Fixation probabilities in weakly compressible fluid flows

Abigail Plummer[a,1], Roberto Benzi[b,c], David R. Nelson[a], and Federico Toschi[d,e,f]

[a]Department of Physics, Harvard University, Cambridge, MA 02138; [b]Department of Physics, University of Rome Tor Vergata, 00133 Rome, Italy; [c]Istituto Nazionale di Fisica Nucleare, University of Rome Tor Vergata, 00133 Rome, Italy; [d]Department of Applied Physics, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands; [e]Department of Mathematics and Computer Science, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands; and [f]Istituto per le Applicazioni del Calcolo, Consiglio Nazionale delle Ricerche, 00185 Rome, Italy

Edited by Andrea Rinaldo, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, and approved November 26, 2018 (received for review August 2, 2018)

Competition between biological species in marine environments is affected by the motion of the surrounding fluid. An effective 2D compressibility can arise, for example, from the convergence and divergence of water masses at the depth at which passively traveling photosynthetic organisms are restricted to live. In this report, we seek to quantitatively study genetics under flow. To this end, we couple an off-lattice agent-based simulation of two populations in 1D to a weakly compressible velocity field—first a sine wave and then a shell model of turbulence. We find for both cases that even in a regime where the overall population structure is approximately unaltered, the flow can significantly diminish the effect of a selective advantage on fixation probabilities. We understand this effect in terms of the enhanced survival of organisms born at sources in the flow and the influence of Fisher genetic waves.

population genetics | turbulence | Fisher waves | stochastic processes | selective advantage

Oceanic flows can affect competition between marine species in important ways, particularly at the submeso- and mesoscales where the characteristic timescales of fluid motion are comparable to the generation time (inverse growth rate) of phytoplankton (1–3). Recent observational and computational work on marine fronts, such as boundary currents and upwelling regions, has noted the likely importance of strong vertical velocities and turbulent eddies to the high productivity and genetic diversity of these regions (4–7).

However, few quantitative connections have been made between the observational and numerical data on fronts and the literature of population genetics, which often relies on simplifying assumptions such as a constant carrying capacity, discrete subpopulations, no vertical mixing, and/or fixed migration patterns (8–13). Bringing these mature fields closer together via simplified models could provide a clearer understanding of biological processes in marine environments, as well as the effect of global climate change on our oceans and atmosphere (14).

Consider a population of passive organisms that are restricted to live at a specific depth. When parcels of incompressible water carrying organisms come together in a convergence zone, or water from deeper ocean layers rises toward the surface in an upwelling zone, organisms experience an effectively compressible velocity field. Other possible sources of effective compressibility include large Stokes numbers and gyrotaxis (15).

We report here results from a 1D agent-based stochastic model of two-species competition that allows for nonuniform occupation in continuous space, coupled to a compressible flow. Although fundamentally a 3D problem, a 1D approach has proved fruitful in the past. The analysis and computational overhead are significantly simplified, and trends observed in 1D often hold in higher dimensions as well (15, 16). We first study a sinusoidal velocity field to understand the effect of a stationary source (positive slope zero crossing) and sink (negative slope zero crossing) pair and then apply this understanding to a shell model of turbulence.

We find that Kimura's famous formula for the fixation probability in well-mixed systems (17) and spatially extended systems with only diffusive motion (18) breaks down for even weakly compressible flows, dramatically increasing the influence of organisms born near source regions and lowering the overall probability of fixation for a given selective advantage, initial fraction, and system size. We explore this deviation and are able to predict the scaling behaviors observed in simulations with simple theoretical models. We believe that the source-oriented view presented here provides a promising framework for predicting ecosystem outcomes in the presence of species mutation and invasion. Our results suggest that the effect of vertical velocities must be treated with care, even in simple models.

Model for Spatial Population Genetics with Compressible Advection

The coarse-grained dynamics in the deterministic limit of our two-species stochastic model in 1D are described by two coupled partial differential equations (ref. 15 and *SI Appendix*, sections A and B),

$$\frac{\partial c}{\partial t} + \partial_x(uc) = D\partial_x^2 c + \mu c(1-c), \qquad [1]$$

$$\frac{\partial f}{\partial t} + u\partial_x f = D\partial_x^2 f + \frac{2D}{c}\partial_x f \partial_x c + sc\mu f(1-f), \qquad [2]$$

Significance

Marine fronts, where two different water masses meet, are characterized by high biological productivity. Determining which organisms are able to thrive in these environments is crucial to understanding ocean ecosystems. However, many questions about how natural selection functions in an advective environment remain unanswered. We study flows with sources and sinks to model the upwellings and convergences experienced by plankton at marine fronts. Remarkably, even when the flows are so weak that they cannot significantly alter the spatial distribution of organisms, they substantially affect genetic outcomes. The flow reduces the effective population size and makes the fate of organisms strongly dependent on local flow conditions. In particular, organisms born in upwelling regions experience a considerably enhanced probability of survival.

