



# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

## Gene Surfing in Expanding Populations

The Harvard community has made this article openly available.  
[Please share](#) how this access benefits you. Your story matters.

<b>Citation</b>	Hallatschek, Oskar, and David Nelson. 2008. Gene surfing in expanding populations. <i>Theoretical Population Biology</i> 73(1): 158-170.
<b>Published Version</b>	<a href="https://doi.org/10.1016/j.tpb.2007.08.008">doi:10.1016/j.tpb.2007.08.008</a>
<b>Accessed</b>	July 20, 2018 7:56:18 PM EDT
<b>Citable Link</b>	<a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:7982716">http://nrs.harvard.edu/urn-3:HUL.InstRepos:7982716</a>
<b>Terms of Use</b>	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a>

*(Article begins on next page)*

# Gene surfing in expanding populations

Oskar Hallatschek\* and David R. Nelson

Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

(Dated: February 6, 2008)

Spatially resolved genetic data is increasingly used to reconstruct the migrational history of species. To assist such inference, we study, by means of simulations and analytical methods, the dynamics of neutral gene frequencies in a population undergoing a continual range expansion in one dimension. During such a colonization period, lineages can fix at the wave front by means of a “surfing” mechanism [Edmonds C.A., Lillie A.S. & Cavalli-Sforza L.L. (2004) Proc Natl Acad Sci USA 101: 975-979]. We quantify this phenomenon in terms of (i) the spatial distribution of lineages that reach fixation and, closely related, (ii) the continual loss of genetic diversity (heterozygosity) at the wave front, characterizing the approach to fixation. Our simulations show that an *effective population size* can be assigned to the wave that controls the (observable) gradient in heterozygosity left behind the colonization process. This effective population size is markedly higher in *pushed* waves than in *pulled* waves, and increases only sub-linearly with deme size. To explain these and other findings, we develop a versatile analytical approach, based on the physics of reaction-diffusion systems, that yields simple predictions for any deterministic population dynamics.

Population expansions in space are common events in the evolutionary history of many species [1, 2, 3, 4, 5, 6, 7] and have a profound effect on their genealogy. It is widely appreciated that any range expansion leads to a reduction of genetic diversity (“Founder Effect”) because the gene pool for the new habitat is provided only by a small number of individuals, which happen to arrive in the unexplored territory first. In many species, the genetic footprints of these *pioneers* are still recognizable today and provide information about the migrational history of the species. For instance, a frequently observed south-north gradient in genetic diversity (“southern richness to northern purity” [8]) on the northern hemisphere is thought to reflect the range expansions induced by the glacial cycles. In the case of humans, the genetic diversity decreases essentially linearly with increasing geographic distance from Africa [2, 3], which is indicative of the human migration out of Africa. It is hoped [9], that the observed patterns of neutral genetic diversity can be used to infer details of the corresponding colonization pathways.

Such an inference requires an understanding of how a colonization process generates a gradient in genetic diversity, and which parameters chiefly control the magnitude of this gradient. Traditional models of population genetics [10], which mainly focus on populations of constant size and distribution, apply to periods before and after a range expansion has occurred, when the population is at demographic equilibrium. However, the spatio-temporal dynamics in the transition period, on which we focus in this article, is less amenable to the standard analytical tools of population genetics, and has been so far studied mostly by means of simulations [11, 12, 13, 14, 15]. An analytical understanding is available only for a linear stepping stone model in which demes (lattice sites)

are colonized one after the other, following deterministic logistic growth [16] or instantaneously [17], in terms of recurrence relations.

Recent computer studies suggest that the neutral genetic patterns created by a propagating population wave might be understood in terms of the mechanism of “gene surfing” [13, 14]: As compared to individuals in the wake, the pioneers at the colonization front are much more successful in passing their genes on to future generations, not only because their reproduction is unhampered by limited resources but also because their progeny start out from a good position to keep up with the wave front (by means of mere diffusion). The offspring of pioneers thus have a tendency to become pioneers of the next generation, such that they, too, enjoy abundant resources, just like their ancestors. Therefore, pioneer genes have a good chance to be carried along with the wave front and attain high frequencies, as if they “surf” on the wave. Thus, the descendants of an individual sampled from the tip of the wave have a finite probability to take over the wave front. In this case of “successful surfing”, further colonization will produce only descendants of the relevant pioneer because the wave front has been “fixed”. The process of fixation at the front of a one-dimensional population wave is illustrated in Fig 1.

