Published in final edited form as:
Am Nat. 2011 October ; 178(4): 538–552. doi:10.1086/661897.

A Quantitative Test of Population Genetics Using Spatiogenetic Patterns in Bacterial Colonies

Kirill S. Korolev[1,*], **João B. Xavier**[2], **David R. Nelson**[3], and **Kevin R. Foster**[4,†]

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

[2]Program in Computational Biology, Memorial Sloan-Kettering Cancer Center, New York, New York 10065

[3]Department of Physics and FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts 02138

[4]FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts 02138; Department of Zoology, Oxford University, South Parks Road, Oxford, OX1 3PS, United Kingdom; and Oxford Centre for Integrative Systems Biology, Oxford University, South Parks Road, Oxford, OX1 3QU, United Kingdom

Abstract

It is widely accepted that population-genetics theory is the cornerstone of evolutionary analyses. Empirical tests of the theory, however, are challenging because of the complex relationships between space, dispersal, and evolution. Critically, we lack quantitative validation of the spatial models of population genetics. Here we combine analytics, on- and off-lattice simulations, and experiments with bacteria to perform quantitative tests of the theory. We study two bacterial species, the gut microbe *Escherichia coli* and the opportunistic pathogen *Pseudomonas aeruginosa*, and show that spatiogenetic patterns in colony biofilms of both species are accurately described by an extension of the one-dimensional stepping-stone model. We use one empirical measure, genetic diversity at the colony periphery, to parameterize our models and show that we can then accurately predict another key variable: the degree of short-range cell migration along an edge. Moreover, the model allows us to estimate other key parameters, including effective population size (density) at the expansion frontier. While our experimental system is a simplification of natural microbial community, we argue that it constitutes proof of principle that the spatial models of population genetics can quantitatively capture organismal evolution.

Keywords

spatial population genetics; genetic drift; range expansion; neutral evolution; bacterial colonies; biofilms

Introduction

The neo-Darwinian synthesis brought population genetics into the heart of evolutionary biology (Fisher 1930; Wright 1932; Dobzhansky 1937), something that has since been crystallized in a vast array of sophisticated models that incorporate factors such as

*Corresponding author; papers.korolev@gmail.com. †Corresponding author; kevin.foster@zoo.ox.ac.uk.

stochasticity, spatial structure, and epistasis (Crow and Kimura 1970; Felsenstein 1976; Loewe and Hill 2010). Despite this and the widespread acceptance of population genetics as the cornerstone of evolutionary theory, the low ratio of empiricism to modeling has long been noted in the literature (Mayr 1959; Rousset 2004). Although genomics studies can be a very useful for testing qualitative predictions of population genetics—for example, the importance of effective population size for purifying selection (Lynch 2002)—a controlled experiment in the laboratory is better suited for the specific goal of validating the theory.

One of the key barriers to quantitative tests is that many theories of population genetics consider well-mixed populations—that is, populations without spatial structure. The assumption of a lack of structure can often be a drastic approximation. Moreover, the ever-growing number of theories that do incorporate spatial structure (Rousset 2004; Nagylaki and Lou 2008; Excoffier et al. 2009) present a new set of problems because empirical tests require the knowledge of certain demographic parameters (such as dispersal rates) for which few data are typically available (Rousset 2004). An open challenge for population genetics, therefore, is to develop new empirical methods and tests of the theories themselves and, in particular, quantitative tests that can rigorously verify models of spatially structured populations.

We focus on neutral population-genetics theory of effectively one-dimensional populations applied to range expansions. The importance of spatial separation for the development of genetic diversity was first emphasized by Wright (1943) and was further developed by Malécot (1955, 1975; see also Kimura and Weiss 1964; Nagylaki 1974; Barton et al. 2002). In addition to stationary populations, Malécot's analysis also describes populations undergoing range expansions, as was shown in Korolev et al. (2010). While the theory is in principle general to a range expansion of any species, we focus here on one of the most tractable of experimental systems, bacteria growing on a solid surface.

Microbial populations are often used to elucidate the mechanisms of evolution because of their short generation times and the ease of genetic manipulation (Elena and Lenski 2003; Barrett et al. 2005; West et al. 2007). With a few exceptions (Rainey and Travisano 1977; Rainey and Rainey 2003; Kerr et al. 2006), the homogeneous environment of a chemostat or a vigorously shaken test tube is preferred because of its simplicity and the ease of comparison to mathematical models, which are often intractable for populations with spatial structure. However, even for microbes it is now clear that most natural populations are not well mixed, and the neighborhood of any individual is enriched in the individuals that are ancestrally related to it. This is especially true for many bacteria that commonly form dense surface-attached communities known as "biofilms" (Kolter and Greenberg 2006; Nadell et al. 2009). These resilient microbial communities are found in numerous contexts, both medical and industrial, and are often highly resistant to antibiotics. For this reason, there is now a large body of literature on the formation of biofilms, which underlines their complexity by demonstrating major differences in the regulation and development of biofilms among strains and species (Beloin and Ghigo 2005). Key factors affecting the structure of biofilms include physical interactions among cells, the diffusivity of metabolites (Dockery and Klapper 2001; Stewart 2003), and the genetic regulation of growth rate in response to both nutrient levels (Vulic and Kolter 2001) and quorum sensing (Parsek and Greenberg 1999; Vulic and Kolter 2001; Hammer and Bassler 2003; Nadell et al. 2008). Growth in microbial groups also often relies on growth-promoting secretions, which themselves can be subject to nutrient and quorum sensing regulation (Xavier et al. 2011).

Despite the potential complexities, the spatiogenetic patterns in microbial biofilms can be visualized by fluorescent labeling part of the population (Klausen et al. 2003; Hallatschek et al. 2007). With this setup, cells of the same color can be ancestrally related, whereas cells of

different colors cannot (provided mutations that switch colors can be neglected). Colony biofilms are particularly amenable to such visualizations (Rani et al. 2007). As an example, the fluorescent image of a bacterial colony is shown in figure 1*A*. The most striking feature of this image is that an initially well-mixed population of two alleles spatially separates or demixes into monoallelic sectors. The sectoring pattern in this colony is due to genetic drift at the front of the expansion—that is, stochastic changes in the allele frequencies due to random variation in the number of offspring keeping up with the outgoing expansion. Since at any given time only a few cells at the edge of the expansion colonize a petri dish, the fluctuations can be large and lead to a random local increase in the number of cells of one color at the expense of the other. Once one of the colors becomes extinct as a result of this stochastic process in a small region of the expanding colony, a sector is formed, and it persists as long as its boundaries do not coalesce. Such spatiogenetic correlations significantly affect the evolutionary dynamics of populations (Korolev et al. 2010). General examples range from the neutral theory in ecology (Houchmandzadeh and Vallade 2003) and evolution (Kimura and Weiss 1964; Charlesworth et al. 2003), through the propensity for malignant tumor formation (Michor et al. 2003), to the evolution of cooperation (Hamilton 1964; Nowak and May 1992; Hauert and Doebeli 2004; Rousset 2004). Indeed, the latter has direct relevance to microbes and the evolution of growth-promoting secretions that are cooperative in the sense that the secretion of one cell will help others grow. These secretions often require a strong spatiogenetic structure that prevents exploitative genotypes from using the secretions of the cooperative genotypes (Nadell et al. 2009, 2010).

Here, we use the spatiogenetic patterns in bacterial colonies to test the population-genetics models of range expansions. We begin by synthesizing and extending the theory of range expansions in a way that allows us to compare the predictions of this theory to our experimental data. In particular, two aspects of the theory are novel. First, we find that chiral or nonneutral growth results in sector boundaries being logarithmic spirals (eq. [9]). Second, we present a detailed discussion of the role played by population density on the number of surviving sectors. This theory of range expansions is tested using experiments with two ecologically distinct bacterial species: the gut microbe *Escherichia coli* and the opportunistic pathogen *Pseudomonas aeruginosa*. Specifically, we test the dependence of the genetic diversity at the population front (sector number) on the physical size of the founding population (initial radius of the colony) and check the assumption that boundaries between genotypes perform radial random walks. We show that the model not only captures the qualitative relationship between demographic parameters and spatiogenetic patterns but also allows us to estimate these parameters with considerable accuracy. This quantitative validation of the theory supports the utility of population-genetics models to capture the evolutionary process.

Methods

Data have been deposited at Dryad (http://dx.doi.org/10.5061/dryad.3147q).

Bacterial Strains and Plasmids

We used the same two strains of *Escherichia coli* as in Hallatschek et al. (2007), where it was shown that these strains are equally fit in well-mixed liquid culture and are genetically identical except for the different fluorescent protein gene carried on plasmids. The *Pseudomonas aeruginosa* strain was generously provided by Roberto Kolter and was originally isolated from freshwater in Columbia. We labeled this strain with cyan and yellow fluorescent protein (Xavier et al. 2009), resulting in two strains that are genetically identical except for the fluorescence gene inserted into the chromosome. The pairs of colored strains within each species grew at the same rate and behaved neutrally on the plates. This was confirmed from the shape of the sector boundaries in the experiments (Hallatschek and

Nelson 2010). In particular, strains with a fitness advantage would tend to curve outward and cut off neighboring strains (see eq. [9]). This boundary bending was not observed, and the analysis of sector boundaries showed that, for both *E. coli* and *P. aeruginosa*, the pairs of colored strains had the same fitness in our experiments (see Results).

Growth Conditions

Each strain was first grown separately in 3 mL of liquid LB medium (lysogeny broth) supplemented with the appropriate antibiotic (carbenicillin for *E. coli* and gentamycin for *P. aeruginosa*) at 100 μg/mL. The test tubes were incubated at 37°C and constantly shaken. After overnight growth, bacteria were diluted to an optical density of 0.05 (at a wavelength of 600 nm) in 3 mL of liquid LB medium without antibiotics and grown for 2 h. Then, the two strains of interest were mixed at a 1 : 1 ratio as measured by optical density, and a small volume of bacterial mix ranging from 1 to 64 μL was pipetted onto an LB agar plate with 1.5% agar and the appropriate antibiotic at 100 μg/mL. The plates were then incubated at 21°C or 37°C with high humidity. *Escherichia coli* strains did not grow at 21°C.

Microscopy and Image Processing

Fluorescent images were obtained using a fluorescence stereoscope (Zeiss Lumar V.12) and a scanner (Amersham 392 Typhoon 9400; GE Healthcare). Scanned plates were imaged from the bottom at 50-μm resolution with an Alexa 526 filter.

