The adaptation rate of a quantitative trait in an environmental gradient

R Hermsen
Theoretical Biology Group, Biology Department, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands
E-mail: r.hermsen@uu.nl

Keywords: evolution, environmental gradients, species range, antibiotic resistance, front propagation, stochastic modeling, simulation methods

Abstract

The spatial range of a species habitat is generally determined by the ability of the species to cope with biotic and abiotic variables that vary in space. Therefore, the species range is itself an evolvable property. Indeed, environmental gradients permit a mode of evolution in which range expansion and adaptation go hand in hand. This process can contribute to rapid evolution of drug resistant bacteria and viruses, because drug concentrations in humans and livestock treated with antibiotics are far from uniform. Here, we use a minimal stochastic model of discrete, interacting organisms evolving in continuous space to study how the rate of adaptation of a quantitative trait depends on the steepness of the gradient and various population parameters. We discuss analytical results for the mean-field limit as well as extensive stochastic simulations. These simulations were performed using an exact, event-driven simulation scheme that can deal with continuous time-, density- and coordinate-dependent reaction rates and could be used for a wide variety of stochastic systems. The results reveal two qualitative regimes. If the gradient is shallow, the rate of adaptation is limited by dispersion and increases linearly with the gradient slope. If the gradient is steep, the adaptation rate is limited by mutation. In this regime, the mean-field result is highly misleading: it predicts that the adaptation rate continues to increase with the gradient slope, whereas stochastic simulations show that it in fact decreases with the square root of the slope. This discrepancy underscores the importance of discreteness and stochasticity even at high population densities; mean-field results, including those routinely used in quantitative genetics, should be interpreted with care.

1. Introduction

In those classical models of evolving populations that include spatial degrees of freedom, the habitat typically has a fixed range and geometry [1]. Adaptation, then, is driven by the selection of traits that improve an organism’s performance in that habitat. In reality, the habitat and spatial range of a species are themselves evolvable properties. Generally, the species range is affected by the species’ ability to cope with biotic, abiotic, or antibiotic conditions that are heterogeneous in space, and traits that affect this ability are clearly subject to natural selection [2]. Many basic results of population genetics, based on standard models such as the Wright–Fisher and Moran models [3], are not applicable to this mode of adaptation, and attempts to formulate and study suitable alternatives are few.

Many range boundaries observed in nature appear relatively stable, and yet cannot be explained by an obvious shift in environmental conditions [4]. For this reason, a significant amount of theoretical work has been devoted to identifying the factors that determine the range limits of species and the mechanisms that contribute to their stability (e.g., see [2, 4–9] and references therein). These studies often rely on the framework of quantitative genetics. Perhaps because of this focus on understanding stable boundaries, less attention has been given to the dynamics of range expansion when the boundaries are not evolutionarily stable.

In the recent years, however, adaptation driven by environmental heterogeneity and range expansion
have received additional attention because of their possible implications for the evolution of drug resistant bacteria and viruses [10–16]. In humans and livestock treated with antibiotics, drug concentrations differ between tissues, organs, and individuals [17]. Environments with low concentrations can function as ‘sanctuaries’, ‘reservoirs’, or ‘sources’ [10–12, 18–21] of poorly adapted strains, and local concentration gradients can provide a ‘staircase’ for adaptation, allowing it to proceed step by step, with adaptation and range expansion going hand in hand [14–16]. Indeed, in experiments using micro-environments with an antibiotic gradient, E. coli rapidly evolved drug resistance [22]. Despite this recent work, many basic questions remain unanswered. If heterogeneity can promote adaptation, then what geometries optimize the rate of adaptation? What parameters and quantities tend to affect the rate of adaptation most strongly? What determines the ‘Goldilocks point’ for the rate of adaptation [23], or in other words, what magnitude of the gradient leads to the optimal ‘resistance-selective environment’ [24]?

Here, we present a toy model designed to explore such questions. It describes an asexual population adapting to a linear environmental gradient that limits its range. We choose a highly simplified description that allows basic intuitions to be developed and tested. While motivated by, and possibly relevant to, the evolution of antibiotic resistance in bacteria, we deliberately keep the formulation general.

Many phenotypes, including antibiotic resistance levels, are determined by a combination of genetic factors as well as many intrinsic and extrinsic non-genetic (environmental) influences. Nevertheless, previous models of resistance evolution in antibiotic gradients assumed a discrete, one-dimensional, unbranched space of genotypes and a one-to-one genotype-to-phenotype mapping [14, 15]. (We discuss these models in more detail in the Discussion and Conclusions section.) While such a fitness landscape can perhaps approximate more complex, higher-dimensional systems projected onto a single reaction coordinate, the regular discretization is certainly an idealization. Here, we take an orthogonal approach by assuming that the evolving phenotype is a quantitative trait, that is, a phenotype that can be characterized by a continuous variable [25]. Examples of quantitative traits are height, weight, and blood pressure, and their (apparent) continuity usually derives from a combination of polygenic effects and/or environmental influences. Even though the drug resistance level of a bacterium, often expressed as the minimal inhibitory concentration (MIC), can in principle take on continuous values and is certainly affected by many genetic and non-genetic factors, we do not claim that this approach is more realistic per se; rather, the approaches should be seen as complementary.

Below, we first define the model. We then proceed to study its behavior in the mean-field limit, where we can derive an analytical expression for the adaptation rate. Next, we compare the mean-field result to discrete-stochastic simulations. In order to simulate large numbers of interacting individuals dispersing and evolving in continuous space and time, we have used an event-driven simulation method that is both exact and surprisingly efficient.

We will see that the behavior of the model shows two regimes: a dispersion-limited regime and a mutation-limited regime. In the dispersion-limited regime, the mean-field prediction provides useful intuitions and qualitatively sound results, even though, quantitatively, its deviations from stochastic simulations can be substantial. In the mutation-limited regime, however, the mean-field prediction becomes highly misleading even at very large population densities, demonstrating that such approximations, including those routinely used in quantitative genetics, must be used with caution.