The present study hinges on the question as to *where* lineages that reach fixation originate within the wave front. Clearly, the probability of successful surfing must increase with the proximity to the edge of the wave [14]. On the other hand, more surfing *attempts* originate from the bulk of the wave where the population density is larger. We show that, due to this tradeoff, the origins of successful lineages have a bell-like distribution inside the wave front. Furthermore, this ancestral probability distribution, together with the population-density profile of the population wave itself, is found to control the observable gradients in genetic diversity. The genetic pattern directly behind the moving colonization front turns out to mimic that of a small well-mixed (*panmictic*) population.

---

\*To whom correspondence should be addressed. E-mail: ohalats@physics.harvard.edu

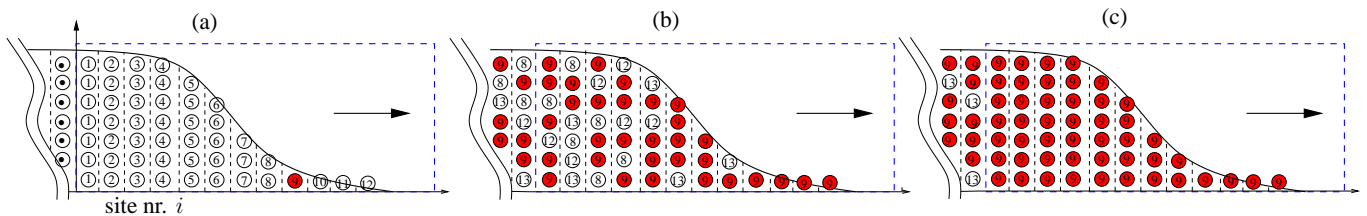


FIG. 1: Illustration of gene surfing by means of three consecutive snapshots of the genetic composition at the edge of an expanding population (in one spatial dimension). (a) A neutral *red* mutant arises at the wave front. (b) After some time, the genetic make-up at the wave front is drastically changed due to random number fluctuations and it is apparent that descendants of *red* will take over the wave front. (c) Fixation in the co-moving frame of descendants of *red*. Numbers in these sketches represent “inheritable” labels that are used in our simulations to trace back the spatial origin of individuals in the wave front. In this example, descendants of *red* are associated with position “9” in the co-moving frame. The dashed blue frame indicates the co-moving simulation box.

The *effective* size  $N_e$  of this population “bottleneck” is shown to be *smaller* than the typical number of individuals in the colonization front and very sensitive to the growth conditions in the very tip of the front. Colonization fronts in which individuals need to be accompanied by others in order to grow (*Allee* effect [18]) have a much larger effective population size than those in which individuals grow even if they are isolated from the rest.

The outline of the paper is as follows. We first introduce a stochastic computer model that we use to generate both *pulled* and *pushed* one-dimensional colonization waves. Tracer experiments within this model are then used to reveal the probability distribution of successful surfers and the decrease of genetic differentiation at the colonization front. Our succeeding theoretical treatment reveals to what extent both measures are related, and how they can be predicted for continuous models with quasi-stationary demography. After a comparison between theory and simulations, we discuss the significance of our results in the light of inferring past range expansions from spatially resolved genetic data.

## I. SIMULATIONS

Most models for range expansions can be classified as describing *pulled* or *pushed* population fronts [19, 20]. The distinction between the two cases corresponds to a difference in behavior. Suppose individuals need to be in proximity to other individuals in order to grow in number (*Allee* effect [18]). The presence of conspecifics can be beneficial due to numerous factors, such as predator dilution, antipredator vigilance, reduction of inbreeding and many others [21]. Then, the individuals in the very tip of the front do not count so much, because the rate of reproduction decreases when the number density becomes too small. Consequently, the front is *pushed* in the sense, that its time-evolution is determined by the behavior of an ensemble of individuals in the boundary region. On the other hand, a population in which an individual reproduces, even if it is completely isolated from the rest, will be “spearheaded” by these front individu-

als. These *pulled* fronts are responsive to small changes in the frontier and, therefore, are prone to large fluctuations [20].

