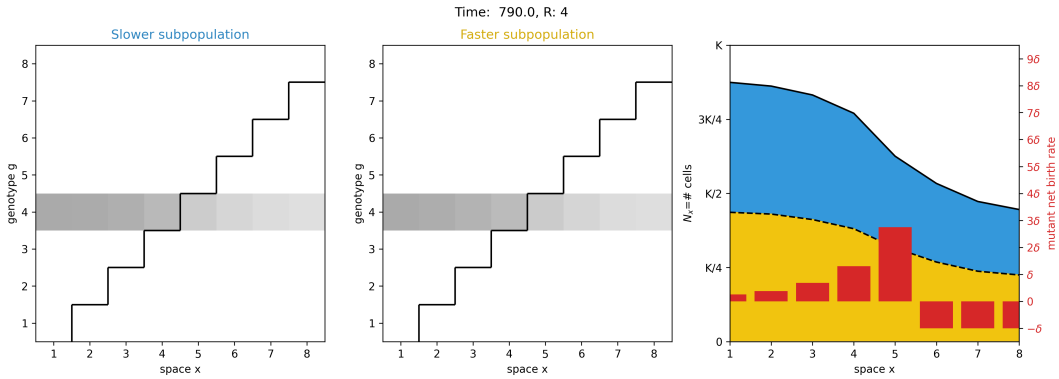
MOVIE S3. **Effect of positive chemotaxis**. Similar to Movie S2 but now $p > 0.5$ and the population survives. Parameters used: $L = 8$, $K = 10^5$, $r = 1$, $\delta = 0.1$, $\mu_f = 10^{-7}$, $\mu_b = 10^{-4}$, $\nu = 1$, $p = 0.7$.



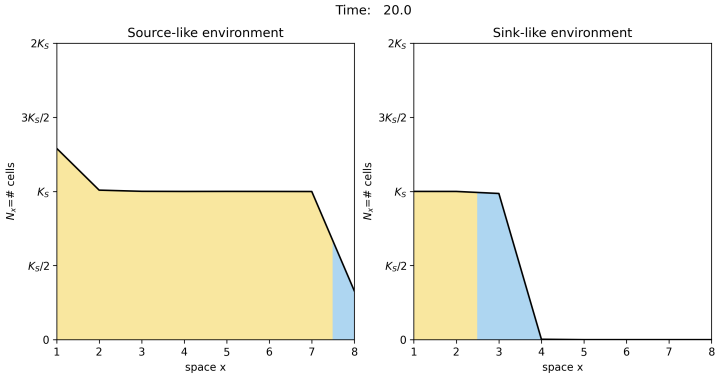
MOVIE S4. **Effect of negative chemotaxis**. Similar to Movie S2 but now $p < 0.5$. Parameters used: $L = 8$, $K = 10^5$, $r = 1$, $\delta = 0.1$, $\mu_f = 10^{-7}$, $\mu_b = 10^{-4}$, $\nu = 1$, $p = 0.3$.



MOVIE S5. **Effect of stochastic switching when the switching rate is low**. Simulation of a bacterial population evolving antibiotic resistance in the staircase model with stochastic switching between motility phenotypes for the mixed motility combination $\min(\nu_1, \nu_2) < \delta < (\nu_1 + \nu_2)/2$. Find a full description in the caption of Fig. 3. Parameters used: $L = 8$, $K = 10^5$, $r = 1$, $\delta = 0.1$, $\mu_f = 10^{-7}$, $\mu_b = 10^{-4}$, $\nu_1 = 10^{-4}$, $\nu_2 = 10^{0.5}$, $s = 10^{-3}$.



MOVIE S6. **Effect of stochastic switching when the switching rate is high**. Similar to Movie S5 but with high rates of switching between motility phenotypes. Find a full description in the caption of Fig. 3. Parameters used: $L = 8$, $K = 10^5$, $r = 1$, $\delta = 0.1$, $\mu_f = 10^{-7}$, $\mu_b = 10^{-4}$, $\nu_1 = 10^{-4}$, $\nu_2 = 10^{0.5}$, $s = 5$.