Author contributions: A.P., R.B., D.R.N., and F.T. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: plummer@g.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1812829116/-/DCSupplemental.

Published online December 26, 2018.

where $c(x)$ is the fraction of the no-flow carrying capacity at position x, $f(x)$ is the fraction of organisms of a given species at x, $u(x)$ is a compressible velocity field, D is the diffusion constant, μ is the growth rate when either species is dilute, and s is the selective advantage of one species over the other when the system is near its no-flow carrying capacity, as defined by the microscopic rates given in *SI Appendix*, section A.

In the limit of incompressible flow (i.e., when $u(x)$ = const., as must be the case in 1D), Eq. **1** is the Fisher equation and admits traveling-wave solutions with speed $v_p = 2\sqrt{D\mu} + u$. Eq. **2**, describing the more complicated genetic dynamics, also reduces to a similar form in the limit $c(x) \to 1$, with a genetic wavefront speed of $v_g = 2\sqrt{D\mu s} + u$. We assume the selective advantage is small ($s << 1$), allowing us to define three parameter regimes, which will also be important in the compressible case, using the local value of $u(x)$:

i) $|u| > 2\sqrt{D\mu}$: An opposing flow can arrest a Fisher population wave. Compressible flows of this magnitude can localize the population, for example, near a sink. Population dynamics in this localized regime have been studied in 1D and 2D (16, 19, 20).

ii) $2\sqrt{D\mu s} < |u| < 2\sqrt{D\mu}$: An opposing flow can arrest fragile Fisher genetic waves, but can only slow down the more robust Fisher population waves. A compressible flow near a 1D sink is not able to create a localized steady-state population structure, but can nevertheless localize genetic boundaries.

iii) $|u| < 2\sqrt{D\mu s}$: An opposing flow is so weak that it can arrest neither genetic nor population waves.

Here, we examine regimes *ii* and *iii*, which, to our knowledge, have not yet been systematically explored and have clear biological relevance. While some vertical velocities at strong upwellings certainly satisfy condition *i* [for example, we can estimate $D \approx 10^{-13}\text{m}^2/\text{s}$ by assuming unflagellated microorganisms with a Stokes–Einstein diffusivity $D = k_B T/6\pi\eta R$, $R \approx 1$ μm, and assume μ to be $1\ \text{d}^{-1}$, giving us $2\sqrt{D\mu} \approx 2 \times 10^{-4}$ m/d, whereas vertical velocities contribute roughly $|u|_{max} \approx 5$ m/d (21–24)], strongly localized structures at upwellings/convergence zones are not the only interesting situation. Conditions *ii* and *iii* describe weaker upwelling/convergence events, other sources of compressibility, and/or a strong upwelling event in a population that has a greater effective diffusivity due to active flagella (15). In these cases, the steady-state concentration profile is almost identical to the no-flow case, making them more theoretically tractable. Note that the presence of noise will change the location of the boundaries between the behaviors. For a detailed treatment of noisy Fisher waves, see refs. 25 and 26.

We further focus our investigation on fixation probabilities, a central topic in population genetics describing the stochastic process by which one species outcompetes another. The fixation probability for a species with selective advantage s in a population of size N that makes up an initial fraction f of all organisms in the absence of advection and mutation is given by Kimura's formula, which neglects terms of order s/N (17):

$$P_{fix} = \frac{1 - \exp(-sNf)}{1 - \exp(-sN)}. \quad [3]$$

This formula was first derived for the well-mixed case. However, it can be extended to 1D systems with diffusive motion using an argument inspired by Maruyama (18) (*SI Appendix*, section C) and is confirmed by our simulations. This insensitivity to spatial dimension makes the fixation probability an interesting object of study in the presence of fluid flows, because it allows us to isolate the effect of advection from that of diffusion. In contrast, fixation times, studied in the presence of advection by Pigolotti et al. (15), depend sensitively on the spatial structure of the system, and the well-mixed result for fixation times does not hold in the presence of diffusion.

In this paper, we study the behavior of fixation probabilities for weakly compressible flows. As is often the case in population genetics (27), we characterize the altered fixation probabilities by an effective population size. For small s, the derivative of Eq. **3**, with respect to s, is

$$\frac{dP_{fix}}{ds} = \frac{1}{2} Nf(1-f) + \mathcal{O}(s). \quad [4]$$

The slope at $s = 0$ in the well-mixed case is proportional to the probability of a competitive encounter between the two different species. With advection, we can interpret the slope as measuring only encounters that have an impact on the future of the system, occurring for an effective population size such that $N \to N_{eff}$ in Eq. **4**. By avoiding spatially localized structures associated with strong flows, where $|u| > 2\sqrt{D\mu}$, we can approximate the overall density of organisms as constant. This assumption greatly simplifies the determination of the effective population size, reducing it to an effective length scale. As we shall see, the result is to replace N by N_{eff} in the small-s limit appropriate to Eq. **4** and to replace N by $N_g(s)$ in the large-s limit, where the population genetics are dominated by Fisher genetic waves.