2. The ‘ramp’ model

We envision a population of organisms in a one-dimensional continuous space. Each organism is characterized by its position \(x\) and its trait value \(y\); the population can therefore be represented by a cloud of points \((x, y)\) in a two-dimensional phenotype–position plane (see figure 1). The organisms disperse in real space (horizontally) by diffusion, with diffusion constant \(D_d\). Assuming that mutations cause small changes in the trait value \(y\), we also model phenotypic

\[
\begin{array}{c}
\text{reproduction} \\
\text{parameters:} \\
\text{death} \quad d \\
\text{reproduction} \quad r \\
\text{dispersion} \quad D_d \\
\text{mutation} \quad D_m \\
\text{carrying capacity} \quad k \\
\text{slope} \quad a \\
\text{interaction range} \quad s
\end{array}
\]

Figure 1. The ‘ramp’ model of adaptation in an environmental gradient. Asexual organisms (black dots) inhabit a one-dimensional environment, plotted vertically, and are characterized by the value of a quantitative trait \(y\), plotted horizontally. They disperse diffusively (horizontally) with diffusion constant \(D_d\). Their quantitative trait value can change due to mutation and/or environmental factors, modeled as diffusion in the vertical direction with diffusion constant \(D_m\). Organisms die at a rate \(d\). They reproduce (duplicate) logistically, that is, at a rate decreasing linearly from \(r \to 0\) as the local density increases from \(0\) to the carrying capacity \(k\); the local density is assessed by a kernel density estimate with standard deviation (interaction range) \(s\). Crucially, reproduction is possible for organisms with trait value \(y\) at position \(x\) only if \(y > ax\), where the slope \(a\) represents the steepness of an environmental gradient that limits the species range. This defines a diagonal ‘ramp’ (red line), below which organisms are not sufficiently adapted to reproduce. As the population adapts, it ‘climbs’ the ramp at an asymptotic rate \(v_r\) that can be decomposed into the invasion rate \(v_i\) and the adaptation rate \(v_a\).
change as a diffusion process (vertically), with diffusion constant \(D_y\).

Organisms reproduce asexually by division; the daughter is placed at the exact same coordinates as the mother. To implement local resource competition, the stochastic rate of reproduction \(g_i\) of an organism \(i\) depends logistically on the local population density. This local density is measured by a kernel density estimate, that is, by the convolution of the population distribution with a compact interaction kernel \(O(x)\). In the simulations below, we choose \(O(x)\) to be Gaussian with standard deviation (‘interaction range’) \(s\). The effect of the environmental gradient is included by stipulating that individuals at position \(x\) can reproduce only if they have a trait value \(y > ax\). In the diagram in figure 1, this defines a diagonal ‘ramp’ (red line) dividing the organisms that are sufficiently adapted to reproduce from those that are not; its slope \(a\) reflects the steepness of the environmental gradient. To summarize, the rate of reproduction of individual \(i\) can be expressed as

\[
g_i = r \left(1 - \sum_{j \neq i} \frac{O(x_i - x_j)}{k} \right) H(y_i - ax_i). \tag{1}
\]

where \(r\) is the maximal rate of reproduction, \(k\) is the carrying capacity per unit of length, and \(H(y)\) is the Heaviside step function. We stress that, through equation (1), the growth rate of each organism depends on its own coordinates, as well as on the relative position of every other individual in the population.

Lastly, organisms die at a fixed stochastic rate \(d\).

Below we demonstrate that an initially compact population distribution starts to ‘climb’ the ramp, simultaneously invading new territory and adapting to the gradient. In time, the population front obtains a fixed average velocity in the phenotype–position space, denoted \(v\). The component of this velocity in the \(x\)-direction \(v_x\) we identify as the rate of invasion, and the component in the \(y\)-direction \(v_y\) as the rate of adaptation, as illustrated in figure 1 with black arrows.

\[
\frac{\partial v}{\partial t} = \nabla^2 v + \left\{ \nu \left(1 - \int_{-\infty}^{\infty} \Omega(x - \chi|\sigma|\nu) \mathrm{d}x \right) \right\} \delta v
\]

\[
\text{for } \psi > \alpha \chi,
\]

\[
\text{for } \psi < \alpha \chi.
\]

3. Results

In this section, we first study the mean-field limit of the ramp model and calculate the rate of adaptation in this limit analytically. Next, we compare the mean-field results to discrete-stochastic simulations.

### Table 1. Dimensionless variables and parameters

| Original quantity | Dimensionless alternative |
|-------------------|---------------------------|
| \(t\)             | \(t \equiv (r - d)t\)     |
| \(x\)             | \(x \equiv \sqrt{r/dD_x}\) |
| \(y\)             | \(y \equiv \sqrt{r/dD_y}\) |
| \(d\)             | \(d \equiv d/r - d\)     |
| \(r\)             | \(r = 1 + \delta\)       |
| \(D_x\)           | \(D_y = 1\)              |
| \(a\)             | \(a \equiv \sqrt{D_x/D_y}\) |
| \(k\)             | \(k \equiv D_y/\sqrt{d - k}\) |
| \(s\)             | \(s \equiv \sqrt{d/D_\sigma}\) |
| \(n(x, y)\)       | \(n(x, y) \equiv \sqrt{D_y/(r - d)^\gamma} n(x, y)/k\) |

#### 3.1. The mean-field model

If the population density is large, the following mean-field description would seem appropriate:

\[
\frac{\partial n}{\partial t} = D_x \frac{\partial^2 n}{\partial x^2} + D_y \frac{\partial^2 n}{\partial y^2} + \tau n \left(1 - \int_{-\infty}^{\infty} \frac{O(x - x'|\psi|)n}{k} \mathrm{d}x \mathrm{d}x'\right) \times H(y - ax) - d n.
\]