One might suspect that the genetic pattern left behind a population wave should reflect whether the colonization process is controlled by a small or large number of individuals. Hence, we have set up a computer model that allows us to investigate the surfing dynamics for both classes of waves.

### A. Population dynamics

The population is distributed on a one dimensional lattice, whose sites (demes) can carry at most  $N$  individuals. The algorithm effectively treats individuals ( $\bullet$ ) and vacancies ( $\circ$ ) as two types of particles, whose numbers must sum to  $N$  at each site. A computational time step consists of two parts: (i) a migration event, in which a randomly chosen particle exchanges place with a particle from a neighboring site. This step is independent of the involved particle types. (ii) A duplication *attempt*: Two particles are randomly chosen (with replacement) from the same lattice site. A duplicate of the first one replaces the second one ( $1^{\text{st}} \rightarrow 2^{\text{nd}}$ ) with probabilities based on their identities: proposed replacements  $\circ \rightarrow \circ$ ,  $\bullet \rightarrow \bullet$  and  $\bullet \rightarrow \circ$  (growth) are realized with probability 1, whereas  $\circ \rightarrow \bullet$  (death) is carried out only with probability  $1 - s$ , depending on a growth parameter  $0 < s < 1$ . This asymmetry controls the effective local growth advantage of  $\bullet$  over  $\circ$ .

In terms of individuals and vacancies instead of particles, we see that our model describes migration and local logistic growth of a population distributed over demes with carrying capacity  $N$ . Starting with a step-function initial condition, the simulation generates an expanding *pulled* population wave. The above algorithm represents a discretized version [22] of the stochastic Fisher-Kolmogorov equation [23] with a Moran-type of breeding scheme [10]. To generate *pushed* waves as well, we extend our model by the following rule: In demes in which the number of individuals falls to  $N_c$  or below, we set their

effective linear growth rate  $s$  to zero. This represents, for  $N_c > 0$ , a simple version of the above mentioned Allee effect of a reduced growth rate when the population density is too small.

## B. Tracer dynamics

Tracer experiments within this computer model allow us to extract the genealogies of front individuals. After the population had enough time to relax into its propagating equilibrium state, all individuals are labeled according to their current position  $i \in \{1 \dots n\}$  within the simulation box of length  $n$ , see Fig. 1a. These labels are henceforth inherited by the descendents, which thereby carry information about the spatial position of their ancestors. The randomness in the reproduction and migration processes (genetic drift) during the succeeding dynamics inevitably leads to a reduction in the diversity of labels present in the simulation box, see Fig. 1b. Labels are lost due to either extinction or because they cannot keep up with the simulation box, which *follows* the propagating wave front[44].

In our simulation, the gradual loss of diversity of labels at the wave front is measured by the quantity

$$H(t) = \sum_{i=1}^n p_i(t) [1 - p_i(t)] , \quad (1)$$

which depends on the frequency  $p_i(t)$  of label  $i$  at time  $t$  after the wave has been labeled.  $H(t)$  represents the time-dependent probability that two individuals, randomly chosen from the bounded simulation box, carry different labels. Provided that mutations are negligible on the time-scale of the range expansion, we may think of our inheritable labels as being neutral genes at one particular locus (alleles). We may thus identify  $H(t)$  with the probability that two alleles randomly chosen from the front region are different conditional on the well-mixed labeling state at  $t = 0$  imposed by our simulation. Hence, we refer to  $H(t)$  as the time-dependent *expected heterozygosity* [10] at the wave front [45].

The perpetual loss of labels in our model without mutations eventually leads to the fixation of one label in the simulation box, see Fig. 1c. The value of this label indicates the origin *within the co-moving frame* of this successful “surfer”. It contributes one data point to the spatial distribution  $P_i$  of individuals whose descendents came to fixation. After fixation, the algorithm proceeds with the next labeling event.

## C. Results

The parameters of our computer models are the deme size  $N$ , i.e. the maximal number of individuals per lattice site, the linear growth rate  $s$  per generation, and the critical occupation number  $N_c$ , below which the growth rate

drops to zero (Allee effect). In our simulations, we set  $s = 0.1$  throughout, and determine, for varying  $N$  and  $N_c$ , the averages of the ancestral distribution function  $\langle P_i \rangle$ , the scaled occupation number  $\langle n_i \rangle / N$ , both being functions of the lattice site  $i$  in the co-moving frame, and the time-dependent probability of non-identity,  $\langle H(t) \rangle$ . Here, angle brackets indicate that the enclosed quantities have been averaged in time, i.e. over many fixation events, *and* over multiple realizations of the same computer experiment[46].