MOVIE S7. **Effect of density-dependent motility**. Simulation of a bacterial population evolving antibiotic in the staircase model with slow-to-fast density-dependent motility. If the environment is source-like (left), the faster high-density population bulk (yellow) expands its range by pulling the slower low-density population front (blue). If the environment is sink-like (right), and the faster high-density motility phenotype moves above the critical motility, this expansion is halted. As a result of this reduced antibiotic exposure, the adaptation rate is decreased in the top row of Fig. 4b (sink-like environment) but not in the top row of Fig. S5 (source-like environment). Parameters used: $L = 8$, $K = 10^5$, $r = 1$, $\delta = 0.1$ (left) and $\delta = 0.3$ (right), $\mu_f = 10^{-7}$, $\mu_b = 10^{-4}$, $\nu_L = 10^{-3}$, $\nu_H = 10^{0.5}$, $S = 10^4$.

REFERENCES

[S1] Bulmer, M. Multiple niche polymorphism. *The American Naturalist* **106**, 254–257 (1972).

[S2] Holt, R. D. & Gomulkiewicz, R. How does immigration influence local adaptation? a reexamination of a familiar paradigm. *The American Naturalist* **149**, 563–572 (1997).

[S3] Lenormand, T. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution* **17**, 183–189 (2002).

[S4] Hastings, A. Dynamics of a single species in a spatially varying environment: the stabilizing role of high dispersal rates. *Journal of mathematical biology* **16**, 49–55 (1982).

[S5] Lipsitch, M. The rise and fall of antimicrobial resistance. *Trends in microbiology* **9**, 438–444 (2001).

[S6] Andersson, D. I. The biological cost of mutational antibiotic resistance: any practical conclusions? *Current opinion in microbiology* **9**, 461–465 (2006).

[S7] Hermsen, R., Deris, J. & Hwa, T. On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient. *Proceedings of the National Academy of Sciences* **109**, 10775 – 10780 (2012).

[S8] Hermsen, R. & Hwa, T. Sources and sinks: a stochastic model of evolution in heterogeneous environments. *Physical review letters* **105**, 248104 (2010).

[S9] Oliveira, N. M. *et al.* Suicidal chemotaxis in bacteria. *Nature Communications.* Accepted (2022).

[S10] Bru, J.-L. *et al.* Pqs produced by the pseudomonas aeruginosa stress response repels swarms away from bacteriophage and antibiotics. *Journal of bacteriology* **201**, e00383–19 (2019).

[S11] Von Wintersdorff, C. J. *et al.* Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Frontiers in microbiology* 173 (2016).

[S12] Niehus, R., Mitri, S., Fletcher, A. G. & Foster, K. R. Migration and horizontal gene transfer divide microbial genomes into multiple niches. *Nature communications* **6**, 1–9 (2015).

[S13] Cohan, F. M. The effects of rare but promiscuous genetic exchange on evolutionary divergence in prokaryotes. *The American Naturalist* **143**, 965–986 (1994).

[S14] Cohan, F. M. Does recombination constrain neutral divergence among bacterial taxa? *Evolution* **49**, 164–175 (1995).

[S15] Shapiro, B. J., David, L. A., Friedman, J. & Alm, E. J. Looking for darwin's footprints in the microbial world. *Trends in microbiology* **17**, 196–204 (2009).

[S16] Gevers, D. *et al.* Re-evaluating prokaryotic species. *Nature Reviews Microbiology* **3**, 733–739 (2005).

[S17] Levin, B. R. Periodic selection, infectious gene exchange and the genetic structure of e. coli populations. *Genetics* **99**, 1–23 (1981).

[S18] Hastings, A. Can spatial variation alone lead to selection for dispersal? *Theoretical Population Biology* **24**, 244–251 (1983).

[S19] Barrett, P., Hunter, J., Miller, J. T., Hsu, J.-C. & Greenfield, P. matplotlib–a portable python plotting package. *Astronomical data analysis software and systems XIV* **347**, 91 (2005).