(Remember that \(d\) is the death rate; \(dn\) is not a differential.) Here \(n(x, y)\) is the (now continuous) density of organisms with trait value \(y\) and position \(x\). For convenience, we replace all variables and parameters by dimensionless ones denoted by corresponding Greek characters. Effectively, we scale time, position and phenotype such that the net growth rate above the ramp \((r - d)\) and both diffusion constants \(D_x\) and \(D_y\) are set to unity; see table 1. Thus, equation (2) simplifies to

\[
\frac{\partial \psi}{\partial \tau} = \nabla^2 \psi + \left\{ \nu \left(1 - \int_{-\infty}^{\infty} \Omega(\chi - \chi'|\sigma|\nu) d\chi d\chi'\right) \right\} \delta \psi
\]

\[
\text{for } \psi > \alpha \chi,
\]

\[
\text{for } \psi < \alpha \chi.
\]

This reveals that the behavior of the mean-field model depends on just three lumped, dimensionless parameters: the slope of the ramp \(\alpha\), the death rate \(\delta\), and the interaction range \(\sigma\) (the standard deviation of the interaction kernel).

For given parameters, equation (3) can be integrated numerically (see the Methods section). As an example,
After some time, the leading edge of the pulse obtains a fixed shape (red curve) with a constant invasion rate. The asymptotic pulse shape in the two-dimensional phenotype–position space of figure 1 is displayed in figure 2(b).

We wish to derive the asymptotic velocity of the front analytically. To do so, we exploit that the pulse is a pulled front [26], similar to the famous Fisher–Kolmogorov–Petrovsky–Piscounov (F–KPP) front [27, 28]. The asymptotic velocity of pulled fronts developing from a compact initial distribution is determined entirely by the propagation velocity of small perturbations of the linearly unstable state $\nu = 0$ [26]. Therefore, it can be obtained from linearized equations. Linearizing equation (3) around $\nu = 0$ results in

$$\frac{\partial \nu}{\partial t} = \nabla^2 \nu + \begin{cases} \nu & \text{for } \psi \geq \alpha \chi \\ -\delta \nu & \text{for } \psi < \alpha \chi. \end{cases} \tag{4}$$

This equation describes a diffusing population in two dimensions of individuals that reproduce above the ramp at rate 1 and die below the ramp at a rate $\delta$. Two important conclusions follow. First, the asymptotic velocity of the pulse does not depend on the interaction kernel chosen. Second, we note that a rotation of the coordinate frame leaves equation (4) invariant except for a change in $\alpha$; conversely, that means that the solutions of equation (4) for different values of $\alpha$ are identical up to a rotation. This directly proves that the velocity of the front parallel to the ramp $v_\parallel$ (see figure 1) is independent of $\alpha$. The rate of adaptation $v_\perp$ then follows by projecting $v_\parallel$ onto the $\psi$-axis. The result is

$$v_\perp = \sqrt{\frac{\alpha^2}{1 + \alpha^2} v_\parallel}. \tag{5}$$

This analysis does not exclude that $v_\parallel$ might depend on $\delta$. However, we derive in the appendix that $v_\parallel = 2$ (as expected for F–KPP-type fronts), independent of $\delta$.

In figure 3 we verify equation (5) by measuring the adaptation rate using numerical solutions of the full nonlinear equation (equation (3)), for a range of values $\alpha$ and $\delta$ (see the Methods section below for technical details). Also included are results for the limit $\delta \to \infty$, where the ramp effectively becomes an absorbing boundary. All results are in excellent agreement with the analytical prediction (solid red line). The rates calculated for different values of $\delta$ (symbols) fall right on top of each other, confirming that the adaptation rate is independent of $\delta$.

In terms of the original dimensional parameters, the adaptation rate reads:

$$v_\perp = \sqrt{\frac{4(r - d)D_y D_x a^2}{D_y + D_x a^2}}. \tag{6}$$

It is useful to distinguish two regimes in this result. If $a^2 \ll D_x/D_y$ (that is, $\alpha \ll 1$), the rate of adaptation approaches $v_\perp = 2\sqrt{(r - d)D_x a}$ expected for an F–KPP front. We call this the dispersion-limited regime. If, instead, $a^2 \gg D_x/D_y$ (that is, $\alpha \gg 1$), then $v_\perp \approx 2\sqrt{(r - d)D_y}$, and the adaptation is directly limited by the mutation rate. In this regime the rate becomes independent of both $D_x$ and $a$.

Remarkably, equations (5) and (6) show that the adaptation rate increases strictly monotonically with the gradient slope. This is unexpected: intuitively, adaptation against a very steep gradient should be slow, because in this limit only those rare mutations that cause an unusually large increase in the quantitative trait allow for a significant range expansion. And yet, in the mean-field limit a larger slope always yields faster adaptation, even in the mutation–limited regime of $\alpha \gg 1$. To test whether this is an artifact of the mean-field limit, we now compare the above results to
stochastic simulations of discrete organisms at a finite carrying capacity.

3.2. Discrete-stochastic simulations

Exact simulations of the stochastic ramp model are challenging. Through equation (1), the reproduction rate of each individual depends on the coordinates of all organisms. Because each organism diffuses in continuous time and space, the reproduction rates of all individuals fluctuate rapidly and in unpredictable ways, which disqualifies the standard Gillespie algorithm [29]. We therefore used an event-driven simulation scheme that is both exact and remarkably efficient. We call it the WHINE algorithm. It can deal with reaction rates that depend on time or on any number of other variables, and could therefore be applied to a large variety of stochastic systems studied within and outside of population biology.

The crucial idea is to extend the stochastic model with a ‘dummy’ process: an additional process that has no consequences for the behavior of the system. This process is given a reaction rate such that the sum of all rates of reactions affecting each given individual (in this case growth, death, and ‘dummy’) becomes constant—that is, independent of time and other coordinates, and the same for all individuals. Consequently, the next-reaction time and the 2D diffusion (in continuous space and time) of the organisms become statistically independent and can therefore be sampled from their exact continuous probability distributions, without a need to discretize space, quantitative trait values, or time. The exact algorithm is detailed in the Methods section below.