Figure 2 illustrates the relation between the front profiles  $\langle n_i \rangle / N$  and the ancestral distribution  $\langle P_i \rangle$  in the co-moving frame. Whereas the wave profiles have the familiar sigmoidal shapes of reaction-diffusion waves [19, 20], the ancestral distribution functions are bell-curves with most of its support beyond the inflection point of the wave front. The fact that  $\langle P_i \rangle$  has a maximum inside the wave front reflects a tradeoff, mentioned earlier, between a larger fixation probability in the tip of the wave versus a larger number of surfing *attempts* originating from the bulk. Notice from Fig. 2a that, for increasing deme size, the distribution becomes wider and shifts further into the tip of the wave, which is in contrast to the almost  $N$ -independent scaled wave profiles. Fig. 2b shows that the opposite effect is caused by increasing the cutoff value  $N_c$ , which changes the type of the wave from *pulled* to *pushed*.

Next, we measured the temporal decay of the heterozygosity  $H(t)$ , defined in Eq. (1). In Fig. 3, time-traces of  $H(t)$  are depicted for various parameters and show an exponential decay after an initial transient. This allows us to characterize the strength of genetic drift at the wave front by a single number, the (asymptotic) exponential decay rate,  $-\partial_t \log \langle H(t) \rangle$ , which can be extracted from logarithmic plots of  $\langle H(t) \rangle$ . By analogy with well-mixed (panmictic) populations, in which the heterozygosity decays exponentially with rate  $2/N$  (Moran model[10]), it is convenient to express the decay rate by  $2/N_e$ , in terms of an *effective population size*  $N_e$ . The theoretical part below will further clarify to what extent the genetic diversity at the wave front mimics that of a population “bottleneck” of constant size  $N_e$ .

Figure 4 depicts  $N_e$  as a function of the deme size  $N$  on a double logarithmic scale for  $N_c = 0$  and  $N_c = 10$ . Naively, one might expect  $N_e$  to be, roughly, the characteristic number of individuals in the width of the wave front, since these individuals contribute (by growing) to the advance of the wave. Thus, a linear relationship between deme and effective population size would not be surprising. In contrast, we find that  $N_e$  increases much slower than linearly with increasing deme size. Furthermore, the effective population size turns out to be very sensitive to the presence of an Allee effect ( $N_c > 0$ ), which has the ability strongly increase the effective population size. This point is illustrated, in particular, by the inset of Fig. 4 which depicts the effective population size  $N_e$  in a simulation of fixed deme size ( $N = 1000$ ) and varying strength of the Allee effect ( $10 < N_c < 500$ ).