The initial radii of the colonies were measured within an hour of inoculation by fitting a circle to the colony boundary using the fluorescent microscope software. The number of sectors in each colony was calculated manually. MatLab R2007 was used to extract the radii of the colonies at later times and the shape of sector boundaries from the images. In particular, we used the routine edge to identify sector boundaries and a least-square-fitting routine to find the circumference of colonies (Pratt 1987).

To calculate the heterozygosity $H(t, \phi)$ defined by equation (3) and shown in figure 2, we obtained the allele frequencies $f(t, \phi)$ along a set of concentric circles (which stand for different time points). While in simulations the frequencies are known, we had to calculate $f(t, \phi)$ in experiments from the fluorescence intensities of the colony images. This was done by choosing a threshold intensity such that any pixel with higher intensity is assigned $f = 1$ and any pixel with lower intensity is assigned $f = 0$. We then used equation (3) to calculate the heterozygosity, and the averaging was done over all the points ϕ_1 and ϕ_2 such that $\phi = \phi_1 - \phi_2$ in a given circle and over all the circles corresponding to the same time point in the colonies used in that experiment.

Estimating D_g and D_s

Here we describe how D_g and D_s were estimated from fluorescent images of genetic demixing. The diffusion constant (D_s^m in table 1) was estimated from the random walks of sector boundaries. We first obtained the polar coordinates $\varphi(r)$ of individual sector boundaries as described above. Once the coordinates were known for $r \in (r_i, r_f)$, we evaluated $\varphi(r) - \varphi(r_i)$ for each boundary. We then estimated the mean and variance of this quantity as a function of $\ln(r/r_i)$ and $1/r_i - 1/r$, respectively. The model accurately described boundary motion, and the variance of $\varphi(r) - \varphi(r_i)$ was a linear function of $1/r_i - 1/r$; therefore, $2D_s/v_\parallel$ was found from the least mean square estimate of the slope.

In addition, both D_g and D_s (D_s^e in table 1) were estimated from the experiments done with different volumes of inoculant. Different volumes of inoculant led to different initial radii of the colonies, which were measured as described above. The observed values of radii were then binned, and the average number of sectors at the end of the experiment (when

coalescence of sectors ceased) were calculated within each bin. The average number of sectors was a linear function of the square root of the initial radius. The least mean square estimate equals $H_0(2\pi v_\parallel/D_s)^{1/2}$ for the slope and $2\pi H_0 v_\parallel/D_g$ for the intercept (see eq. [7]). Since both H_0 and $v_\parallel$ were measured directly, these least mean square estimates allowed us to measure D_g and D_s.

On-Lattice Simulations

We also tested our genetic demixing model described below via simulations designed to mimic range expansions of natural populations. In these simulations, the habitat was a square lattice of islands that could support at most N individuals, similar to that in the stepping-stone model (Kimura and Weiss 1964). During every time step, each island had one opportunity for outward migration and one opportunity for growth.

During the migration update, an individual from an island could migrate to a randomly selected nearest-neighbor island with a probability proportional to the fraction of cells on the island and the fraction of empty spots (vacancies) on that neighbor island. The coefficient of proportionality ($d \leq 1$) determined the magnitude of the effective diffusion constant and was used to vary the rate of migration. Since the migration rate was proportional to the fraction of vacancies, an organism could exchange positions only with a vacancy, not another organism. Hence, there was no migration in the dense-packed interior of the simulated colony. Its spatiogenetic pattern was thus a frozen record of the genetic composition of the expansion front, similar to Hallatschek et al. (2007). Hence, we analyzed both the experimental data and the simulation data on an equal footing. Periodic boundary conditions were used on all four sides of the two-dimensional habitat.

During the growth update, a randomly selected individual in each island reproduced with a probability proportional to the fraction of the cells on the island and the fraction of vacancies on that island; the coefficient of proportionality ($p \leq 1$) allowed us to vary the growth rate. We checked that this update rule could reproduce simple logistic growth (Murray 2003). To compare with the experiments, we simulated only two alleles and tracked the genetic identity of the individuals throughout the simulation. A typical simulated range expansion is shown in figure 1*B*.

Off-Lattice Simulations

We also carried out simulations using an off-lattice agent-based approach that explicitly described the dynamics of each agent based on the microenvironment it perceived. The agents represented cells, and their behavior was governed by the rules that mimic the behavior of real bacteria: they increased in radius when taking up nutrients, divided once they reached a certain critical size, and moved continuously (off lattice) when pushed by neighboring cells. The pushing algorithm was based on a hard-sphere model where each agent moved in the direction that minimized the overlap with neighboring agents. The model used a fast numerical method (FAS multigrid) to solve partial differential equations for the nutrient concentration. The algorithms were described in a previous publication (Xavier et al. 2005).

We simulated neutral evolution experiments by implementing two genotypes with equal phenotypes except for their color. Agent growth was determined by the local concentration of the growth-limiting nutrient following Monod kinetics (see Nadell et al. 2010). Simulations were initiated with a number of individuals placed in a circular homeland. The growth of the agents and colony was then followed in time, producing structures such as the one shown in figure 1*C*.

Models of Range Expansions

The importance of space to evolutionary processes has been studied in many disciplines, including population genetics, ecology, and probability theory, where the classic spatial models—for example, the stepping-stone model (Kimura and Weiss 1964; Nagylaki 1974; Malécot 1975), Hubbell's model (Hubbell 2001; Houchmandzadeh and Vallade 2003), and the voter model (Cox and Griffeath 1986)—have been developed for stationary (nonexpanding) populations. Following Korolev et al. (2010), we use coalescent theory (Kingman 1982) to track the genetic diversity at different points in space and adapt the classic results from spatial population genetics (Malécot 1955, 1975; Kimura and Weiss 1964; Nagylaki 1974; Barton et al. 2002) to understand the evolutionary dynamics of clonal populations during a range expansion.

While neutral expansions typically occur on a two-dimensional surface, a one-dimensional model is sufficient to understand how genetic diversity changes at the front of the expanding population. For stationary (nonexpanding) populations, the dimensionality of space plays an important role (Barton et al. 2002; Korolev et al. 2010), and dimensional reduction from two spatial dimensions to one is impossible. In expanding populations, however, this dimensional reduction is possible because all relevant dynamics occur at the perimeter of the population, which is effectively one-dimensional. Indeed, the organisms far behind the expansion front cannot affect the future genetic composition at the frontier because their offspring are not part of the outgoing expansion wave. Although the dynamics of allele frequencies at the front can be described by an effective one-dimensional model, this is not in general true for the population in the wake of the expansion, for which a full two-dimensional model is necessary. Nevertheless, the results of the one-dimensional model are applicable behind the front on the timescales when migration in the interior can be neglected (Hallatschek and Nelson 2010).

Although a one-dimensional model is easier to analyze, the demographic parameters in this model are in general complicated functions of the demographic parameters of the original two-dimensional population. For example, the effective linear density (carefully defined in the appendix) of the one-dimensional population at the frontier depends on the carrying capacity of the habitat, the shape of the density profile at the front, and the probability that individuals slightly behind the expansion edge will migrate forward1 (Hallatschek and Nelson 2008). The exact nature of these dependences varies for different reproduction and migration models, but the effective one-dimensional description is universal, once model-specific details are absorbed in the effective demographic parameters.

To make a direct connection with the subsequent experiments, we assume that only two genotypes are present and that they have the same fitness. We also assume that the colony radius R increases linearly with time t:

$$R(t)=R_0+v_{\|}t, \tag{1}$$

where $v_{\|}$ is the radial expansion velocity and R_0 is the radius of the colony when $t = 0$. Expansions with a constant velocity are expected for reaction-diffusion population dynamics (Murray 2003) and have been observed in microbial colonies (Wakita et al. 1994; Hallatschek et al. 2007). Range expansions in our simulations also obey equation (1).

[1]Organisms expanding in spatially homogeneous habitats typically reach the carrying capacity in the interior of the population, but the population density decays to 0 beyond the expansion front. Here, we assume that the density profile at the front remains constant during the expansion, resulting in a constant population density in the effective one-dimensional model. This behavior is typical for reaction-diffusion systems with saturation (Fisher 1937; Kolmogorov et al. 1937; Murray 2003) and is confirmed in our simulations (see also Nadell et al. 2010).

Spatial genetic demixing (fig. 1) is one of the key predictions of this model. Although the total number of individuals is very large, locally only a finite number reproduce, leading to genetic drift (i.e., fluctuations in the genetic composition of a population due to the random sampling of offspring to form the next generation). Over time, these demographic fluctuations tend to reduce genetic diversity at the growing front of the population. Eventually, a single genotype will reach fixation at any point on the expanding edge. In bacterial colonies, this fixation manifests as the formation of a sector (fig. 1*A*).

Below we present two complimentary ways to rigorously analyze the behavior of the one-dimensional population at the expansion front. The first approach is an extension of classic spatial population genetics (Malécot 1955, 1975; Kimura and Weiss 1964; Nagylaki 1974, 1998; Barton et al. 2002) that uses coalescent theory (Kingman 1982) to track the genetic diversity at different points in space. At the core of this approach is the stepping-stone model of Kimura and Weiss (1964), which describes the migration and reproduction of individuals in a population. A concise description of this model is given in the appendix. The second modeling approach was proposed in Hallatschek et al. (2007) and Hallatschek and Nelson (2010), and it tracks the domain boundaries between the regions with different genotypes. Although it is applicable only after monoallelic domains (sectors) form, the second approach yields predictions complimentary to those of the first approach. We use these additional predictions to conduct a quantitative experimental test of our theory.

Model 1: Tracking Local Genetic Diversity

Spatial genetic demixing can be most readily understood by calculating the genetic diversity of the population over time. The genetic composition of the front at time t is fully described by the relative fraction (frequency) of one of the two alleles $f(t, \phi)$, where ϕ is the azimuthal angle around the center of the population. The relative fraction of the other allele is then $1 - f(t, \phi)$. These fractions change in time because of migration and reproduction. One can think of the population as a circle of islands (demes), each carrying a certain number of organisms (this number of organisms sets the strength of the genetic drift). Each generation, the organisms reproduce within a deme—by, say, a Wright-Fisher process—and migrate between neighboring demes. This conceptual population model is the stepping-stone model of Kimura and Weiss (1964); we present a mathematical formulation of the stepping-stone model in the appendix.