We performed simulations for a large number of parameter combinations. Figure 4(a) is a snapshots from one of these, showing a dense cloud of nearly 30 000 organisms steadily climbing the ramp ($\kappa = 500$). Figure 4(b) displays a kernel density estimate of the exact same configuration. In this simulation, the same slope and death rate ($\alpha$ and $\delta$) were used as in the mean-field snapshot of figure 2(b). Comparison between figures 4(b) and 2(b) confirms the expected similarity at high carrying capacities between the pulse of the mean-field model and that of the stochastic model.

We have used such simulations to measure the adaptation rate for a broad range of parameter values. Some results are shown in figure 5(a), were the adaptation rate is plotted against the slope $\alpha$ for many values of the carrying capacity $\kappa$. In the dispersion-limited regime ($\alpha \ll 1$), we see that the results are qualitatively similar to the mean-field result (red line), even though significant quantitative deviations are found even at large carrying capacities (this was to be expected; see Discussion). In the mutation-limited regime ($\alpha \gg 1$) however, the mean-field result deviates qualitatively from the stochastic simulations: in the mean-field, the adaptation rate saturates at $v_\phi = 2$, whereas the stochastic simulations show a rapid decrease in $v_\phi$ with increasing $\alpha$. For any fixed slope $\alpha$, the adaptation rate approaches the mean-field prediction as the carrying capacity is increased. Nevertheless, for each fixed value of the carrying capacity, the adaptation rate is non-monotonic as a function of $\alpha$, in striking contrast with the mean-field result. At small carrying capacities, the slope providing the maximal rate of adaptation is near $\alpha = 1$; at larger carrying capacities the optimum slowly creeps to the right. In all cases, however, the adaptation rate is already close to optimal at $\alpha = 1$.

In order to understand the observed discrepancy, we study the convergence of the simulation results to the mean-field prediction. We consider the deviation from the mean-field prediction $\Delta \equiv v_{\phi,\text{MF}} - v_{\phi,\text{S}}$, where we denoted the adaptation rate in the stochastic model by $v_{\phi,\text{S}}$ and the mean-field expectation by $v_{\phi,\text{MF}}$. Figure S1 shows the relative deviation $\Delta / v_{\phi,\text{MF}}$ for three values of the slope $\alpha$ as a function of the carrying capacity $\kappa$. In the simulated regime of $\kappa = 1$ to 1000, we find that the convergence to the mean-field result is well described by a power law $\Delta(\kappa) \equiv Q \kappa^{-\mu}$, where the prefactor $Q$ and exponent $\mu$ depend on the parameters other than $\kappa$ (to wit, $\alpha$, $\delta$ and $\sigma$).

Figure 6(a) plots exponent $\mu$ as a function of the slope $\alpha$, for two values of interaction range $r$. In the dispersion-limited regime, $\mu$ is roughly constant, but in the mutation-limited regime the scaling is well described by $\mu \propto \alpha^{-1/2}$. This does not depend on a particular choice of $\sigma$ or $\delta$: While the convergence becomes slower over-all if $\sigma$ is increased, the scaling is not affected (figure 6(a)), and if the death rate $\delta$ is reduced, the convergence becomes slower in the
mutation-limited regime, but the same square-root behavior is found for large enough $\alpha$ (see figure S2(a)). We conclude that, in the mutation-limited regime, the convergence to the mean-field limit slows down strongly as $\alpha$ increases. This non-uniform convergence explains how a non-monotonic dependence of the adaptation rate on the slope at any finite population density can, in fact, be consistent with a monotonic dependence in the mean-field limit.

Figure 6(b) plots the prefactor $Q$ as a function of $\alpha$, again for two values of $\sigma$. It appears that $Q$ is proportional with $\alpha$ when $\alpha \lesssim 1$ but converges to $2 - 1/\alpha$ when $\alpha \gtrsim 1$. This is remarkably insensitive to changes in $\sigma$ (figure 6(b)) or $\delta$ (figure S2(b)).

As a result of these scalings, we can reasonably summarize our simulation results as follows:

\[
\begin{align*}
Q \approx \begin{cases} 
   c_1 \alpha \kappa^{-c_2} & \text{for } \alpha < 1 \\
   (2 - 1/\alpha) \kappa^{-(2/\alpha) c_2} & \text{for } \alpha > 1.
\end{cases}
\end{align*}
\]

Here, the coefficients $c_1$, $c_2$, and $c_3$ depend on $\sigma$ and $\delta$, even though $c_1$ and $c_2$ appear to be very insensitive to changes in these parameters (as seen in figures 6(b) and S2(b)). Figure S3 demonstrates the quality of equation (7) as an approximation to the data of figure 5. Incidentally, we note that an expansion of equation (7) in the limit of large $\alpha$ shows that $v_\psi$ scales as $\propto \alpha^{-1/2}$; this behavior can indeed be seen in figures 5 and S3 for sufficiently large $\alpha$ and $\kappa$.

One rather drastic assumption of the ramp model is that the rate of reproduction switches discontinuously from zero to a finite value as an organism moves from below to above the ramp. To test whether this discontinuity crucially affects the above results, we
performed additional simulations in which the step function $H(y)$ of equation (2b) is replaced by a sigmoidal (logistic) function $L(y)$ defined as:

$$L(y) = \frac{1}{1 + e^{-y/w}}. \quad (8)$$

Here $w$ is the width of the transition region, arbitrarily defined as the interval where $L(y)$ increases from 12% to 88% of its maximal value. As shown in figure S4, the results are essentially the same as for the step function, even if the transition region is wide ($w = 16$).