Sine Wave Flows

To better understand the impact of sources (positive slope zero crossings) and sinks (negative slope zero crossings) on fixation before tackling time-dependent turbulence, we first study a steady sine wave flow given by

$$u(x) = A_0 \sin(x - \pi/2), \quad [5]$$

in a domain of size 2π with periodic boundary conditions. The source and sink associated with this velocity profile have a characteristic time, given by $\tau_s = 1/A_0$, the inverse gradient of the velocity field at the zero crossings. Organisms are more likely to die near the sink at $3\pi/2$, where there is a constant influx of organisms, and more likely to flourish near the source at $\pi/2$. These effects violate the conditions necessary for Kimura's formula to hold in 1D (*SI Appendix*, section C) (18) and give an advantage to the source population relative to the sink population, even if there is no microscopic selective advantage involved. To determine the effective population size, we need to characterize the width of the advantageous source region.

Quasi-Neutral Competitions with Sine Wave Flows. For the case of zero selective advantage, we can analytically compute an approximate fixation probability as a function of position, using a random-walk model. As shown in Fig. 1, simulations are initialized so that only a small, localized window (length $\pi/16$ or ≈3% of the interval $[0, 2\pi]$ in Fig. 1) contains all of one species. The fixation probability is then measured as a function of the location of this window. By treating these genetic boundaries as random walkers biased by the flow, we find in the limit of small Δ,

$$P_{fix}(x, \Delta, A_0) \approx \Delta \cdot \mathcal{N}(x|\pi/2, D/A_0), \quad [6]$$

where Δ is the width of the spatial window, x is the leftmost genetic boundary, and $\mathcal{N}(x|\pi/2, D/A_0)$ is the normalized probability density function at x, given by a Gaussian with mean $\pi/2$ and variance D/A_0. Details are presented in *SI Appendix*, section D.

This function defines a length scale for our source, $l_s = \sqrt{D/A_0}$, given by the balance between diffusion and the advecting velocity field. Any organism that can diffuse to the source in a time faster than the source time $\tau_s = l_s^2/D$ is relevant to the genetic future of this sine wave system. The farther away

an organism is from the source, the stronger the outward velocity field it experiences. Any organism that moves significantly farther than l_s from the source is unlikely to be able to return and has a negligible chance of fixation as it is drawn into the sink.

We hypothesize that for zero or small selective advantage the effective population size scales as the source length scale, times the density of organisms, ρ_0, which is a constant well approximated by the no-flow limit since we are not in the regime of spatial localization. Based on these considerations, we conclude that

$$N_{eff} = B_1 \rho_0 \sqrt{\frac{D}{A_0}}, \qquad [7]$$

where B_1 is a constant of order unity. We see from Fig. 2 that the fixation probability is dramatically different from the case without an external velocity field and is well described by these considerations for small s, with $B_1 = 3.5$. The large difference between N_{eff} ($N_{eff} = 71$) and the total number of organisms in the system ($N_{tot} = 394$) is striking, considering that the presence of the flow cannot be easily detected in the snapshots of the organismic density alone without observing genetic interfaces.

Although a stochastic model was the key to predicting the magnitude of the fixation probability given by Eq. **6**, the Gaussian enhancement of the source region is also evident in a purely deterministic model. Deterministic simulations of the neutral case, which directly solve Eqs. **1** and **2** for $s = 0$ without number fluctuations, are shown in Fig. 3. Although these simulations, which are equivalent to the agent-based model in the limit $N \to \infty$, cannot directly observe fixation, we nevertheless see that a small population of one species initially localized close to the source (Fig. 3, *Inset A*) grows in size until it reaches a steady-state population (Fig. 3, *Inset B*) more than five times its original size. In contrast, a population starting near the sink shrinks to values that are less than $1/N$, approximating extinction in a simulation with discrete organisms. A Gaussian with SD l_s centered on the source provides a good fit for Fig. 3, just as in Fig. 1.