Because $f(t, \phi)$ is a random variable, which is different for every expansion, we need to calculate quantities averaged over repeated expansions to compare the predictions of the model with experimental data. The first moment (mean) of $f(t, \phi)$ does not contain information about genetic demixing because the relative fraction of two neutral genotypes does not change on average:

$$\mathbb{E}[f(t,\phi)]=f_0, \tag{2}$$

where f_0 is the initial frequency of allele one and $\mathbb{E}$ denotes averaging with respect to different realizations of the range expansion.

The second moment (variance), however, captures the spatiogenetic correlations. We define the average spatial heterozygosity as

$$H(t,\phi_1-\phi_2)=\mathbb{E}\{f(t,\phi_1)[1-f(t,\phi_2)]\}+\mathbb{E}\{f(t,\phi_2)[1-f(t,\phi_1)]\}, \tag{3}$$

which is the probability of sampling two different alleles separated by distance $|\phi_1 - \phi_2|R(t)$ along the circular frontier. Note that H depends only on the relative separation between the two points because of rotational invariance of the averages.

At the beginning of the expansion, $H(0, \phi) = H_0 = 2f_0 (1 - f_0)$ because the individuals are well mixed. As the expansion progresses and large monoallelic domains appear, the average spatial heterozygosity decreases for ϕ smaller than the typical domain size because sampling two different alleles in close proximity is unlikely when the population is demixed. For ϕ much larger than the typical domain size, the average spatial heterozygosity is close to H_0 because the two sampling points belong to different and uncorrelated sectors carrying allele 1 with probability f_0 or allele 2 with probability $1 - f_0$. Hence, the spatial scale over which $H(t, \phi)$ increases from its lowest value to H_0 characterizes the typical domain size (see fig. 2*A* and the discussion below).

As in well-mixed populations, the depression of $H(t, \phi)$ for small ϕ can be understood by working backward in time and calculating the probability that two individuals have a common ancestor. Apart from t, this probability depends on the spatial separation of the individuals, the extent of lateral movement during the expansion, and the density of the population. Indeed, for their ancestral lineages to coalesce (converge on the common ancestor), the spatial random walks performed by the lineages must move to the same spatial location, and the lineages must originate from the same individual in that location. Therefore, the forward-in-time equation for the dynamics of $H(t, \phi)$ must have a diffusion term to account for the lateral random walk of the ancestral lineages and a sink term to account for the loss of genetic diversity due to lineage coalescence. This equation is rigorously derived in Korolev et al. (2010; see also the appendix) and is given below:

$$\frac{\partial}{\partial t} H(t,\phi) = \frac{2D_s}{(R_0 + v_\parallel t)^2} \frac{\partial^2}{\partial \phi^2} H(t,\phi) - \frac{D_g}{R_0 + v_\parallel t} H(t,0)\delta(\phi), \tag{4}$$

where $\delta(\phi)$ is Dirac's delta function, D_s is the spatial diffusion constant of an ancestral lineage, and $D_g\delta(x)$ is the coalescence rate of two ancestral lineages separated by distance x along the front. This rate and, therefore, D_g are inversely proportional to the generation time and decreases with the population density at the front because the probability of two nearby organisms to be siblings is inversely proportional to the effective population density (Ewens 1982), defined in the appendix. In other words, D_g is a phenomenological parameter characterizing the strength of genetic drift. Although one may naively expect that D_g is inversely proportional to the actual population density (or carrying capacity), this is often not the case (Hallatschek and Nelson 2008). The factors of $R_0 + v_\parallel t$ are necessary to relate the angular coordinate ϕ to distances along the constantly inflating circumference of the population.

From the forward-in-time perspective, the first term on the right-hand side of equation (4) represents the diffusive mixing in the population, and the second term is responsible for the loss of genetic diversity due to sampling. Smaller population densities lead to more severe bottlenecks, faster reduction in local genetic diversity, and, therefore, larger values of D_g.

The exact solution of equation (4) is given in Korolev et al. (2010) and is plotted in figure 2*A* for a particular choice of parameters. The coalescence term in equation (4) acts as a sink and is responsible for the development of the dip in $H(t, \phi)$ around $\phi = 0$. This decrease in the probability of finding two different alleles at the same spatial location (given by $H(t, 0)$) tracks local fixation and the formation of sectors.

There are two important timescales for the genetic demixing process in the model. The first timescale characterizes the interplay of migration and genetic drift and is the time over which genetic diversity is lost locally (monoallelic domains form) when radial growth is slow. As was shown in Korolev et al. (2010), this timescale is given by

$$\tau_s = D_s / D_g^2. \quad (5)$$

The other timescale characterizes the rate of the spatial expansion and is the time it takes the colony to double its initial radius:

$$\tau_d = R_0 / v_{\|}. \quad (6)$$

As we shall show below, the ordering of these timescales ($\tau_d > \tau_s$ vs. $\tau_d < \tau_s$) determines whether the domain formation is sensitive to the constant increase in the front circumference.

For long times ($t \gg \tau_d$ and $t \gg \tau_s$), the average spatial heterozygosity $H(t, \phi)$ reaches a steady state (see fig. 2*A* and Korolev et al. 2010). At this point, the diffusive motion of the sector boundaries is negligible compared with the deterministic increase in separation among sector boundaries due to colony expansion; therefore, sector boundaries do not meet, and the number of sectors does not change. The number of sectors in this asymptotic state is calculated in Korolev et al. (2010) and is given by

$$\mathcal{N} = \frac{2\pi H_0 v_{\|}}{D_g} + H_0 \sqrt{\frac{2\pi R_0 v_{\|}}{D_s}}. \quad (7)$$

The two terms on the right-hand side of equation (7) reflect the dominant contributions of the two possible regimes for different orderings of τ_d and τ_s. The second term on the right-hand side of equation (7) gives the dominant contribution to $\mathcal{N}$ (R_0) when sectors form before the diffusive motion of sector boundaries becomes weaker than their deterministic separation because of colony expansion. This condition is equivalent to $\tau_s \ll \tau_d$—that is, the sectors form before the colony expands significantly.2 This term was first calculated in Hallatschek et al. (2007) and Hallatschek and Nelson (2010) by neglecting the initial stage of genetic demixing when distinct sectors appear. The first term on the right-hand side of equation (7) is an additional contribution to $\mathcal{N}$ (R_0) from the early stage of genetic demixing and gives the dominant contribution in the opposite limit $\tau_s \gg \tau_d$ when sectors do not coalesce after formation.

Equation (7) allows us to estimate D_s and D_g (see "Methods") and forms the basis of the quantitative test of spatial population genetics discussed below.

Model 2: Tracking Domain Boundaries

Although equation (4) explains spatial genetic demixing during range expansions, it does not directly describe the dynamics of the domain boundaries, one of the most prominent features of the spatiogenetic patterns shown in figure 1. We now introduce a second modeling approach to understand genetic demixing that ignores the dynamics of individual organisms and instead focuses on a larger spatial scale of domain boundaries among monoallelic

2A sector of angle Φ increases in size as $\Phi R(t)$, so the change of the sector size due to the colony growth in time Δt is $\Delta t \Phi v_{\|}$. The change of this sector size due to the diffusive motion of the boundaries is on the order of $(D_s \Delta t)^{1/2}$. These two changes are equal for $\Delta t^* \sim D_s / (v_{\|}^2 \Phi^2)$. The typical size of sectors when they first appear is about $(D_s \tau_s)^{1/2} = D_s / D_g$; therefore, $\Delta t^* \sim R_0^2 D_g^2 v_{\|}^{-2} D_s^{-1} = \tau_d^2 / \tau_s$. Thus, for $\tau_s \ll \tau_d$, $\Delta t^* \gg \tau_d \gg \tau_s$, which implies that sector boundaries appear much earlier than the time when the deterministic expansion moves them far apart. Consequently, sector boundaries have a significant chance to coalesce, eliminating some of the sectors.

regions. This forward-in-time approach is simpler than and complementary to the coalescent analysis outlined in the previous section.

After the sectors form, the dynamics of the population can be described in terms of sector boundaries. As ancestral lineages, sector boundaries diffuse and coalesce upon meeting. When domain boundaries meet, the number of domains decreases and the average sector size increases, further increasing the average separation between the alleles. Random walks that disappear upon meeting each other (so-called annihilating random walks) have been extensively studied in linear geometries (Bramson and Lebowitz 1991; Masser and ben Avraham 2000; Ódor 2004). In the circular geometry of interest to us here, a phenomenological description of domain boundary motion consistent with equation (4) is given in Hallatschek et al. (2007) and Hallatschek and Nelson (2010) and is summarized by the following stochastic differential equation:

$$d\varphi(r)=\frac{v_\perp}{v_\parallel r}dr+\sqrt{\frac{2D_s}{v_\parallel r^2}}dB(r), \tag{8}$$

where $\varphi(r)$ is the trajectory of a sector boundary in polar coordinates, $B(r)$ is a Weiner process, and $v_\perp$ is the velocity of the boundary perpendicular to the direction of the expansion. As in equation (4), the factors of r are needed to switch from Cartesian to polar coordinates. Note that D_s, the diffusion constant of the boundaries in equation (8), is exactly the same as the diffusion constant of ancestral lineages in equation (4). The equality of diffusion constants can be established by comparing the average domain sizes calculated using model 2 and model 1 in the limit $D_g \to \infty$. This equality has also been derived in Hallatschek and Korolev (2009) for the one-dimensional stepping-stone model, which describes range expansions with linear fronts.

In the model of Hallatschek et al. (2007) and Hallatschek and Nelson (2010), the perpendicular velocity $v_\perp$ accounts for the growth rate difference between the organisms on the opposite sides of the boundary. For neutral genetic demixing, $v_\perp$ equals 0, unless the colony growth is chiral. By chiral growth we mean that sector boundaries and ancestral lineages twist exclusively clockwise or exclusively counterclockwise. This should be distinguished from sector boundary bending due to unequal fitnesses of the strains, which causes sector boundaries around the fitter strain to bend in opposite directions.