4. Discussion and conclusions

We have presented results from stochastic simulations and a corresponding mean-field model. These results identified two qualitatively different regimes: a dispersion-limited regime ($a^2 \ll D_y/D_x$) and a mutation-limited regime ($a^2 \gg D_y/D_x$). In the dispersion-limited regime, the invasion rate $v_\alpha$ is not affected by the gradient. As a consequence, the adaptation rate $v_\alpha = a_\alpha v_\alpha$ is linear in $a$. In this regime, the velocity of the pulse is insensitive to the mutation rate $D_\alpha$, which is sufficiently high to provide the phenotypic variation required for the population to adapt sufficiently fast. The dispersion rate $D_\alpha$ however, does directly limit the invasion rate, which is $v_\alpha = 2\sqrt{(r - d)D_\alpha}$ in the mean field—the rate expected for an F–KPP front—with quantitative corrections at finite carrying capacities. In the mutation-limited regime the situation becomes more complex. The mean-field model predicts an adaptation rate $v_\alpha = 2\sqrt{(r - d)D_\alpha}$ independent of $a$ or $D_\alpha$. This prediction, however, proves highly misleading: the adaptation rate in fact decreases with $a$ at any finite carrying capacity. This discrepancy is caused by a non-uniform convergence to the mean-field limit for different values of $a$. At any finite carrying capacity, then, a non-monotonic dependence of the adaptation rate on the slope is found. The optimal gradient for adaptation (all else equal) occurs when these two regimes meet, that is, at $a \approx (D_y/D_x)^{1/2}$. The effect of stochasticity on the propagation of pulled fronts has been studied extensively, in particular in the context of (variants of) the F–KPP front in one and two dimensions (e.g., [26, 30–37]). This work has shown that, in such fronts, the convergence to the mean-field limit with increasing population density is anomalously slow, so that significant quantitative differences can be found between mean-field and finite-density stochastic models even at large population densities. The result above show, however, that the consequences are not necessarily limited to quantitatively wrong conclusions.

Heuristically, the large difference between the mean-field result and the stochastic simulations in the mutation-limited regime can be understood by considering the effects of scaling the spatial dimension (the $x$ axis). In both models, scaling space must leave the rate of adaptation (i.e., the propagation in the $y$ direction) unchanged, but does change the slope $a$. In the mean field, the velocity of the pulse is determined by the behavior far ahead of the bulk, where the linearized equation reigns and the nonlinear interaction kernel can be ignored. Then, the effect of scaling the $x$ axis by a factor $\epsilon$ is only to replace the slope by $a/\epsilon$ and the dispersion rate by $D_\epsilon \epsilon^{-2}$. The adaptation rate at a given $a$ and $D_\epsilon$ must therefore equal the adaptation rate at slope $a$ and dispersion rate $D_\epsilon a^2$. That is, increasing $a$ has the same effect on the adaptation rate as increasing $D_\epsilon$; it leads to an increase in the adaptation rate that saturates when dispersion is no longer limiting. In the stochastic model, however, the interaction kernel cannot be ignored. Scaling of the $x$ axis now has two
additional effects: it simultaneously decreases the carrying capacity \( k \) and increases the standard deviation \( s \) of the interaction kernel by a factor \( \epsilon \). Both of these effects tend to reduce the rate of adaptation and move the system away from the mean-field limit towards a strong-noise regime. In addition, as \( D_k \) is increased, individuals at the leading edge of the pulse rapidly diffuse away, into the regime \( x > \gamma/a \) where they cannot reproduce; this further reduces the population density at the front. Unfortunately, we have not yet been able to find a mathematical derivation of the precise scalings found in figure 6; we hope that future work will fill this void, perhaps based on the scalings found for the strong-noise regime of the F–KPP wave \([35, 37]\).

In the past few years, two models have been published that seem very similar to the ramp model on the surface, but differ crucially from it upon closer inspection \([14, 15]\).

The first model is the so-called staircase model \([14]\). The assumptions of this model are closely analogous to the ramp model. The main difference is that in the staircase model both space and phenotype are discretized. Bacteria can stochastically migrate between neighboring ’compartments’, but in order to reproduce in compartment \( i \) a bacterium must accumulate \( i - 1 \) mutations. Consequently, instead of the continuous ramp of the ramp model, (red line in figure 1) it features a discrete ’staircase’ allowing adaptation and invasion to proceed step by step. This seemingly minor difference makes a direct comparison between the models highly problematic. First and foremost, in the staircase model the steepness of the gradient is not a parameter; because the staircase is directly linked to the discretization of genotype and space, the steepness cannot be varied independently. Consequently, the main results of the ramp model presented above have no direct counterpart in the staircase model. Second, it cannot be varied independently. Consequently, the extinction limit of the staircase model; that is, as the limit in line with the ramp model. In the mutation-limited regime, the adaptation rate to increase linearly with the gradient, respectively. Greulich \( et \ al \) do not discuss the dispersion-limited regime, but we expect the adaptation rate to increase linearly with the gradient, in line with the ramp model. In the mutation-limited regime, the adaptation rate is limited by the arrival and fixation of consecutive mutations. Under the assumption that the dispersion rate is not too high, the authors derive a scaling of \( \nu \propto \mu k/a \), where \( \mu \) is the mutation probability per cell division. This differs from the scaling of \( a^{-1/2} \) found in the ramp model. This seems to be related to another difference between the models: the Greulich model assumes that the growth rate of a given genotype reduces gradually with the drug concentration, whereas the ramp model assumes a step function. The gradual decrease results in a spatial domain within the habitat of each given genotype in which the next mutation is under positive selection; new mutations are most likely to establish in this domain. Importantly, the width of this domain decreases linearly with \( a \), which produces the observed scaling of the dispersion rate. We conclude that, in the Greulich model, the rate of adaptation is non-monotonic as a function of the slope, as in the ramp model—but the scaling is qualitatively different due to several differences in the modeling choices and the parameter regimes considered.