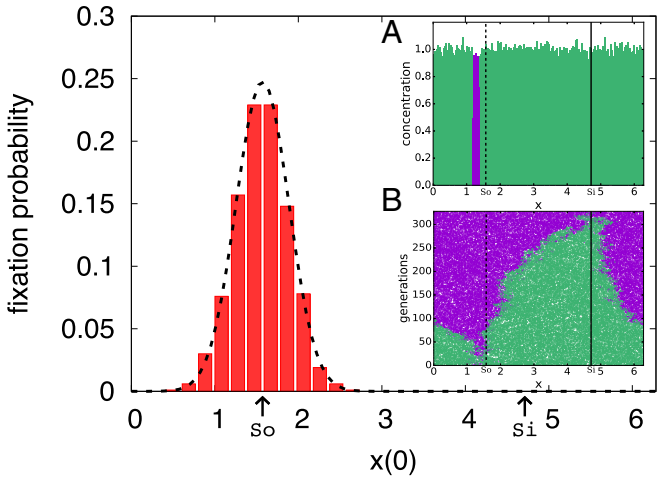

Fig. 1. The fixation probability of an initially localized purple species in a background of green, obtained from 1,000 independent realizations of an agent-based simulation of two neutral populations, as a function of the initial position of the purple distribution. Arrows indicate the locations of the source (*So*) and sink (*Si*). Populations that begin near the source in the sine wave flow are more likely to fix. The dotted line shows agreement with the random-walk model, Eq. **6**, with no fitting parameter. The flow for this plot is $u(x) = 0.05 \sin(x - \pi/2)$. (*Inset A*) One possible initial condition, with 1/32 of the domain holding only purple organisms and the rest filled with green organisms. (*Inset B*) One realization of a fixation event corresponding to the initial condition in *Inset A*. The dashed line shows the source at $\pi/2$ and the solid line shows the sink at $3\pi/2$.

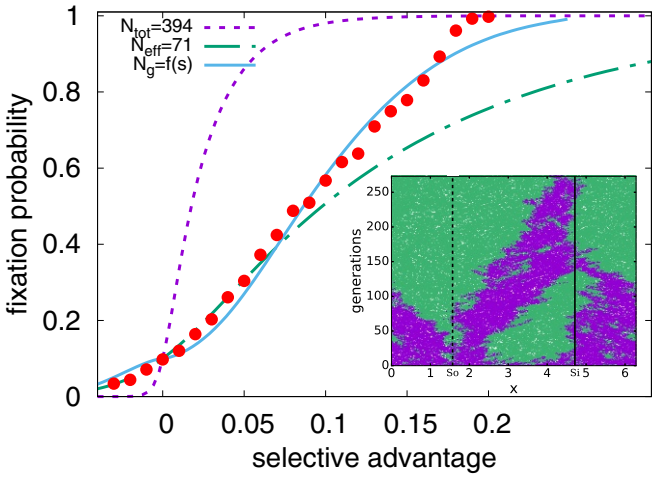

Fig. 2. The fixation probability (red circles) for an initial condition with a random 10/90 mixture of two species with $u(x) = 0.05 \sin(x - \pi/2)$, varying the selective advantage ($P_{fix}(s = 0) = 0.1$, since all organisms are equally likely to take over the population in this limit). The purple line (short dashes) shows Eq. **3**, the Kimura no-flow result, for the measured number of organisms in the simulation, $N_{tot} = 394$. The green line (short and long dashes) gives the prediction of Eq. **7**, valid in the limit of small s with $N_{eff} = 71$. The blue line (solid) gives the prediction of Eq. **9** for the large s case, with a selective advantage-dependent effective population size. Error bars inferred from 2,000 independent simulations are too small to be visible. (*Inset*) A single realization for a 50/50 mixture of two neutral species (thus, $P_{fix}(s = 0) = 0.5$). The dashed line marks the position of the source, *So*, and the solid line marks the position of the sink, *Si*. Note that the genetic interfaces between purple and green tend to annihilate in the sink at $3\pi/2$. The total concentration of organisms remains approximately uniform.

Strong Selective Advantage with Sine Wave Flows. When the selective advantage is sufficiently high (although still small compared to 1), it becomes a significant term in Eq. **2** and plays a role in setting the source length. We must now balance the effects of selection, diffusion, and the velocity.

We can do this using the framework of Fisher genetic waves. Consider a species with a significant selective advantage that is sharply localized within a background population, in a system with a sine wave flow with an amplitude that nevertheless satisfies $A_0 > 2\sqrt{D\mu s}$. This initial condition will typically produce two Fisher genetic wavefronts traveling in opposite directions. If the initial population starts on a source, both wavefronts will be supported by the external flow as they move across the system. However, if the initial population starts on a sink, both wavefronts will face an opposing flow.

There is a window of initial conditions around the source at $x = x_s = \pi/2$, defined by $u(x) = A_0 \sin(x - \pi/2) = \pm 2\sqrt{D\mu s}$, where an initial population can produce Fisher genetic waves that deterministically travel across the entire system. Within this window, the flow velocity is small enough that wavefronts can reach the source even if they do not start there. Outside of this window, Fisher genetic waves cannot reach the favorable source region.

Upon treating Fisher waves crossing the system as a proxy for fixation, this argument suggests another way to define an effective population size in the Kimura formula: the number of organisms around the source within the spatial window with boundaries given by $u(x) = \pm 2\sqrt{D\mu s}$.