In fact, we found that *E. coli* colonies have a noticeable chirality3 (see fig. 3*A*). Indications of weak chirality for the same strains were reported in Hallatschek et al. (2007); however, chirality is more pronounced under our growth conditions. From equation (8), the average twisting of the sector boundary should obey the following equation:

$$\mathbb{E}[\varphi(r)]=\varphi(r_i)+\frac{v_\perp}{v_\parallel}\ln\left(\frac{r}{r_i}\right), \tag{9}$$

where the radius r_i and angle $\varphi(r_i)$ define the initial point from which we follow the boundary. (The subscript "i" in r_i refers to "initial" and is not used for indexing purposes.) The initial radius r_i is larger than R_0 because the boundaries are clearly visible only after the

[3]Macroscopic chiral patterns are often related to microscopic symmetry breaking. After this article was submitted for publication, we learned of work by Wang et al. (S. Wang, L. Furchtgott, K. Huang, and J. Shaevitz, unpublished manuscript, 2011) that reveals a left-handed twisting motion of the *E. coli* cell wall by tracking various markers. Given the approximate statistical bias observed by Hallatschek et al. (2007) for the first few close-packed rows of actively growing *E. coli* cells to be oriented tangent to the border of an actively growing frontier, the screwlike growth observed by Wang and colleagues on the micron scale might explain the macroscopic chiral patterns on a centimeter scale. Indeed, the advance of a left-handed screw at the frontier would result in sector boundaries twisting counterclockwise when viewed from the bottom of a petri dish, as is seen in figure 3*A*.

initial stage of genetic demixing. The perpendicular velocity $v_\perp$ can then be estimated from the slope of $\varphi(r)$ versus $\ln(r/r_i)$. Once the chiral parameter $v_\perp$ is known, we can subtract the deterministic part from the motion of the boundary and focus on the diffusive part described by equation (8) with $v_\perp = 0$, as we shall do in the rest of this article. We shall further assume that $\varphi(r_i) = 0$ since we can always rotate our reference frame by an arbitrary angle. Note that equation (4) remains valid for colonies with chiral growth such as the one shown in figure 3*A* because the average spatial heterozygosity, defined as the probability that two sampled individuals carry different alleles, depends on the separation between the two sampled individuals (and not their absolute angular positions), which is unaffected by the deterministic twisting of the ancestral lineages. We note in passing that the average sector shape given by equation (9) defines an equiangular logarithmic spiral (Huntley 1970), with the angle determined by the dimensionless ratio $v_\perp/v_\parallel$.

To characterize the randomness of domain boundary motion, we compute the variance of $\varphi(r)$ from equation (8):

$$\mathrm{Var}\,(\varphi)=\frac{2D_s}{v_\parallel}\left(\frac{1}{r_i}-\frac{1}{r}\right). \tag{10}$$

As $r \to \infty$, the variance of φ reaches a finite limit, which is consistent with the prediction of equation (4) that the coalescence of domain boundaries ceases at long times. The predicted inverse dependence of Var (φ) on r can be used to check that domain boundaries behave as circular random walks and to estimate D_s from the experimental data. Indeed, $2D_s/v_\parallel$ is the slope of the line Var (φ) versus $1/r_i - 1/r$.

The annihilating random walks of sector boundaries and ancestral lineages provide an intuitive explanation of genetic demixing, but a full test of the validity of the models requires a quantitative test. Accordingly, the experiments and simulations presented here were designed to test whether the dynamics of bacterial colonies could be accurately described by equations (4) and (8) by checking their predictions: equations (7), (9), and (10).

Results

We performed experiments and on- and off-lattice simulations (see fig. 1) to test whether our models accurately describe the development of spatiogenetic correlations during range expansions. In the on-lattice simulations, range expansions are driven by migration into unoccupied territories followed by growth to saturation, similar to the range expansions of natural populations over long distances. In contrast, the off-lattice simulations were designed to mimic microbial colonies whose range expansions are driven by the mechanical forces that cells exert on each other as they grow. The utility of the simulations is threefold. First, they allow us to confirm that it is indeed possible to describe two-dimensional expansions with a one-dimensional model. Second, the simulations show that spatial genetic demixing is a generic phenomenon that occurs in populations with very different biological properties and, therefore, growth dynamics. This in turn enables us to test whether our phenomenological model can accurately describe genetic demixing in these quite-different populations in addition to our experiments. Third, the simulations allow us to explore whether sector boundaries are affected by roughening of the expansion front (Kardar et al. 1986; Hallatschek and Nelson 2010) for parameter values of interest to us here.

For both experiments and simulations, we found that the expansion velocity remained constant during the time of observation (see fig. 4). The values of the expansion velocities are given in table 1. We also found that the motility and growth of the cells behind the front

were significantly reduced, and the sectoring pattern did not change in the interior of the colony. Therefore, the sectoring pattern was a record of the genetic composition of the colony front throughout the expansion: a ring of radius $r > R_0$ around the center of the colony represents the state of the colony front at time $t = (r - R_0)/v_{\parallel}$. This mapping of time onto radius allowed us to relate the time dependence of any variable in our one-dimensional model to the two-dimensional pattern in a colony.

A very direct way to verify the theory is to compare the solution of equation (4) to $H(t, \phi)$ obtained in simulations and experiments (see fig. 2). We found a qualitative agreement between the experiments, simulations, and theory, but quantitative comparisons could not be made because we were unable to reliably correlate the fluorescence intensity to the fraction of the corresponding cells at a given pixel. Therefore, we focused on the two variables in the model that can be examined quantitatively: the random-walk-like motion of the sector boundaries $\varphi(r)$ and the number of surviving sectors at long times $\mathcal{N}R_0$). For both variables, we first confirm that the experimental data are consistent with the corresponding analytical dependence. After this, we estimate all four model parameters—D_s, D_g, R_0, and $v_{\parallel}$—and show that the estimate of D_s corresponds well to an independent estimate that comes from direct measurement of the wandering of boundaries between genotypes.

Lateral Movement of Genotypes

By tracing a large number of sector boundaries $\varphi(r)$, we verified that their trajectories had the statistical properties of random walks in the circular geometry. For *Pseudomonas aeruginosa* colonies, the average angular positions of sector boundaries remained constant, indicating equal fitness of the strains (see fig. 5). *Escherichia coli* colonies experienced chiral growth, and the boundaries twisted counterclockwise on average (see fig. 3*A*). This twisting (also observed in Hallatschek et al. 2007) was consistent with constant lateral velocity $v_{\perp}$ and equation (9), as shown in figure 3*B*. We also analyzed the deviations of sector boundaries from their average position. As shown in figure 3*C*, the variance of $\varphi(r)$ was consistent with the random walk behavior in the circular geometry summarized by equation (10).

Genetic Diversity at the Population Front

Our model of genetic demixing during bacterial range expansions relies on three assumptions: equal fitness of the strains, expansion with a constant velocity, and diffusive migrations at the colony front. Since these assumptions were confirmed by the experimental data, we turned to the number of surviving sectors given by equation (7), which relies not only on the aforementioned assumptions but also on the solution of the model. We found that our one-dimensional model provided a good description of two-dimensional simulations (fig. 6*B*, 6*C*). The experimentally observed $\mathcal{N}R_0$) was also consistent with the predicted square-root dependence for both *E. coli* and *P. aeruginosa*, as can be seen in figure 6*A*.

Quantitative Test of the Models

The agreement between the experiments and the theory not only suggests that bacterial expansions can be described by the one-dimensional stepping-stone model but also allows us to put this agreement to a quantitative test. From the fit of equation (7) to the dependence of sector number on the initial population radius, we estimate the two key parameters of the model, $D_s/v_{\parallel}$ and $D_g/v_{\parallel}$. The estimate of $D_s/v_{\parallel}$—the degree of lateral movement by the genotypes as the population expands—forms a prediction that allows us to test the ability of equation (4) to capture the process of population expansion and genetic drift in the bacteria. This test is performed by taking the predicted value of $D_s/v_{\parallel}$ and comparing it to a direct estimate that is made using our measurements of the sector boundary trajectories. Specifically, the diffusion constant of random walks performed by sector boundaries was

calculated by fitting the observed variance of $\varphi(r)$ to the expected dependence for a random walk in the circular geometry given by equation (10) (see fig. 3*C*). The parameters estimated by the two methods are summarized in table 1. The quantitative agreement between these quite different estimates of $D_s/v_\parallel$ suggests that our spatial population-genetics model can indeed describe the simple bacterial range expansions discussed here.

Although it is beyond the scope of this article to quantitatively relate the microscopic properties of cells and colonies to the macroscopic parameters of our phenomenological model, we can nevertheless understand how to infer some information about the processes within the colony from the measured values of D_s and D_g. To make this connection, we need to know the cell size a_c and the generation time τ. These parameters do not have well-defined values because cells are rod shaped and the generation time presumably depends on the location of a cell relative to the front, where the nutrients are readily available. Therefore, we did not measure a_c and τ, but we believe that it is reasonable to assume that a_c is on the order of 1–5 μm and that τ is on the order of 10^3–10^4 s (Bartlett et al. 1985; Yang et al. 2008). Given the wide range of uncertainty and the differences in a_c and τ between different experimental conditions, we can make only very rough estimates, and the numbers given below should be considered accurate within an order of magnitude. The velocity of expansion is given by $N_v a_c/\tau$, where N_v is the number of actively growing cells at the front (Hallatschek et al. 2007). From table 1, N_v is on the order of 10 cells, which is similar to the earlier results (Hallatschek et al. 2007). The square root of the product of the diffusion constant and the generation time is proportional to the root mean square lateral displacement of a cell during one generation. From the data, we estimate that this lateral displacement is on the order of 10 cell sizes, which is comparable to the advancement of the front during the same period of time. As we show in the appendix (eq. [A6]), the strength of genetic drift D_g can be expressed in terms of effective population density ρ_e; in fact, $D_g = 1/(\rho_e \tau)$. We find that the effective number of cells in a spatial region of the size of a cell, a_c, is smaller or much smaller than 1, as has been predicted by Hallatschek and Nelson (2008).

Conclusions

Expanding bacterial colonies are subject to strong genetic drift, resulting in local extinctions of the genotypes present in the population. These local extinctions manifest themselves in the formation of monoallelic sectors, provided the organisms migrate very slowly behind the expansion front. Despite the potential for complexity in the physical and biotic interactions within bacterial groups, our experiments strongly suggest that these boundaries behave as annihilating random walks, and we show how to estimate their diffusion coefficients from the fluorescent images of spatial genetic demixing (see “Methods”).

The spatial separation of two alleles in an expanding bacterial colony can also be described by a generalization of the stepping-stone model, which accounts not only for the behavior of the sector boundaries but also for their formation from the initially well-mixed population. Although the expansion takes place on a two-dimensional surface, our experiments and simulations strongly suggest that its effects on the genetic composition of the population front can be described by a simple one-dimensional model with only four parameters (D_s, D_g, R_0, and $v_\parallel$), all of which can be measured experimentally as described in “Methods.”