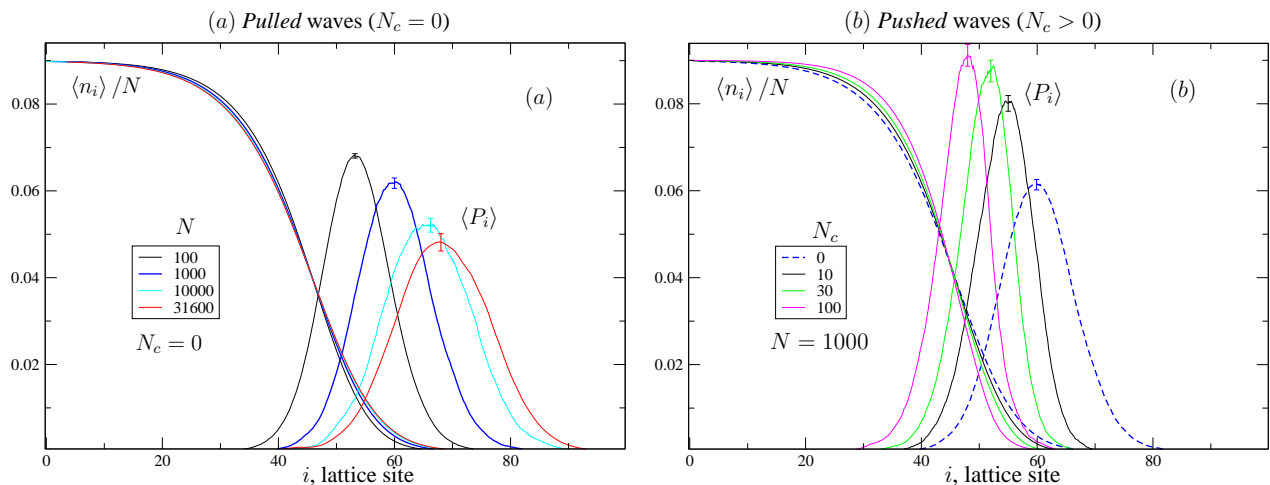


FIG. 2: Measured distributions  $\langle P_i \rangle$  (bell-curves) of “successful surfers” together with the normalized occupation numbers  $\langle n_i \rangle / N$  (sigmoidal curves; scaled along the vertical axis to fit the figure) as a function of the site number  $i$  in the co-moving frame; (a) for *pulled* waves ( $N_c = 0$ ) with varying deme sizes  $N$ ; (b) for various *pushed* waves ( $N_c > 0$ ) with deme size  $N = 1000$  compared to the corresponding *pulled* wave (dashed blue lines), which is also present in (a).

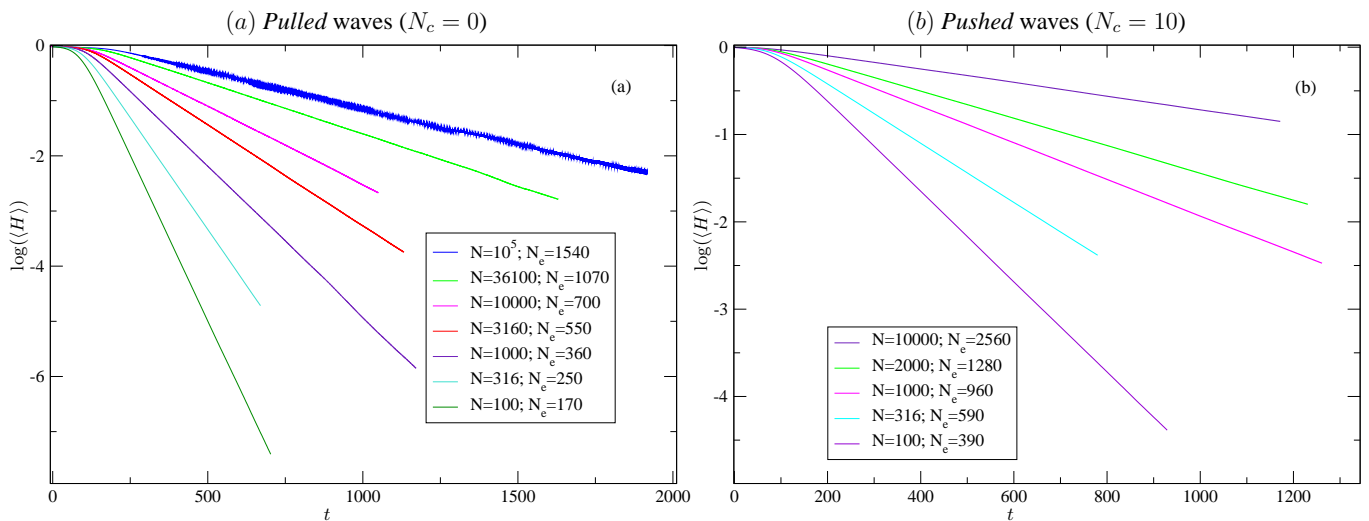


FIG. 3: The decay of genetic diversity,  $\langle H(t) \rangle$ , with time (in units of generations) on a log-linear scale for varying deme sizes; (a) for *pulled* waves ( $N_c = 0$ ); (b) for *pushed* waves with  $N_c = 10$ . Both cases show, that an asymptotic exponential decay of  $\langle H(t) \rangle$  is reached after an initial transient where the decay is weak. The duration of this transient is dependent on the size of the simulation box: The larger the simulation box, the larger the time until the exponential decay is approached. The asymptotic exponential decay rate, however, has been checked to approach a constant for a sufficiently large box size. This exponential decay rate is therefore well-defined and can be used to characterize the decrease of genetic diversity at the wave front. By analogy with panmictic populations, in which the heterozygosity decays exponentially with rate  $2/N$  (Moran model), it is convenient to express the decay rate as  $2/N_e$ , i.e., in terms of an *effective population size*  $N_e$ , which is noted in the legends, and plotted in Fig. 4.