When we solve for this window, we include a fitting parameter, B_2, because the traveling-wavefront solution occurs only for special initial conditions. The window is then given by half-width $\delta(s)$ such that

$$A_0 \sin(\delta) = B_2(2\sqrt{D\mu s}). \qquad [8]$$

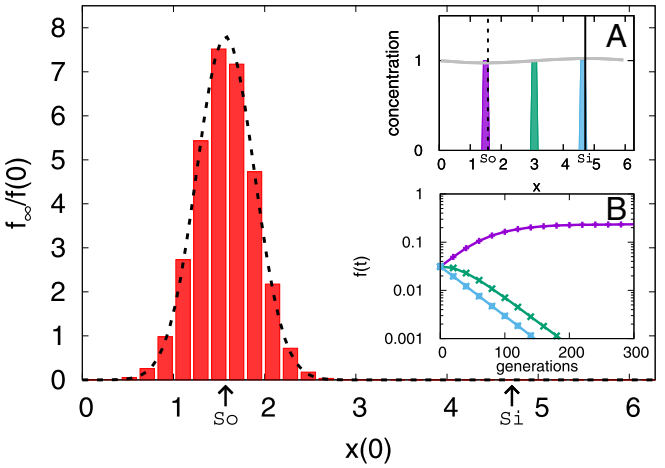

Fig. 3. The final fraction of the initially localized species, $f_\infty = \lim_{t\to\infty} f(t)$, divided by its initial fraction, as a function of its initial location. Populations that begin near the source, *So* ($x = \pi/2$), grow relative to their initial condition, and populations that begin far from the source shrink. Fit is given by a Gaussian of variance $l_s^2 = D/A_0$. Here $u(x) = 0.05\sin(x - \pi/2)$, as in Fig. 1. (*Inset A*) The gray line shows the approximately uniform steady-state total concentration of organisms [the sum of the concentrations of both species, $c(x, t)$]. The colored boxes show the initial spatial distribution of one of the species for three different simulations, relative to the source (dashed line) and the sink (solid line). (*Inset B*) The total fraction of the initially localized species in time, with the colors corresponding to the initial conditions in *Inset A*. The leftmost initial population prospers, while the middle and rightmost initial populations fall off rapidly.

After solving for $\delta(s)$, we can use the approximation of constant density to write the number of organisms in our genetic wave-defined source population as

$$N_g = 2\rho_0\delta(s). \qquad [9]$$

Note that our effective population size N_g now has s dependence. We find that $B_2 = 0.5$ provides good agreement in Fig. 2. At even higher values of selective advantage, where A_0 is no longer greater than $2\sqrt{D\mu s}$, we observe a crossover to well-mixed behavior as expected.

Upon combining these two arguments by using the largest effective population size, $\max(N_{eff}, N_g)$, in the original Kimura formula, Eq. **3**, we can explain the fixation probability for a weak steady sine wave, as shown in Fig. 4.

Turbulent Advection

To generate a 1D chaotic signal with multiscale correlations similar to turbulent flow, we use a well-established set of ordinary differential equations called a shell model (28). For specific details, see ref. 16 and *SI Appendix*, section E. For the steady sine wave flows studied in the previous section, we saw that the presence of a source can dominate fixation events. It seems reasonable to conjecture that transient, time-dependent sources have a similarly important effect in chaotic turbulent flows and that we can still characterize the system with an effective source dimension and an associated effective population size. Remarkably, we can understand much of the fixation probability with generalizations of the simple theoretical arguments we applied to the stationary sine wave case.

Quasi-Neutral Competitions with Turbulent Advection. Even in the absence of selection, turbulent dynamics provide new factors to consider when determining the source length scale. The (multiple) sources in the flow can have long or short lifetimes. Their locations move intermittently, and their slopes vary with time. Depending on how the source times compare with the other timescales dictating organism motion, the source-enhanced population at a given time may or may not be able to move with the source and retain its competitive advantage. An example of competitive turbulent dynamics for the neutral case and some further discussion are given in *SI Appendix*, section F. Although it is clear that understanding the details of fixation probabilities in turbulence is complicated, we can nevertheless apply lessons from the sine wave case to determine scaling behavior.

We assume that our effective population size is the sum of the effective population sizes corresponding to each source, averaged over time. Longer-lived sources thus contribute more to the average and have a greater effect on the effective population size. As in the case of the weakly compressible sine wave, a source's effective population size should be able to be represented as the density of organisms times a characteristic length. It is not obvious how to define a length for an arbitrary source i, but we can easily define a characteristic source time, τ_i, by taking the reciprocal of the slope of the positive zero crossing $(\frac{du(x)}{dx}|^{-1}_{x=x_i})$. Then, we construct a length by assuming the existence of a constant with units of velocity, v_C, that can depend on the diffusion constant and some details of the dynamics, but does not depend on the root mean-square velocity. We can think of v_C as related to the speed with which domain boundaries explore the system.

Therefore, our estimate for N_{eff} in turbulence is

$$N_{eff} = \rho_0 v_C \overline{\sum_i \tau_i}. \qquad [10]$$

Here, the sum is over all sources present at a given time, and the overbar indicates a time average.