We have presented a model that was highly idealized; our aim was to only include the minimal ingredients required for adaptation of a quantitative trait to
an environmental gradient. In future work, more complex and realistic models can be studied; the simple model presented here can provide a null expectation for such extensions. We do expect that several basic results will hold up in more realistic versions. First, we expect that every reasonable model will feature dispersion- and mutation-limited regimes for shallow and steep gradients, respectively. In the dispersion-limited regime, the rate of adaptation should increase linearly with the magnitude of the gradient; in the mutation-limited regime, a decreasing function should be found, the scaling of which may however depend on details of the fitness landscape and genotype-phenotype mapping assumed. Regardless of those details, the optimal adaptation rate should occur in the transition region between these two regimes. Generally, we expect this transition to occur at a modest slope, that is, a slope that hampers the invasion rate mildly. Such qualitative predictions may in principle be tested experimentally, e.g. using bacterial populations in microfluidic devices or on agar plates in which an environmental gradient is set up. We hope that the ramp model, together with its predecessors, will prove useful in interpreting such experiments.

5. Methods

5.1. Numerical solutions of the mean-field equations

To numerically solve the mean-field equation (equation (3)) we used the Method Of Lines: we defined a 2D grid, approximated the PDE using finite differences, and integrated the resulting set of ODEs in Matlab using the ode113() function, which is a variable-step, variable-order Adams–Bashforth–Moulton PECE solver.

For cases with $\alpha \geq 1$, we used a straightforward rectangular grid. For cases with $\alpha < 1$, we instead used a grid that was sheared vertically such as to align the grid points with the ramp. In both cases, the Laplacian at each grid point was evaluated using the $5 \times 5$ array of grid points surrounding it; the proper finite-difference coefficients were calculated specifically for each $\alpha$.

The simulation box itself was shaped as a parallelogram, with sides parallel the vertical (phenotype) axis and to the ramp. This box was shifted periodically (parallel to the ramp) in order to keep the pulse front near the middle of the box. On the left-hand side of the box, reflecting (homogeneous Neuman) boundary conditions were imposed; this introduces small artifacts near this boundary, but does not affect the front of the pulse. The other three sides were absorbing (homogeneous Dirichlet) boundaries.

5.2. Stochastic simulations: the WHINE algorithm

As explained above, we performed individual-based stochastic simulations using an event-driven algorithm that treats the complex density- and position-dependent growth rates exactly. The main challenge for our simulations is that the organisms move diffusively in continuous space, and therefore the reaction rates (of reproduction and death) cannot be assumed constant between reactions, which rules out (variants of) the Gillespie algorithm. The key idea is to extend the model with a dummy reaction, with a reaction rate chosen such that the diffusive path of each organism (in two dimensions) becomes statistically independent of the next-reaction time. This idea can be used to simulate a diverse class of stochastic systems; here we describe its implementation for the current model.

For vividness, we may imagine that our organisms tend to whine, particularly under conditions unfavorable for reproduction; this is the dummy reaction. The rate of whining $w_i$ of individual $i$ is chosen to be

$$w_i = r - g_i.$$  \hfill (9)

(Remember that $r$ and $g_i$ are the maximal rate of reproduction, at low densities, and the actual rate of reproduction of individual $i$; see equation (1).) With this choice, the rates of all discrete reactions that can affect any given organism (growth, death, whining) add up to a constant $R$:

$$g_i + d + w_i = r + d \equiv R.$$  \hfill (10)

Note that, by construction, $R$ is the same for all organisms and independent of density, time, or any coordinates. The algorithm then repeats the following steps as often as desired:

1. **Choose next-reaction time.** Denote the total number of organisms in the system at time $t$ by $N(t)$. Then the sum of all reaction rates in the whole system is $N(t)R$, which is not affected by the 2D diffusion of the individuals and therefore does not change between reactions. Given the time of the previous event $t_p$, this allows a next-event time $t_{n}$ to be sampled as in the standard Gillespie algorithm, by drawing a waiting time $\Delta t \equiv t_n - t_p$ from the exponential waiting-time probability distribution

$$P(\Delta t) \propto \exp[-N(t_p)R\Delta t].$$

2. **Update coordinates.** By construction, the probability distribution of the waiting time $\Delta t$ does not depend on the coordinates of the individuals. Conversely, this means that a particular choice of $\Delta t$ does not bias the diffusive trajectories of the individuals during this interval. Therefore, between reactions the individuals move by pure, unbiased diffusion, and their coordinates at $t_n$ can be chosen by adding Gaussian random numbers to the coordinates at time $t_p$ (with variances $2D_x\Delta t$ and $2D_y\Delta t$ for $x$ and $y$, respectively).

3. **Choose the affected individual.** We now know **when** the next reaction will take place, and **where**
all organisms are at that time; we proceed to choose **whom** the event will affect. The total event rate \( R \) is the same for all individuals; therefore, the affected individual \( a \) may be chosen uniformly at random.

4. **Choose the reaction.** Next, we decide which reaction occurs: reproduction, death, or whining. Now that the next-event time \( t_n \), the coordinates of all individuals at that time, and the affected individual \( a \) are known, the probabilities of the event types are simply proportional to their rates at \( t_n \). To choose the event the ordinary ‘roulette wheel’ algorithm can be used (as in the standard Gillespie scheme) after the reproduction rate of \( a \) has been calculated explicitly using equation (1).

(We note that the expensive calculation of the reproduction rate can be avoided in many time steps: because the probability that a death event occurs is always \( d/R \), the reaction rate needs to be calculated only if it is decided that \( a \) does not die.)

5. **Execute the reaction.** If it turns out that the organism merely whines, nothing needs to be done; otherwise, the chosen reaction is executed.

The introduction of the dummy reaction obviously introduces overhead. In some of our simulations, 5 out of 6 events are dummy events, effectively resulting in a sixfold reduced time step. That said, the algorithm is in many ways remarkably efficient. As we emphasized, the reaction rates affecting a given organism depend on its own coordinates—whether it is above or below the ramp—and on the relative position of all other organisms. All these coordinates fluctuate due to the continuous-time diffusion of each organism. Yet, to deal with this exactly, at each time step of WHINE algorithm the reaction rates for only one organism (at most) need to be calculated explicitly.