Qualitatively this phenomenon may be explained with the *pushed* nature of these waves. An Allee effect shifts the distribution  $P_i$  of successful surfers away from the tip towards the wake of the wave (see Fig. 2b) and hence increases the gene pool from which the next generation of pioneers is sampled. This argument indicates a close relation between the  $N_e$  and  $P_i$ , which also emerges explicitly in the theoretical analysis below.

## II. THEORY

The following employs a continuous reaction diffusion approach to establish a theoretical basis for the relation between the neutral genetic diversity and the population dynamics in non-equilibrium situations like range expansions. It will help us to reconcile the somewhat surprising response of our simulations to parameter changes

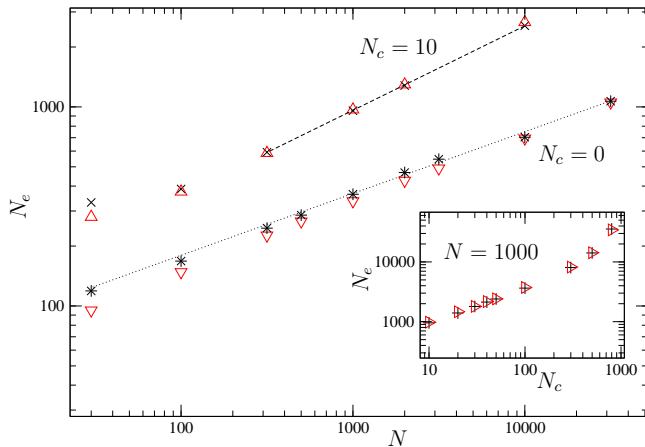


FIG. 4: The measured effective population size  $N_e$  as a function of deme size  $N$  on a log-log scale for *pulled* waves ( $N_c = 0$ , asterisks) and *pushed* waves ( $N_c = 10$ , crosses). The dashed and dotted lines have slope .30 and .42, respectively, i.e., significantly smaller than 1. Triangles represent the effective population sizes as inferred from the strong-migration approximation, Eq. (8), using the *measured* ( $P$ )–distribution and population profiles. The inset shows the behavior of  $N_e$  for varying cutoff-value  $N_c$  and fixed deme size  $N = 1000$ , again, on a log-log scale.

(deme size and Allee effect). Note from Fig. 2 that the changes in the ancestral distribution are dramatic, while the changes in the population profile itself are quite modest. Results obtained from our approximation scheme are tested by direct comparison of simulations and theory.

### A. Gene surfing

In our simulations, as well as in many other models of range expansions, a propagating population wave results from the combination of random short-range migration and logistic local growth. In the continuum limit, a general coarse-grained continuum description of such a reaction-diffusion system of a single species is given by

$$\partial_t c(x, t) = D \partial_x^2 c(x, t) + v \partial_x c(x, t) + K(x, t) \quad (2)$$

formulated in the frame co-moving with velocity  $v$ , where  $c(x, t)$  represents the density of individuals at location  $x$  at time  $t$  and  $D$  is a diffusivity. The first two terms on the right hand side represent the *conservative* part of the population dynamics, for which we make the usual diffusion assumptions [24]. The reaction term  $K(x, t)$  accounts for both deterministic and stochastic fluctuations in the number of individuals due to birth and death processes, and typically involves non-linearities such as a logistic interaction between individuals as well as noise caused by number fluctuations. For instance, our computer model with  $N_c = 0$  maps, in the continuum limit [22], to the stochastic Fisher equation, for which  $K(x, t) = sc(c_\infty - c) + \epsilon \sqrt{c(c_\infty - c)} \eta$ , where  $\eta(x, t)$  is a

Gaussian white noise process in space and time,  $c_\infty \propto N$  is the carrying capacity and  $\epsilon \propto \sqrt{N}$  sets the strength of the noise. We would like to stress, however, that the following analysis does not rely on a particular form of  $K$ . Therefore, we leave the reaction term unspecified.