Since the density of organisms, ρ_0, is approximately constant, it is proportional to μ, the growth rate. Therefore, we can also understand Eq. **10** as a balance between the organism generation time, μ^{-1}, and the source time, τ_i. If $\mu\tau_i$ is large, the generation time is short relative to the source time, and organisms can reproduce many times during the source time. These organisms and

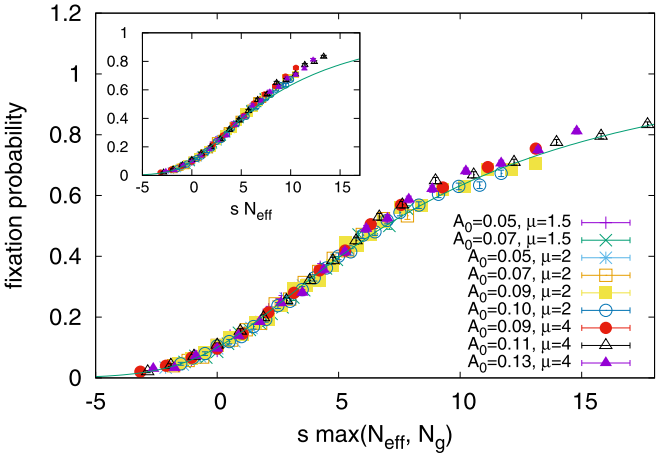

Fig. 4. When we plot the measured fixation probabilities against s times the larger effective population size of our two approximations, Eqs. **7** and **9**, our data collapse to a master curve. The agreement with our sine wave simulations suggests that we have identified how the fixation varies with A_0, μ, and s for this parameter regime. Error bars associated with 2,000 independent realizations are shown. $A_0 > 2\sqrt{D\mu s}$ for all data points shown, avoiding the crossover to well-mixed behavior seen at high s in Fig. 2. (*Inset*) When plotted as a function of the quasi-neutral theory only, Eq. **7**, we see a departure from the theory at high values of s, which motivates the high s approximation given by Eq. **9**.

their offspring thus experience an enhanced fixation probability, and such sources will give a large contribution to N_{eff}. However, if $\mu\tau_i$ is small, few organisms are affected, and the contribution to N_{eff} will be small.

We now make the approximation

$$\overline{\sum_i \tau_i} \approx \overline{n_s \tau_s}, \qquad [11]$$

where n_s is the number of sources and $\tau_s = \left\langle \left(\frac{\partial u(x,t)}{\partial x} \right)^{-2} \right\rangle^{1/2}$ is the root mean-square reciprocal velocity gradient.

We expect that, for a given value of the root mean-square velocity, u_{rms}, the total number of zero crossings, $2n_s$, scales with its gradient. As the root mean-square velocity increases, we expect the number of zero crossings to decrease. These considerations lead to the conjecture

$$n_s \sim \frac{L}{u_{rms}} \left\langle \left(\frac{\partial u(x,t)}{\partial x} \right)^2 \right\rangle^{1/2}. \qquad [12]$$

This is known to be true for Gaussian processes in 1D (29), and we have checked it explicitly via simulations with our shell model. A similar relation has been found experimentally in measurements of turbulent flows, where the number of nodes is proportional to the inverse of the Taylor microscale (30).

We note that n_s and τ_s are instantaneously strongly fluctuating quantities, and we calculate the dependence of their product on Reynolds number in *SI Appendix*, section G as $n_s\tau_s \sim \mathrm{Re}^{0.08}$. Since this dependence is very weak, we neglect it and make a mean-field approximation to find $\overline{n_s\tau_s} \approx L/u_{rms}$, where u_{rms} is now time averaged using the harmonic mean.

Upon combining these arguments, and absorbing v_C into the constant B_3, we obtain

$$N_{eff} = B_3\rho_0\overline{n_s\tau_s} = \frac{B_3\rho_0 L}{u_{rms}}. \qquad [13]$$

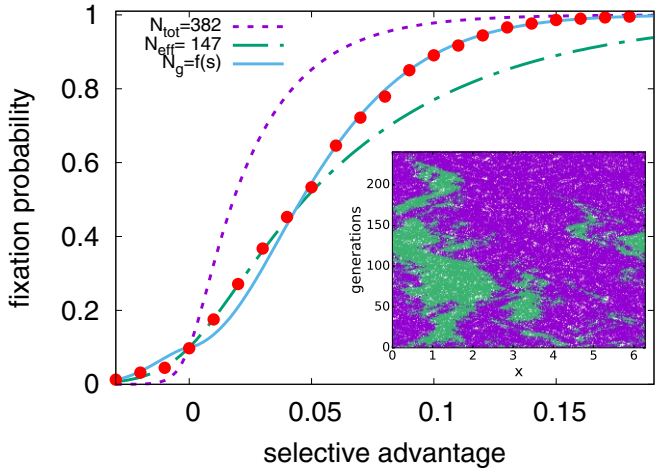

Fig. 5. The fixation probability (red circles) for an initial condition with a random 10/90 mixture of two species advected by a shell model with amplitude $A_0 = 0.3$ and growth rate $\mu = 2$, for variable selective advantage ($P_{fix}(s = 0) = 0.1$). The purple line (short dashes) shows Eq. **3**, the no-flow result, for the measured number of organisms in the simulation, $N_{tot} = 382$. The green line (short and long dashes) gives the prediction of Eq. **7**, valid in the limit of small s with $N_{eff} = 147$. The blue line (solid) gives the prediction of Eq. **9** for the large s case, with a selective advantage-dependent population size. Error bars inferred from 2,000 independent simulations are too small to be visible. (*Inset*) A single realization for a 50/50 mixture of two neutral species ($P_{fix}(s = 0) = 0.5$).