We note that dummy reactions have been used before to efficiently simulate stochastic systems—e.g., see [15]. In these applications, however, spatial coordinates were discretized, with individuals performing a discrete random walk rather than diffusion in continuous space and time (a Wiener process). As a result, in these systems all reaction rates are constant between reactions. The above algorithm is remarkable in that it properly treats systems in which the reaction rates fluctuate unpredictably in continuous time.

During the simulations, the population expands as it invades new territory. To cap the population size, the simulation box we used is the half-plane bounded on the left by a reflective boundary \( x = b(t) \) that moves along with the leading edge of the pulse; there is no need to limit the simulation box in the \( y \) direction or on the right-hand side. As the boundary moves, organisms are ‘lost’ on the left. The boundary \( b(t) \) is frequently updated such that the pulse has a simulated range in the \( x \)-direction of \( \approx 60 \) dimensionless units (as in figure 4). This is enough: extensive test simulations with a range of \( \approx 100 \) gave nearly identical results.

The simulations were performed using custom code written in Matlab (prototype and first results) and Fortran. In the latter case we used the ziggurat method [38] (translated to Fortran by Alan Miller), to efficiently draw Gaussian random numbers (the algorithm needs many).

**Acknowledgments**

Part of this work was performed at the Department of Bionanoscience of the TU Delft; I am grateful for their valuable support. In particular, I thank Cees Dekker for his encouragement and advice. This work was also supported by the NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) (VENI 680-47-419).

**Appendix. Derivation of the velocity of the pulse**

We here derive that the velocity of the pulse parallel to the ramp \( v_l \) equals 2 (in terms of the dimensionless quantities of table 1) independent of the death rate \( \delta \). To obtain \( v_l \), we analyze the linearized equation (equation (4)) with initial condition

\[
\nu(x, \psi, \tau = 0) = D_1(\chi)D_1(\psi),
\]

where \( D_1(x) \) is the Dirac delta function.

Because we established that \( v_l \) is independent of \( \alpha \) (see the Results) we can choose \( \alpha = 0 \) without loss of generality. We then exploit that equation (4) is separable by making the Ansatz

\[
\nu(x, \psi, \tau) = e^{i\psi} f(x, \tau) g(\psi, \tau).
\]

(The prefactor \( \exp(\tau) \) is arbitrary but will simplify the equations below.) Our task is to derive the functions \( f \) and \( g \). We note, however, that we do not need the full solution of \( f(\psi, \tau) \): we are satisfied with an expression for \( g^l(\tau) \equiv \int g(\psi, \tau) d\psi \), because the velocity of the front can be obtained from the (marginal) population distribution in real space, \( v^l(\chi, \tau) \), defined as

\[
v^l(\chi, \tau) \equiv \int v(x, \psi, \tau) d\psi = e^{i\chi} g^l(\tau) f(x, \tau).
\]

We combine equation (A.2) with equation (4) of the main text to obtain separate differential equations for \( f \) and \( g \). We find that \( f(x, \tau) \) simply obeys the diffusion equation in one dimension, with the standard solution
The equation for \( g(\psi, \tau) \) reads:

\[
\frac{\partial g}{\partial \tau} = \frac{\partial^2 g}{\partial \psi^2} - (1 + \delta)gH(-\psi),
\]

with initial condition

\[
g(\psi, 0) = \text{Di}(\psi).
\]

(Again, Di(x) is the Dirac delta function.) This describes an ensemble of particles that diffuse in one dimension and are degraded in the domain \( \psi \leq 0 \) at a rate \( 1 + \delta \), after initially being placed right at the boundary of that domain.

To make progress, we perform a Laplace transformation. We denote the Laplace transform of \( g(\chi, \tau) \) by \( G(\chi, s) \), and use that the Laplace transform of the time-derivative of \( g(\chi, \tau) \) equals \( sG(\chi, s) \) \(-g(\chi, 0)\). Thus equation (A.5) is converted to a second-order differential equation that can be solved exactly. To deal with the delta peak of the initial condition, we temporarily replace it by a rectangular peak with a finite width \( \epsilon \) and height \( 1/\epsilon \), solve the differential equation, and then take the limit of \( \epsilon \to 0 \). We thus obtain:

\[
G(\psi, s) = \begin{cases} 
\frac{e^{-(1+\delta+s)\psi}}{\sqrt{1+\delta+s+\sqrt{s}}} & \text{for } (\psi \leq 0), \\
\frac{e^{-\sqrt{s} \psi}}{\sqrt{1+\delta+s+\sqrt{s}}} & \text{for } (\psi > 0).
\end{cases}
\]

The inverse Laplace transform of equation (A.7) is hard to obtain. Instead, we integrate A.7 over \( \psi \) to obtain the Laplace transform \( G^I(s) \) of \( g^I(\tau) \), which yields the remarkably simple result

\[
G^I(s) = [s(1+\delta+s)]^{-\frac{1}{2}}.
\]

The inverse Laplace transform of this result can be found:

\[
g^I(\tau) = I_0[(1+\delta)\tau/2]e^{-(1+\delta)\tau/2},
\]

where \( I_0(\chi) \) is the modified Bessel function of the first kind.

Since we are after the asymptotic velocity of the front, we look for the large-\( \tau \) behavior of \( g^I(\tau) \). This behavior is determined by the low-\( s \) behavior of the Laplace transform, where \( G^I(s) \to [(1+\delta)s]^{-\frac{1}{2}} \). Consequently, in the limit or large \( \tau \), we find

\[
g^I(\tau) \propto [\pi (1 + \delta) \tau]^{-\frac{1}{2}},
\]

and

\[
\nu^I(\psi, \tau) \propto \exp[\tau - \psi^2/(4\tau)]/2\pi\sqrt{1+\delta+\tau}.
\]

For any \( \delta \), this equation has the same functional form as the equivalent result for the F-KPP front in two dimensions, which is known to have the asymptotic velocity of \( v_F = 2 \). Therefore, the asymptotic velocity of the pulse is equal to that of the 2D F-KPP equation: \( v_{\parallel} = v_F = 2 \), independent of \( \delta \).