This generalization of the stepping-stone model not only explains spatial genetic demixing but also characterizes neutral evolutionary dynamics of expanding populations quantitatively, just as the Wright-Fisher and Moran models capture essential aspects of the dynamics of stationary well-mixed populations. This quantitative description and the methods of estimating the model parameters could be important for understanding evolution within microbial communities. For example, the survival probability of a beneficial mutation

depends on D_g, which characterizes the strength of genetic drift in spatially extended populations (Maruyama 1970; Doering et al. 2003; Korolev et al. 2010), and the evolution of cooperation depends on the relatedness (or assortment) of the organisms, which is described by $H(t, \phi)$ (Nadell et al. 2009, 2010). In addition, a better understanding of bacterial range expansions could also facilitate the development of new methods of genetic inference from spatially resolved data.

More generally, we believe that a key contribution of this study is to illustrate the potential of spatial population-genetics models to make quantitative predictions. Clearly, our bacterial system is not representative of all species and is a system that lends itself to making measurements that will not always be possible in other species. Nevertheless, our study demonstrates the potential for these models to generate accurate quantitative predictions in addition to confirming the utility of population genetics as a method for evolutionary analysis.

Acknowledgments

D.R.N. and K.S.K. are grateful to O. Hallatschek for many helpful conversations and for a critical reading of the manuscript. Some of the computations in this article were run on the Odyssey cluster supported by the FAS Sciences Division Research Computing Group. This work was supported by National Institute of General Medical Sciences Center of Excellence grant 5P50 GM 068763-01 to K.R.F. and D.R.N. and European Research Council grant 242670 to K.R.F. D.R.N. also acknowledges the support of the National Science Foundation through grant DMR-0654191 and the Harvard Materials Research Science and Engineering Center through grant DMR-0820484.

APPENDIX

Genetic Drift in Expanding Populations

In this appendix, we briefly outline the theory leading to equation (4) and discuss the relationship between D_g and effective population size N_e. Because the number of offspring fluctuates, allele frequencies change stochastically. In the classic Wright-Fisher model of a haploid population with two neutral alleles at a locus, the dynamics of the frequency of one of the alleles, $f(t)$, is governed by the following stochastic differential equation:

$$df(t) = \sqrt{\frac{1}{N\tau} f(t)[1 - f(t)]}\, dB(t), \tag{A1}$$

where $B(t)$ is a Weiner process and N is the population size. The generation time τ is shown explicitly to facilitate the transition to the spatial model later. For other population models (e.g., the Moran model), equation (A1) still describes the dynamics of $f(t)$ but with a different coefficient inside the square root (Kingman 1982;Wakeley 2008). It is then convenient to introduce an effective population size N_e instead of N so that equation (A1) is satisfied. Under certain assumptions, the effective population size equals the ratio of the census population size to the variance in the number of offspring (Kingman 1982). This choice of N_e is sometimes called “coalescent effective population size” (Wakeley 2008) or “eigenvalue effective population size” (Ewens 1982).

From stochastic differential equation (A1), we can compute moments (e.g., average or variance) using Itô’s lemma (Gardiner 1985; Risken 1989). For example, the average heterozygosity $H(t) = \mathbb{E}2f(t)[1 - f(t)]\}$ obeys the equation

$$\frac{d}{dt}H(t) = -\frac{1}{N_e\tau}H(t), \tag{A2}$$

which describes the loss of genetic diversity due to genetic drift.

The one-dimensional stepping-stone model is a generalization of equation (A1) to an array of demes labeled by an index k (Kimura and Weiss 1964; Barton et al. 2002):

$$df_k(t)=\frac{m}{2\tau}\,[\,f_{k-1}(t)-2f_k(t)+f_{k+1}(t)]+\sqrt{\frac{1}{N_e\tau}f_k(t)[\,1-f_k(t)]}\;dB_k(t), \tag{A3}$$

where m is the fraction of organisms migrating out of each deme into one of the two nearest neighbors. We can obtain a continuous-in-space description of this lattice model by introducing a spatial variable $x = ka$, where a is the distance between the demes. The corresponding stochastic differential equation reads (Korolev et al. 2010)

$$df(t,x)=D_s\frac{\partial^2}{\partial x^2}f(t,x)dt+\sqrt{D_g f(t,x)[\,1-f(t,x)]}\;dB(t,x), \tag{A4}$$

where

$$D_s=\frac{ma^2}{2\tau} \tag{A5}$$

and

$$D_g=\frac{a}{N_e\tau}. \tag{A6}$$

Note that the factor of a in the definition of D_g arises from the normalization of the spatial Wiener process $B(t, x)$. The ratio $\rho_e = N_e/a$ is the effective population density, which appears in the continuous formulations of the stepping-stone model (Nagylaki 1974). When equation (A4) is applied not to strictly one-dimensional populations but to linear expansion fronts, D_g depends on many demographic parameters in an intricate way (Hallatschek and Nelson 2008).

For circular expansions of interest to us here, it is convenient to introduce polar coordinates $R(t) = R_0 + v_\parallel t$ and ϕ around the center of the expansion. Equation (A4) then takes the following form:

$$df(t,\phi)=\frac{D_s}{R^2(t)}\frac{\partial^2}{\partial\phi^2}f(t,\phi)dt+\sqrt{\frac{D_g}{R(t)}f(t,\phi)[\,1-f(t,\phi)]}\;dB(t,\phi). \tag{A7}$$

To obtain equation (4), we use Itô's lemma to differentiate $H(t, x)$ defined by equation (3) in terms of $f(t, \phi)$.

Literature Cited

Barrett RDH, MacLean RC, Bell G. Experimental evolution of *Pseudomonas fluorescens* in simple and complex environments. American Naturalist. 2005; 166:470–480.

Bartlett R, Mazens M, Testa M. Differentiation of *Pseudomonas aeruginosa* and *Enterobacteriaceae* in direct smears based on measurements by scanning electron microscopy. Diagnostic Microbiology and Infectious Disease. 1985; 3:143–147. [PubMed: 3919992]

Barton NH, Depaulis F, Etheridge AM. Neutral evolution in spatially continuous populations. Theoretical Population Biology. 2002; 61:31–48. [PubMed: 11895381]

Beloin C, Ghigo J. Finding gene-expression patterns in bacterial biofilms. Trends in Microbiology. 2005; 13:16–19. [PubMed: 15639627]

Bramson M, Lebowitz JL. Asymptotic behavior of densities for two-particle annihilating random walks. Journal of Statistical Physics. 1991; 62:297–372.

Charlesworth B, Charlesworth D, Barton NH. The effects of genetic and geographic structure on neutral variation. Annual Review of Ecology, Evolution, and Systematics. 2003; 34:99–125.

Cox JT, Griffeath D. Diffusive clustering in the two dimensional voter model. Annals of Probability. 1986; 14:347–370.

Crow, JF.; Kimura, M. An introduction to population genetics theory. New York: Harper & Row; 1970.

Dobzhansky, T. Genetics and the origin of species. New York: Columbia University Press; 1937.

Dockery J, Klapper I. Finger formation in biofilm layer. SIAM Journal on Applied Mathematics. 2001; 62:853–869.

Doering CR, Mueller C, Smereka P. Interacting particles, the stochastic Fisher-Kolmogorov-Petrovsky-Piscounov equation, and duality. Physica A. 2003; 325:243–259.

Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nature Reviews Genetics. 2003; 4:457–469.

Ewens W. On the concept of the effective population size. Theoretical Population Biology. 1982; 21:373–378.

Excoffier L, Foll M, Petit RJ. Genetic consequences of range expansions. Annual Review of Ecology, Evolution, and Systematics. 2009; 40:481–501.

Felsenstein J. The theoretical population genetics of variable selection and migration. Annual Review of Genetics. 1976; 10:253–280.

Fisher, RA. The genetical theory of natural selection. Oxford: Clarendon; 1930.

Fisher RA. The wave of advance of advantageous genes. Annals of Eugenics. 1937; 7:355–369.

Gardiner, C. Handbook of stochastic methods. New York: Springer; 1985.

Hallatschek O, Korolev KS. Fisher waves in the strong noise limit. Physical Review Letters. 2009; 103:108103–108106. [PubMed: 19792344]

Hallatschek O, Nelson DR. Gene surfing in expanding populations. Theoretical Population Biology. 2008; 73:158–170. [PubMed: 17963807]

Hallatschek O, Nelson DR. Life at the front of an expanding population. Evolution. 2010; 64:193–206. [PubMed: 19682067]

Hallatschek O, Hersen P, Ramanathan S, Nelson DR. Genetic drift at expanding frontiers promotes gene segregation. Proceedings of the National Academy of Sciences of the USA. 2007; 104:19926–19930. [PubMed: 18056799]

Hamilton WD. The genetical evolution of social behaviour. I. Theoretical Population Biology. 1964; 7:1–16.

Hammer B, Bassler B. Quorum sensing controls biofilm formation in *Vibrio cholerae*. Molecular Microbiology. 2003; 50:101–104. [PubMed: 14507367]

Hauert C, Doebeli M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature. 2004; 428:643–646. [PubMed: 15074318]

Houchmandzadeh B, Vallade M. Clustering in neutral ecology. Physical Review E. 2003; 68:61912–61918.

Hubbell, SP. The unified neutral theory of biodiversity and biogeography. Princeton, NJ: Princeton University Press; 2001.

Huntley, HE. The divine proportion: a study in mathematical beauty. New York: Dover; 1970.

Kardar M, Parisi G, Zhang Y. Dynamic scaling of growing interfaces. Physical Review Letters. 1986; 56:889–892. [PubMed: 10033312]

Kerr B, Neuhauser C, Bohannan B, Dean A. Local migration promotes competitive restraint in a host-pathogen “tragedy of the commons”. Nature. 2006; 442:75–78. [PubMed: 16823452]

Kimura M, Weiss GH. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics. 1964; 49:561–576. [PubMed: 17248204]

Kingman JFC. On the genealogy of large populations. Journal of Applied Probability. 1982; 19A:27–43.

Klausen M, Heydorn A, Ragas P, Lambertsen L, Aaes-Jorgensen A, Molin S, Tolker-Nielsen T. Biofilm formation by *Pseudomonas aeruginosa* wild type, flagella and type IV pili mutants. Molecular Microbiology. 2003; 48:1511–1524. [PubMed: 12791135]

Kolmogorov AN, Petrovsky N, Piscounov NS. A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. Moscow State University Bulletin Series International A. 1937; 1:1–25.

Kolter R, Greenberg E. Microbial sciences: the superficial life of microbes. Nature. 2006; 441:300–302. [PubMed: 16710410]

Korolev, KS. PhD thesis. Cambridge, MA: Harvard University; 2010. Statistical physics of topological emulsions and expanding populations.