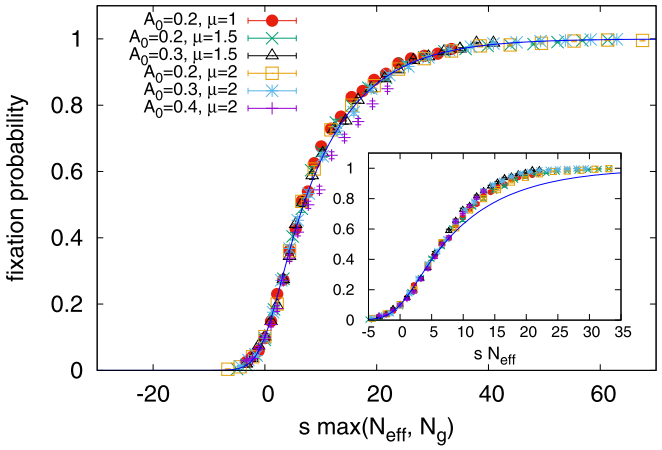

Fig. 6. When we plot the measured fixation probabilities against s times the larger effective population size of our two approximations, Eqs. **13** and **15**, our data collapse to a master curve. The agreement with our shell model simulations suggests that we have identified how the fixation varies with A_0, μ, and s for this parameter regime. Error bars associated with 2,000 independent realizations are shown. $N_{tot} \approx N_0$ for all data points shown, avoiding the crossover to localized behavior. These parameters ensure that approximate spatial uniformity is maintained. (*Inset*) When plotted as a function of the quasi-neutral theory only, Eq. **13**, we see a departure from the theory at high values of s, which motivates the high s approximation given by Eq. **15**.

Our simulations support this form of N_{eff} as shown in Figs. **5** and **6** with the constant (units of velocity to account for unknown v_C factor) $B_3 = 0.031$.

Strong Selective Advantage with Turbulent Advection. As in the sine wave case, we expect a critical value of s beyond which selection must be taken into account in the source size calculation. Unlike the quasi-neutral competitions in turbulence, however, we now have an obvious choice for a velocity that can be used to form a length scale—the Fisher genetic wavefront speed. As before, we include a dimensionless fitting parameter, B_4. Our estimate for N_g, the effective population size associated with genetic waves, is

$$N_g = B_4\rho_0(2\sqrt{D\mu s})\overline{\sum_i \tau_i}, \qquad [14]$$

where τ_i is the characteristic time of source i and the overbar indicates a time average.

Eq. **14** is also the simplest generalization of Eq. **9**, obtained by expanding the sine function to linear order in a Taylor series close to each source.

Upon estimating $\overline{\sum_i \tau_i}$ as before, we obtain

$$N_g = B_4\rho_0 L\frac{2\sqrt{D\mu s}}{u_{rms}}. \qquad [15]$$

This estimate, when $N \to N_g$ in the Kimura formula, Eq. **3**, shows good agreement with Figs. 5 and 6, with $B_4 = 0.747$. As before, we combine the arguments behind Eqs. **13** and **15** by taking the largest effective population size, $\max(N_{eff}, N_g)$, in Fig. 6.

Discussion

For both simple (i.e., sine wave) and turbulent compressible flows for which the population density is approximately uniform, we have shown that fixation probabilities are controlled by an effective population size smaller than the total number of organisms in the system in a previously undescribed and

biologically relevant region of parameter space. This reduction in the effective population size creates a significant reduction in the fixation probability as a function of selective advantage overall, but a greatly enhanced fixation probability for organisms fortunate enough to be born near sources, even those that are very weak. In the ocean, source regions can be associated with upwellings, if we assume organisms are restricted to live at a certain depth, for example. Our results suggest that the genetic compositions of these regions may have a controlling effect on the genetics of a much greater domain.

Furthermore, we have shown the reduced fixation probabilities can be explained by simple theoretical arguments and can be explored with both agent-based and relatively inexpensive deterministic simulations.

Deviations of Kimura's formula in the case of strongly compressible turbulence in 1D (i.e., flows that produce spatial localization) have also been observed in simulations, a problem closely related to gene surfing (31).

ACKNOWLEDGMENTS. We are grateful for the assistance provided by Pinaki Kumar and Francesca Tesser. We thank Amala Mahadevan, Mara Freilich, and Luca Biferale for useful discussions. Work by A.P. and D.R.N. was supported by the National Science Foundation through Grant DMR-1608501 and via the Harvard Materials Science Research and Engineering Center via Grant DMR-1435999. Work by F.T. was partially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek I (NWO-I), The Netherlands.