References

[1] Felsenstein J 1976 The theoretical population genetics of variable selection and migration Annu. Rev. Genet. 10 233–80
[2] Gaston K J 2003 The Structure and Dynamics of Geographic Ranges (Oxford: Oxford University Press)
[3] Rice SH 2004 Evolutionary Theory: Mathematical and Conceptual Foundations (Sunderland, MA: Sinauer Associates)
[4] Kirpatrick M and Barton N H 1997 Evolution of a species’ range Am. Nat. 150 1–23
[5] Bridle J R and Vines T H 2007 Limits to evolution at range margins: when and why does adaptation fail? Trends Ecol. Evol. 22 140–7
[6] Gaston K J 2009 Geographic range limits: achieving synthesis Proc. R. Soc. B: Biol. Sci. 276 1395–406 (1661)
[7] Sexton J P, McIntyre P J, Angert A L and Rice K J 2009 Evolution and ecology of species range limits Annu. Rev. Ecol. Evol. Syst. 40 415–36
[8] Polechov J, Barton N and Marion G 2009 Species’ range: adaptation in space and time Am. Nat. 174 E186–204
[9] Holt R D and Barfield M 2011 Theoretical perspectives on the statics and dynamics of species borders in patchy environments Am. Nat. 178 56–25
[10] Baquero F and Negri M-C 1997 Challenges: selective compartments for resistant microorganisms in antibiotic gradients Biosys. 19 731–6
[11] Kepler T B and Perelson A S 1998 Drug concentration heterogeneity facilitates the evolution of drug resistance Proc. Natl Acad. Sci. 95 11514–9
[12] Hermsen R and Hwa T 2010 Sources and sinks: a Stochastic model of evolution in heterogeneous environments Phys. Rev. Lett. 105 (24).
[13] MacLean R C, Hall A R, Perron G G and Buckling A 2010 The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts Nat. Rev. Genet. 11 405–14
[14] Hermsen R, Deris I B and Hwa T 2012 On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient Proc. Natl Acad. Sci. USA 109 10775–80
[15] Greulich P, Wacław B and Allen R J 2012 Mutational pathway determines whether drug gradients accelerate evolution of drug-resistant cells Phys. Rev. Lett. 109 088101
[16] Hallatschek O 2012 Bacteria evolve to go against the grain Physics 5 93
[17] Elliott A M, Berning S E, Isemann M D and Peloquin C A 1995 Failure of drug penetration and acquisition of drug resistance in chronic tuberculous empyema Tuberc. Lung Dis. 76 663–7
[18] Lipsitch M and Levin B R 1998 Population dynamics of tuberculosis treatment: mathematical models of the roles of non-compliance and bacterial heterogeneity in the evolution of drug resistance Int. J. Tuber. Lung Dis. 2 187–99
[19] Perron G G, Gonzalez A and Buckling A 2007 Source–sink dynamics shape the evolution of antibiotic resistance and its pleiotropic fitness cost Proc. R. Soc. B: Biol. Sci. 274 2351–6 (1623)
[20] Perron G G, Quessy S and Bell G G 2008 A reservoir of drug-resistant pathogenic bacteria in asymptomatic hosts PloS One 3 e3749
[21] Chattopadhyay S, Feldgarden M, Weissman S I, Dykhuizen D E, van Belle G and Sokurenko E V 2007 Haplotyp diversity in source–sink dynamics of escherichia coli urovirulence J. Mol. Evol. 64 204–14
[22] Zhang Q, Lambert G, Liao D, Kim H, Robin K, Tung C-K, Pourmand N and Austin R H 2011 Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments Science 333 1764–7
[23] Zhang Q, Robin K, Liao D, Lambert G and Austin R H 2011 The goldilocks principle and antibiotic resistance in bacteria Mol. Pharm. 8 2063–8
[24] Baquero F, Negri M-C, Morosini M-I and Blázquez J 1998 Antibiotic-selective environments Clin. Infect. Dis. 27 S5–11
[25] Hartl D L and Clark A G 2006 Principles of Population Genetics, Fourth Edition 4th edn (Sunderland, MA: Sinauer Associates)
[26] van Saarloos W 2003 Front propagation into unstable states Phys. Rep. 386 29–222
[27] Fisher R A 1937 The wave of advance of advantageous genes Ann. Eugen. 7 355–69
[28] Kolmogoroff A, Petrovsky I and Piscounoff N 1937 Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem Bull. L’universite Detat Moscow Ser Int Sect. A I
[29] Gillespie D T 2008 Simulation methods in systems biology Formal Methods for Computational Systems Biology (Lecture Notes in Computer Science vol 5016) (Berlin: Springer) pp 125–67
[30] Mueller C and Tribe R 1995 Stochastic pde’s arising from the long range contact and long range voter processes Probab. Theory Relat. Fields 102 519–45
[31] Brunet E and Derrida B 1997 Shift in the velocity of a front due to a cutoff Phys. Rev. E 56 2597–604
[32] Kessler D A, Ner Z and Sander L M 1998 Front propagation: precursors, cutoffs, and structural stability Phys. Rev. E 58 107
[33] Moro E 2001 Internal fluctuations effects on fisher waves Phys. Rev. Lett. 87
[34] Brunet E and Derrida B 2001 Effect of microscopic noise on front propagation J. Stat. Phys. 103 269–82
[35] Doering C R, Mueller C and Smereka P 2003 Interacting particles, the stochastic Fisher–Kolmogorov–Petrovsky–Piscounov equation, and duality Phys. Stat. Mech. Appl. 325 243–59
[36] Cohen E, Kessler D A and Levine H 2005 Front propagation up a reaction rate gradient Phys. Rev. E 72
[37] Hallatschek O and Korolev K 2009 Fisher waves in the strong noise limit Phys. Rev. Lett. 103
[38] Marsaglia G et al 2000 The ziggurat method for generating random variables J. Stat. Softw. 5 1–7