Korolev KS, Avlund M, Hallatschek O, Nelson DR. Genetic demixing and evolution in linear stepping stone models. Reviews of Modern Physics. 2010; 82:1691–1718. [PubMed: 21072144]

Loewe L, Hill WG. The population genetics of mutations: good, bad and indifferent. Philosophical Transactions of the Royal Society B: Biological Sciences. 2010; 365:1153–1167.

Lynch M. Intron evolution as a population-genetic process. Proceedings of the National Academy of Sciences of the USA. 2002; 99:6118–6123. [PubMed: 11983904]

Malécot G. The decrease of relationship with distance. Cold Spring Harbor Symposia on Quantitative Biology. 1955; 20:52–53.

Malécot G. Heterozygosity and relationship in regularly subdivided populations. Theoretical Population Biology. 1975; 8:212–241. [PubMed: 1198353]

Maruyama T. On the fixation probability of mutant genes in a subdivided population. Genetics Research. 1970; 15:221–225.

Masser T, ben Avraham D. Kinetics of coalescence, annihilation, and the q-state Potts model in one dimension. Physics Letters A. 2000; 275:382–385.

Mayr E. Where are we? Cold Spring Harbor Symposia on Quantitative Biology. 1959; 24:1–14.

Michor F, Frank SA, May RM, Iwasa Y, Nowak MA. Somatic selection for and against cancer. Journal of Theoretical Biology. 2003; 225:377–382. [PubMed: 14604590]

Murray, JD. Mathematical biology. Berlin: Springer; 2003.

Nadell C, Xavier J, Levin S, Foster K. The evolution of quorum sensing in bacterial biofilms. PLoS Biology. 2008; 6:e14. [PubMed: 18232735]

Nadell CD, Xavier J, Foster KR. The sociobiology of biofilms. FEMS Microbiology Reviews. 2009; 33:206–224. [PubMed: 19067751]

Nadell CD, Foster KR, Xavier J. Emergence of spatial structure in cell groups and the evolution of cooperation. PLoS Computational Biology. 2010; 6 e1000716.

Nagylaki T. The decay of genetic variability in geographically structured populations. Proceedings of the National Academy of Sciences of the USA. 1974; 74:2523–2525. [PubMed: 267945]

Nagylaki T. The expected number of heterozygous sites in a subdivided population. Genetics. 1998; 149:1599–1604. [PubMed: 9649546]

Nagylaki T, Lou Y. The dynamics of migration-selection models. Tutorials in Mathematical Biosciences IV. 2008; 1922:117–170.

Nowak MA, May RM. Evolutionary games and spatial chaos. Nature. 1992; 359:826–829.

Ódor G. Universality classes in nonequilibrium lattice systems. Reviews of Modern Physics. 2004; 76:663–724.

Parsek M, Greenberg E. Quorum sensing signals in development of *Pseudomonas aeruginosa* biofilms. Methods in Enzymology. 1999; 310:43–55. [PubMed: 10547781]

Pratt V. Direct least-squares fitting of algebraic surfaces. Computer Graphics. 1987; 21:145–152.

Rainey P, Rainey K. Evolution of cooperation and conflict in experimental bacterial populations. Nature. 2003; 425:72–74. [PubMed: 12955142]

Rainey P, Travisano M. Adaptive radiation in a heterogeneous environment. Philosophical Transactions of the Royal Society B: Biological Sciences. 1977; 280:29–101.

Rani S, Pitts B, Beyenal H, Veluchamy R, Lewandowski Z, Davison W, Buckingham-Meyer K, Stewart P. Spatial patterns of DNA replication, protein synthesis, and oxygen concentration within

bacterial biofilms reveal diverse physiological states. Journal of Bacteriology. 2007; 189:4223–4233. [PubMed: 17337582]

Risken, H. The Fokker-Planck equation: methods of solution and applications. Berlin: Springer; 1989.

Rousset, F. Genetic structure and selection in subdivided populations. Princeton, NJ: Princeton University Press; 2004.

Stewart P. Diffusion in biofilms. Journal of Bacteriology. 2003; 185:1485–1491. [PubMed: 12591863]

Vulic M, Kolter R. Evolutionary cheating in *Escherichia coli* stationary phase cultures. Genetics. 2001; 158:519–526. [PubMed: 11404318]

Wakeley, J. Coalescent theory: an introduction. Roberts. CO: Greenwood Village; 2008.

Wakita J, Komatsu K, Nakahara A, Matsuyama T, Matsushita M. Experimental investigation on the validity of population dynamics approach to bacterial colony formation. Journal of the Physical Society of Japan. 1994; 63:1205–1211.

West SA, Diggle SP, Buckling A, Gardner A, Griffin AS. The social lives of microbes. Annual Review of Ecology, Evolution, and Systematics. 2007; 38:53–77.

Wright S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the Sixth International Congress of Genetics. 1932; 1:356–366.

Wright S. Isolation by distance. Genetics. 1943; 28:114–138. [PubMed: 17247074]

Xavier JB, Picioreanu C, van Loosdrecht MC. A framework for multidimensional modelling of activity and structure of multispecies biofilms. Environmental Microbiology. 2005; 7:1085–1103. [PubMed: 16011747]

Xavier JB, Martinez-Garcia E, Foster KR. Social evolution of spatial patterns in bacterial biofilms: when conflict drives disorder. American Naturalist. 2009; 174:1–12.

Xavier J, Kim W, Foster K. A molecular mechanism that stabilizes cooperative secretions in *Pseudomonas aeruginosa*. Molecular Microbiology. 2011; 79:166–179. [PubMed: 21166901]

Yang L, Haagensen J, Jelsbak L, Johansen H, Sternberg C, Hoiby N, Soren M. In situ growth rates and biofilm development of *Pseudomonas aeruginosa* populations in chronic lung infections. Journal of Bacteriology. 2008; 190:2767–2776. [PubMed: 18156255]

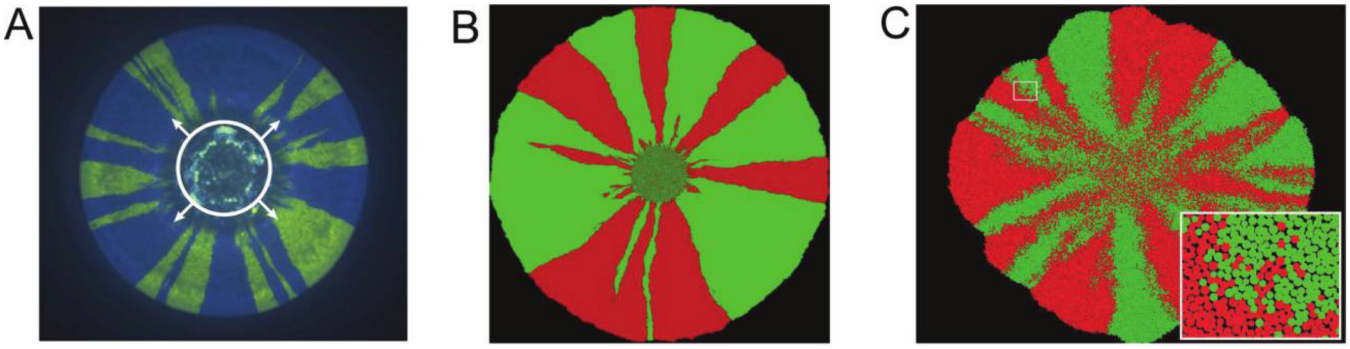

Figure 1.

Genetic demixing in experiments (*A*), on-lattice simulations (*B*), and off-lattice simulations (*C*). Different colors label different alleles, and the initial mixing ratio was 1 : 1. *A*, The petri dish was inoculated with a drop (bounded by the white circle) of a mixture of *Pseudomonas aeruginosa* cells labeled with two different fluorescent markers that do not affect the relative fitness of the cells. As the drop dries out and the colony starts to expand (shown by arrows), the population at the front of the expansion demixes (separates) into sectors of different colors. This genetic demixing is due to the loss of local genetic diversity during the expansion (see Hallatschek et al. 2007; Hallatschek and Nelson 2010; Korolev et al. 2010). This plate was incubated at 37°C. *B*, The habitat was a 1,000 × 1,000 array of islands arranged on a square lattice. Each island could harbor at most $N = 30$ individuals, consistent with the width of the layer of actively growing cells found in Hallatschek et al. (2007). The simulation was started with a well-mixed population occupying a disk of radius $R_0 = 80$ lattice spacings in the center of the habitat. *C*, The simulation was initiated with 1,024 agents placed at random within a circular inoculant in the center of the habitat. The boxed region of the colony is shown in the inset at a higher magnification. We attribute the slightly square shape of the colony to the underlying square grid used to solve the nutrient reaction-diffusion equation.