1. d'Ovidio F, De Monte S, Alvain S, Dandonneau Y, Lévy M (2010) Fluid dynamical niches of phytoplankton types. *Proc Natl Acad Sci USA* 107:18366–18370.
2. Lévy M, Jahn O, Dutkiewicz S, Follows MJ, d'Ovidio F (2015) The dynamical landscape of marine phytoplankton diversity. *J R Soc Interf* 12:20150481.
3. Prairie JC, Sutherland KR, Nickols KJ, Kaltenberg AM (2012) Biophysical interactions in the plankton: A cross-scale review. *Limnol Oceanogr Fluids Environ* 2:121–145.
4. Lévy M, Ferrari R, Franks PJ, Martin AP, Rivière P (2012) Bringing physics to life at the submesoscale. *Geophys Res Lett* 39:L14602.
5. Clayton S, Nagai T, Follows MJ (2014) Fine scale phytoplankton community structure across the Kuroshio front. *J Plankton Res* 36:1017–1030.
6. Clayton S, Dutkiewicz S, Jahn O, Follows MJ (2013) Dispersal, eddies, and the diversity of marine phytoplankton. *Limnol Oceanogr Fluids Environ* 3:182–197.
7. De Monte S, Soccodato A, Alvain S, d'Ovidio F (2013) Can we detect oceanic biodiversity hotspots from space?. *ISME J* 7:2054–2056.
8. Wares JP, Pringle JM (2008) Drift by drift: Effective population size is limited by advection. *BMC Evol Biol* 8:235.
9. Chust G, Irigoien X, Chave J, Harris RP (2013) Latitudinal phytoplankton distribution and the neutral theory of biodiversity. *Glob Ecol Biogeogr* 22:531–543.
10. Giovannoni SJ, Stingl U (2005) Molecular diversity and ecology of microbial plankton. *Nature* 437:343–348.
11. Krieger MS, McAvoy A, Nowak MA (2017) Effects of motion in structured populations. *J R Soc Interf* 14:20170509.
12. Herrerías-Azcué F, Pérez-Muñuzuri V, Galla T (2018) Stirring does not make populations well mixed. *Sci Rep* 8:4068.
13. Whitlock MC (2003) Fixation probability and time in subdivided populations. *Genetics* 164:767–779.
14. Behrenfeld MJ, et al. (2006) Climate-driven trends in contemporary ocean productivity. *Nature* 444:752–755.
15. Pigolotti S, Benzi R, Jensen MH, Nelson DR (2012) Population genetics in compressible flows. *Phys Rev Lett* 108:128102.
16. Benzi R, Nelson DR (2009) Fisher equation with turbulence in one dimension. *Phys D Nonlinear Phenom* 238:2003–2015.
17. Kimura M (1962) On the probability of fixation of mutant genes in a population. *Genetics* 47:713–719.
18. Maruyama T (1974) A simple proof that certain quantities are independent of the geographical structure of population. *Theor Popul Biol* 5:148–154.
19. Perlekar P, Benzi R, Nelson DR, Toschi F (2010) Population dynamics at high Reynolds number. *Phys Rev Lett* 105:144501.
20. Pigolotti S, et al. (2013) Growth, competition and cooperation in spatial population genetics. *Theor Popul Biol* 84:72–86.
21. Jacox MG, Hazen EL, Bograd SJ (2016) Optimal environmental conditions and anomalous ecosystem responses: Constraining bottom-up controls of phytoplankton biomass in the California current system. *Sci Rep* 6:27612.
22. Bravo L, Ramos M, Astudillo O, Dewitte B, Goubanova K (2016) Seasonal variability of the Ekman transport and pumping in the upwelling system off central-northern Chile ($\sim 30^\circ$ S) based on a high-resolution atmospheric regional model (WRF). *Ocean Sci* 12:1049–1065.
23. Pollard R, Regier L (1992) Vorticity and vertical circulation at an ocean front. *J Phys Oceanography* 22:609–625.
24. Moore L (1995) Comparative physiology of Synechococcus and Prochlorococcus: Influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Mar Ecol Prog Ser* 116:259–275.
25. Hallatschek O, Korolev K (2009) Fisher waves in the strong noise limit. *Phys Rev Lett* 103:108103.
26. Doering CR, Mueller C, Smereka P (2003) Interacting particles, the stochastic Fisher-Kolmogorov-Petrovsky-Piscounov equation, and duality. *Phys A Stat Mech Appl* 325:243–259.
27. Hartl DL, Clark AG (1997) *Principles of Population Genetics* (Sinauer Associates, Sunderland, MA).
28. Biferale L (2003) Shell models of energy cascade in turbulence. *Annu Rev Fluid Mech* 35:441–468.
29. Rice S (1955) Mathematical analysis of random noise. *Selected Papers on Noise and Stochastic Processes*, ed Wax N (Dover Publications, New York), pp 133–294.
30. Kailasnath P, Sreenivasan K (1993) Zero crossings of velocity fluctuations in turbulent boundary layers. *Phys Fluids A Fluid Dyn* 5:2879–2885.
31. Hallatschek O, Nelson DR (2008) Gene surfing in expanding populations. *Theor Popul Biol* 73:158–170.