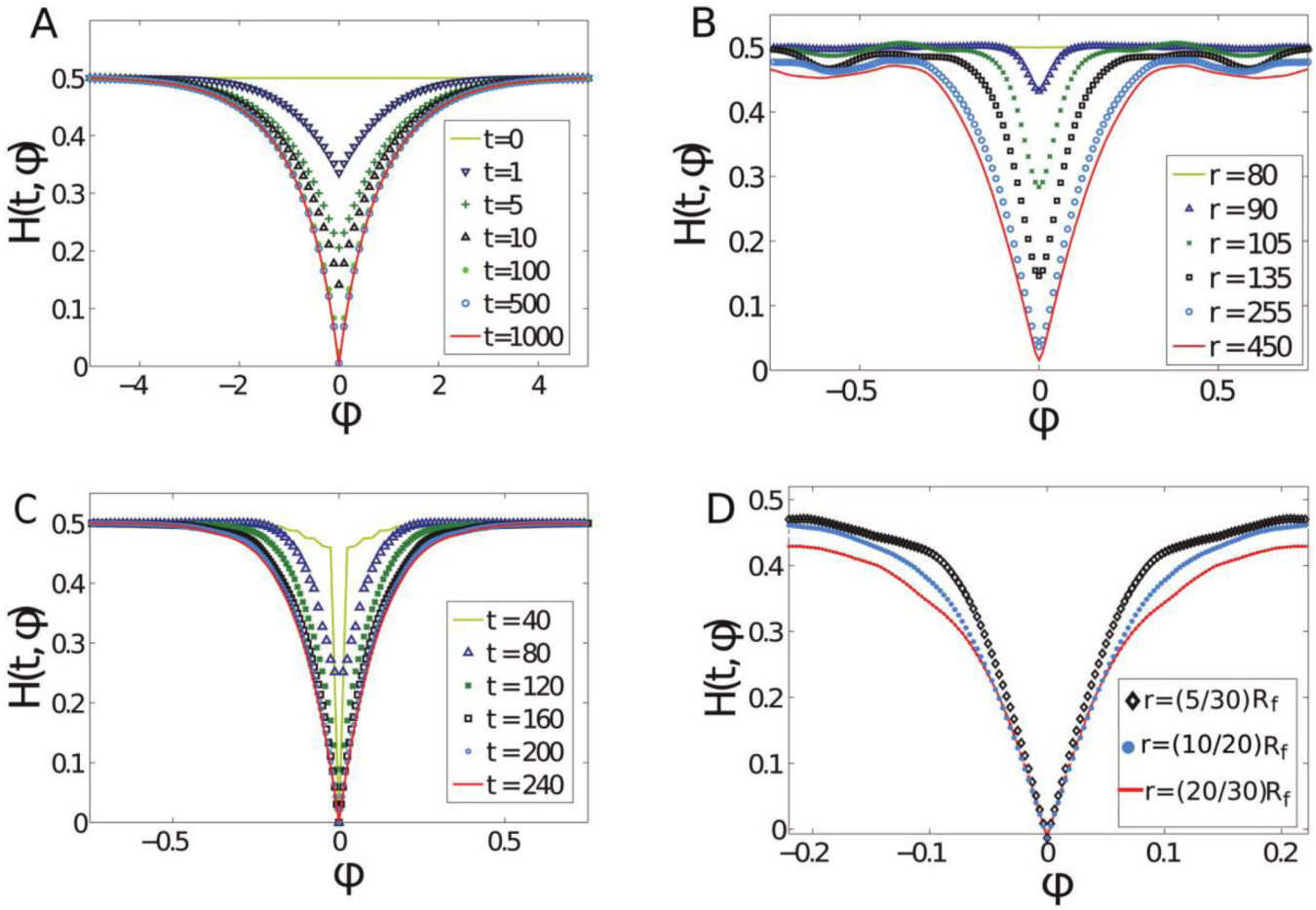

Figure 2.

Plots of average spatial heterozygosity as a function of angle during range expansions from well-mixed 1 : 1 populations in our model (*A*), on-lattice simulations (*B*), off-lattice simulations (*C*), and experiments (*D*). *A*, Solution of equation (4) with $D_s = 1$, $D_g = 1$, $v_{\|} = 1$, and $R_0 = 1$ at various times *t*. Note that there is no significant difference between $H(500, \phi)$ and $H(1000, \phi)$ because $H(t, \phi)$ reaches a nontrivial limit shape as $t \to \infty$. *B*, $H(t, \phi)$ from 24 on-lattice simulations with the same parameters as in figure 1*B*, except that $N = 300$. In agreement with *A*, we see a gradual decrease of $H(t, 0)$ with time. The radius $r = v_{\|}t + R_0$ is in direct correspondence with time *t*. *C*, $H(t, \phi)$ at the expansion frontier from 10 off-lattice simulations, as in figure 1*C*. Because off-lattice simulations model a monolayer of cells, any spatial point in a colony has a unique genetic state; hence, $H(t, 0) = 0$. *D*, Average spatial heterozygosity calculated from eight *Escherichia coli* colonies inoculated with 3 μL of the bacterial mix of cells. As in *B*, radius *r* is directly related to time *t* because the spatiogenetic pattern is a frozen record of the genetic composition of the front. The dip in $H(t, \phi)$ widens with time, in agreement with equation (4). Similar to *C*, $H(t, 0) = 0$ because we assume that $f(t, \phi)$ is either 0 or 1 at every pixel. This assumption is valid only after the initial fixation time, as discussed in the text.

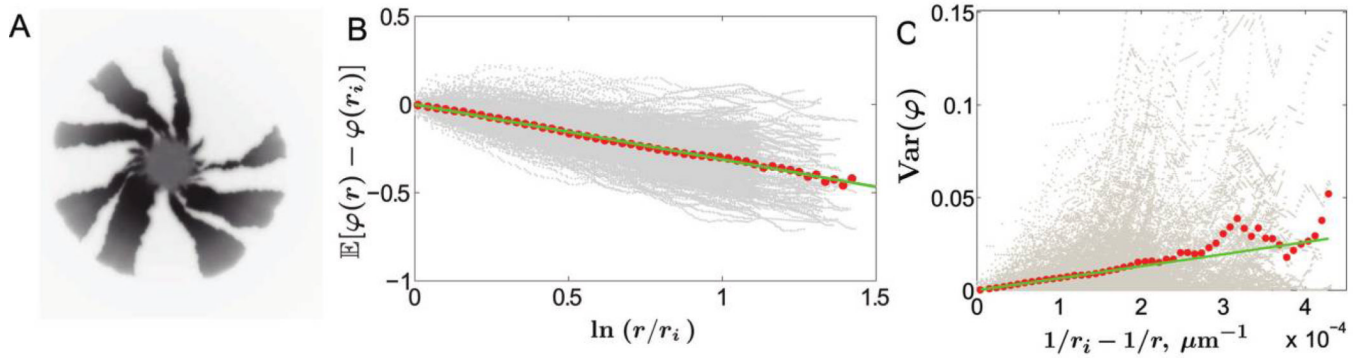

Figure 3.

Chirality and boundary wandering in *Escherichia coli* colonies. *A*, Genetic demixing in a chiral *E. coli* colony. The image, taken from the bottom of the petri dish, shows the fluorescent signal from only one of the two segregating alleles. Bacterial colonies of *Pseudomonas aeruginosa* at 37°C and 21°C did not exhibit chiral growth. *B*, Twisting of sector boundaries in *E. coli* colonies. The gray dots represent individual measurements along the sector boundaries at pixel resolution (one pixel is 50 μm) from 30 colonies (390 sector boundaries). Each red dot is the mean position of the gray dots in one of the 50 equidistant bins along the *X*-axis. The green line is the least square fit to the red dots. According to equation (9), the slope of the green line equals $v_\perp/v_\parallel$, which yields $v_\perp/v_\parallel = 0.32$. As shown in *A* and *B*, all sector boundaries twist on average in the same direction for chiral growth. In contrast, sector boundaries between nonneutral strains bend in both clockwise and counterclockwise directions because the boundaries around the fitter strains bend outward (Hallatschek and Nelson 2010). To exclude possible small fitness differences, we also separately analyzed boundaries that would bend in opposite directions if the strains were nonneutral. Nonneutrality would result in different values of $v_\perp/v_\parallel$ for these two sets of boundaries. Upon applying equivalent tests to those used in figure 5, we confirmed that $v_\perp/v_\parallel$ are the same, which supports the previous finding that these strains are neutral (Hallatschek et al. 2007). *C*, Random walks of sector boundaries. The same data as in *B* are used to plot the variance of $\varphi(r)$; gray dots are the individual data points, and red dots are the averages. The green line is the least square fit to the first 25 red dots; the last 25 dots are not used because of large fluctuations due to a smaller sample size at large *r*. According to equation (10), the slope of the green line equals $2D_s/v_\parallel$. The data sets for *P. aeruginosa* at 37°C and 21°C show similar behavior of the variance of $\varphi(r)$, but the data set for *E. coli* fluctuates less partly because of the larger sample size and larger spatial diffusion constant compared with those of the other two experiments.

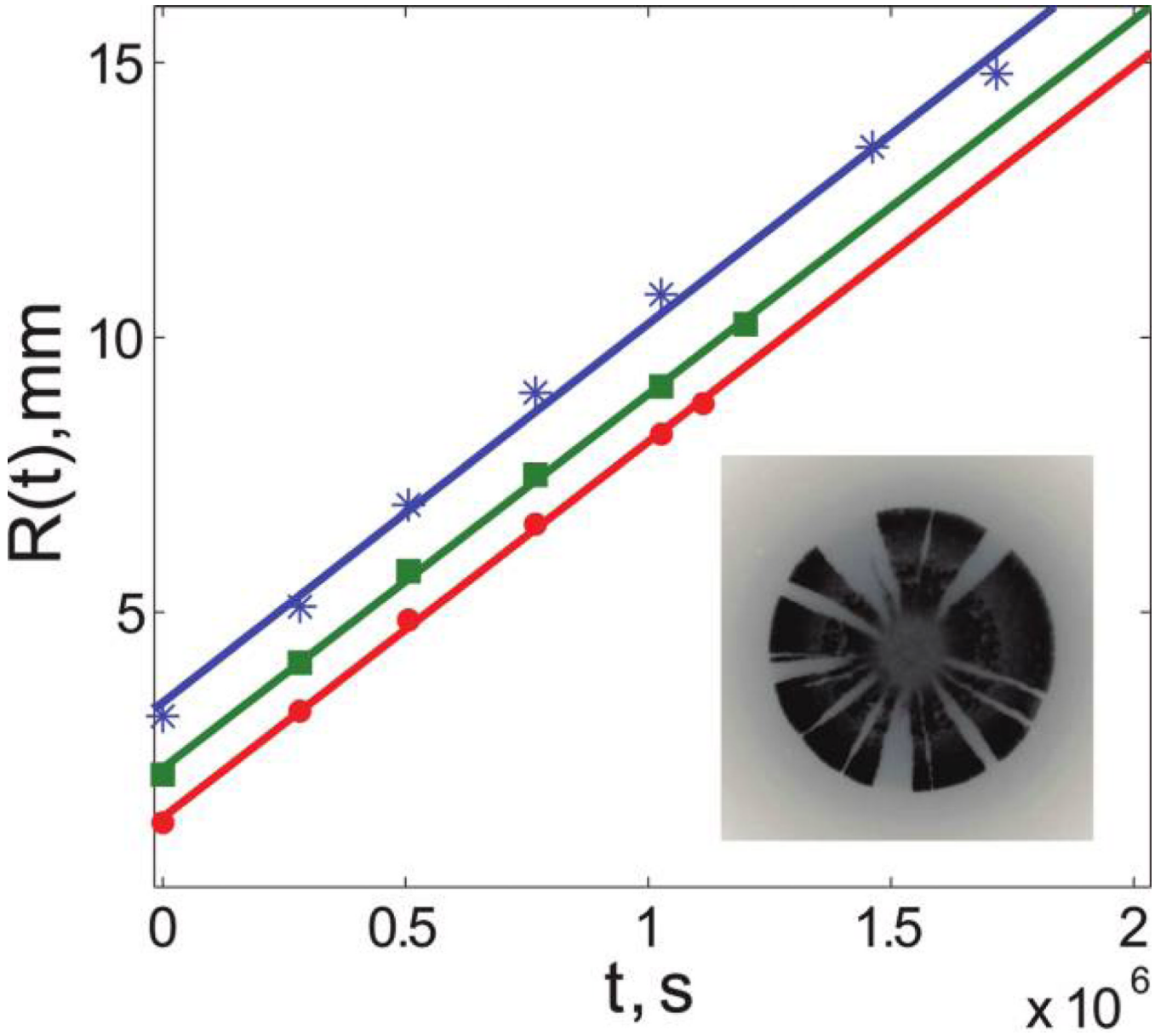

Figure 4.

Linear increase of colony radii with time. The radii of three colonies of *Pseudomonas aeruginosa* grown at 21°C are plotted as a function of time. The three sets of data correspond to three different inoculation volumes: 1 μL (red dots), 3 μL (green squares), and 14 μL (blue stars). The solid lines are the least square fits to straight lines. For all three colonies, the radius increased linearly with time and the rate of the increase was approximately independent of the initial size of the colony, as would be the case under high-nutrient conditions. Note that for the largest colony the expansion seems to slow down slightly at the end, possibly because of nutrient depletion. Similar behavior was observed for other colonies of *P. aeruginosa* and *Escherichia coli* for all growth conditions studied. The obtained values of the expansion velocities are given in table 1. The inset shows a fluorescent image of a *P. aeruginosa* colony grown at 21°C; fluorescence from only one of the two alleles is shown.

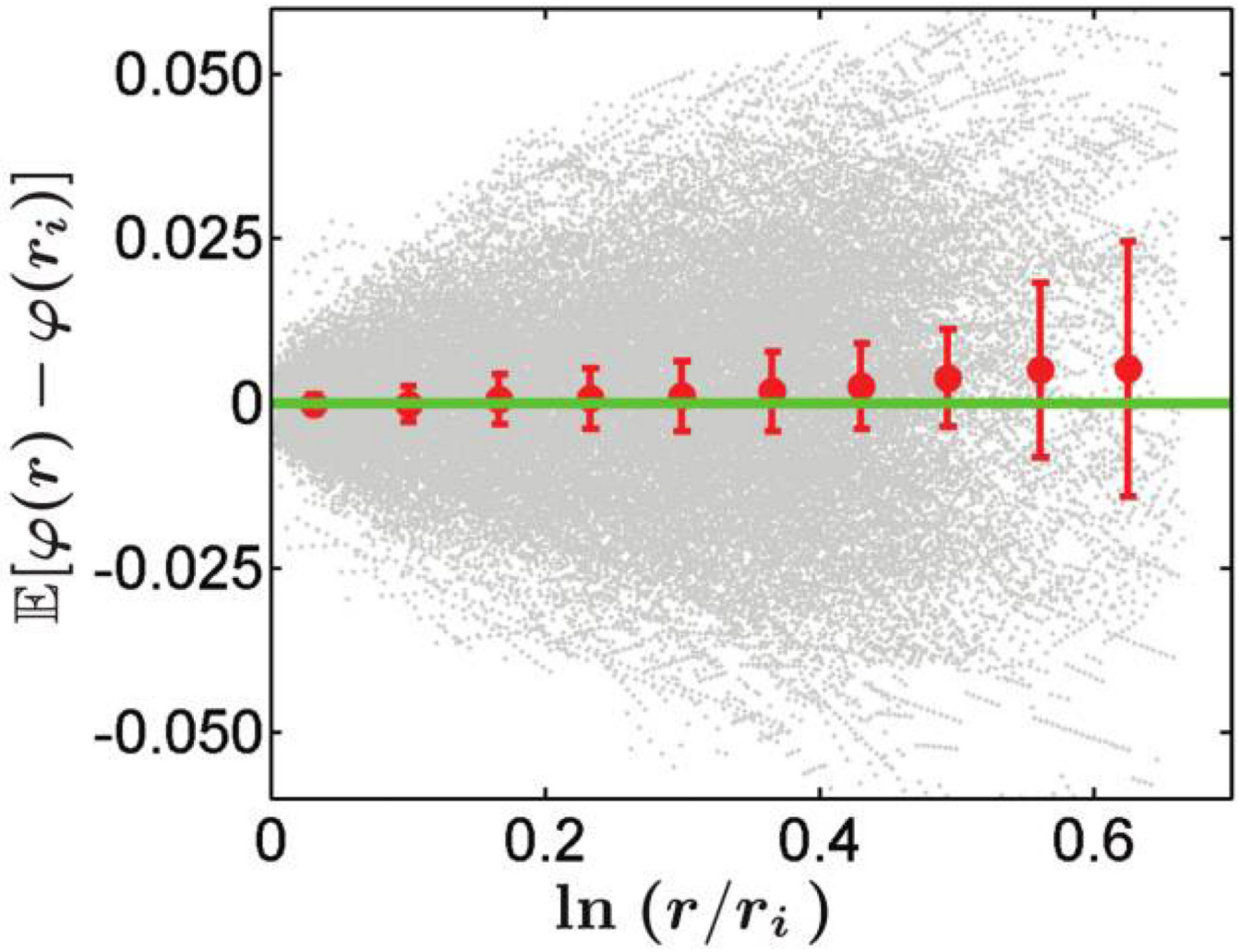

Figure 5.

Neutrality of *Pseudomonas aeruginosa* strains at 37°C. The gray dots represent individual measurements along the sector boundaries at pixel resolution (one pixel is 50 μm) from 24 colonies (524 sector boundaries) of *P. aeruginosa*. Positive φ represent boundary bending from cyan toward yellow sectors. Each red dot is the mean position of the gray dots in one of the 10 equidistant bins along the *X*-axis. The error bars represents 95% confidence intervals of the mean values. The green line is the expected dependence if the strains are neutral, $\mathbb{E}\varphi(r)] = 0$. Since the expected line passes through all of the error bars, we cannot reject the hypothesis that the strains are neutral. In fact, the largest fitness difference consistent with the data is ~10^{-4}. This comes from using equation (9) to find $v_\perp/v_\parallel$ and relating $v_\perp/v_\parallel$ to the fitness difference, as in Hallatschek and Nelson (2010) and Korolev (2010). It is then reasonable to conclude that these strains can be considered equally fit on the timescales of our experiments. The corresponding data for *P. aeruginosa* strains grown at 21°C are also consistent with the assumption that the strains are equally fit. *Escherichia coli* strains are also equally fit (see fig. 3 and Hallatschek et al. 2007).

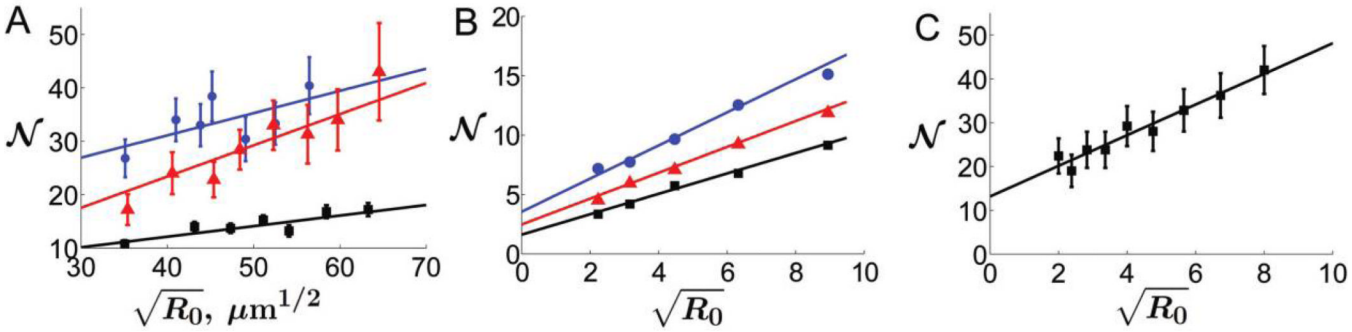

Figure 6.

Dependence of diversity at the expansion front on the initial size of the population in experiments (*A*), on-lattice simulations (*B*), and off-lattice simulations (*C*). The average number of sectors at the end of a range expansion $\mathcal{N}$ is plotted against the square root of the initial radius of the colony R_0. These coordinates are chosen so that the theoretically predicted dependence (see eq. [7]) is a straight line (solid lines). The error bars represent 95% confidence intervals. Parameters $D_s/v_\parallel$ and $D_g/v_\parallel$ can be estimated from the slope of the fit and its intercept with the *Y*-axis, respectively. *A*, *Escherichia coli* (black squares), *Pseudomonas aeruginosa* at 37°C (red triangles), and *P. aeruginosa* at 21°C (blue dots). *B*, Three different island-carrying capacities: $N = 3$ (black squares), $N = 30$ (red triangles), and $N = 300$ (blue dots). Each data point is the average of 40 simulations. Notice that the populations with higher N have more sectors in agreement with equation (3) because populations with higher densities have lower D_g. In this plot, the error bars are represented by the size of the markers used. *C*, Because of the discrete nature of these simulations, we use the square root of the number of cells at inoculation as a proxy for the initial radius.

Table 1

Measured values of the model parameters

Experiment	$v_\parallel$ (μm/s)	$D_g/v_\parallel$	$D_s^e/v_\parallel$ (μm)	$D_s^m/v_\parallel$ (μm)
Escherichia coli	$(8.3 \pm .4) \times 10^{-3}$	.75 (.32, ∞)	40 (10, 90)	32 ± 2
Pseudomonas aeruginosa at 37°C	$(17.3 \pm .9) \times 10^{-3}$	∞ (3, ∞)	4.6 (4.0, 5.2)	3.7 ± 1.1
P. aeruginosa at 21°C	$(6.9 \pm .3) \times 10^{-3}$	.17 (.10, .77)	5.5 (1.2, 13.0)	5.2 ± .5

Note: Shown are expansion velocities, migration rates, and strengths of genetic drift. D_s^e and D_s^m are two different estimates of the diffusion constant; their agreement indicates quantitative validity of our model. The values of $v_\parallel$ were obtained by fitting $R(t)$ to a linear function (see fig. 4); D_g and D_s^e are estimated from the fit of the data shown in figure 6 to equation (7). For $D_g/v_\parallel$ and $D_s^e/v_\parallel$, the 95% confidence intervals are given in parentheses. The infinite values of $D_g/v_\parallel$ are due to the possibility of the linear fit (see fig. 6*A*) passing through the origin. Values of D_s^m are measured from the random walks of sector boundaries (see eq. [10]), and the error estimates represent the standard deviation between different repetitions of the experiments. See "Methods" for a detailed description of the estimating procedures.