As in our tracer experiments, let us assume that inheritable labels, representative of neutral genes, are attached to individuals within the population and ask: Given Eq. (2) is a proper description of the population dynamics, to what extent is the dynamics of these labels determined? To answer this question, it is convenient to adopt a retrospective view on the tracer dynamics. Imagine following the ancestral line of a single label located at  $x$  *backwards* in time to explore which spatial route its ancestors took. This backward-dynamics of a single line of descent will show drift and diffusion *only*; any reaction is absent because among all the individuals living at some earlier time there must be exactly one ancestor from which the chosen label has descended from. We may thus describe the ancestral process of a single lineage by the probability density  $G(\xi, \tau|x, t)$  that a label presently, at time  $t$  and located at  $x$ , has descended from an ancestor that lived at  $\xi$  at the earlier time  $\tau$ . In this context, it is natural to choose the time as increasing towards the past,  $\tau > t$ , and to consider  $(\xi, \tau)$  and  $(x, t)$  as *final* and *initial* state of the ancestral trajectory, respectively. With this convention, the distribution  $G$  satisfies the initial condition  $G(\xi, t|x, t) = \delta(x - \xi)$ , where  $\delta(x)$  is the Dirac delta function, and is normalized with respect to  $\xi$ ,  $\int G(\xi, \tau|x, t) d\xi = 1$ .

Since  $G(\xi, \tau|x, t)$  as function of  $\xi$  and  $\tau$  is a probability distribution function generated by a diffusion process that is continuous in space and time, we expect its dynamics to be described by a generalized diffusion equation (Fokker-Planck equation [24]). Indeed, in the Appendix A we show that,  $G(x, t|\xi, \tau)$  obeys

$$\begin{aligned} \partial_\tau G(\xi, \tau|x, t) &= -\partial_\xi J(\xi, \tau|x, t) \\ J(\xi, \tau|x, t) &\equiv -D \partial_\xi G + \{v + 2D \partial_\xi \ln[c(\xi, \tau)]\} G, \end{aligned} \quad (3)$$

where all derivatives are taken with respect to the ancestral coordinates  $(\xi, \tau)$ . The drift term in Eq. (3) has two antagonistic parts. The first term,  $v$ , tends to push the lineage into the tip of the wave, and is simply a consequence of the moving frame of reference. The second term proportional to *twice* the gradient of the logarithm of the density is somewhat unusual. It accounts for the purely “entropical” fact that, since there is a forward-time flux of individuals diffusing from regions of high density to regions of low density, an ancestral line tends to drift *into* the wake of the wave where the density is higher.[47]

Our computer experiments measure the spatial distribution  $P$  of the individuals whose descendants came to fixation. This information is encoded in the long-time behavior of  $G$

$$P(\xi, \tau) = \lim_{t \rightarrow -\infty} G(\xi, \tau|x, t), \quad (4)$$

because it represents the probability that, in the far future ( $t \rightarrow -\infty$ , in our notation) when the population is fixed, an individual of the extant lineage has descended from an individual who lived at location  $\xi$  of the co-moving frame at time  $\tau$ . Equation (4) has to be independent of  $x$  if fixation occurs: For  $t \rightarrow -\infty$ , all individuals irrespective of their position  $x$  must have descended from the same ancestor and, thus, from the same location at the earlier time  $\tau$ .

In principle, it is thus possible to relate the ancestral distribution to the population dynamics by solving Eq. (3) in the long-time limit. Unfortunately, this task is usually difficult to achieve analytically because the number fluctuations in the density  $c(\xi, \tau)$  of the total population add noise to the drift term in Eq. (3). As is customary in many spatially explicit models of population genetics, let us suppose, however, that rules of “strict density regulation” [26] are imposed in order to guarantee a stationary demography, so that in the co-moving frame,

$$c(x, t) \approx c_{\text{st}}(x). \quad (5)$$

Even though real systems and our discrete particle simulations exhibit density fluctuations even in equilibrium, we take Eq. (5) as a first approximation in cases where the total number of particles is large enough, such that the relative magnitude of the density fluctuations is small (law of large numbers). We will call assumption Eq. (5) the “deterministic approximation” as it neglects stochastic fluctuations in the total population density.

With Eq. (5), all parameters in the Fokker-Planck equation, Eq. (3), of the ancestral distribution  $G(x, t|\xi, \tau)$  are time-independent and its analysis considerably simplifies: If a unique stationary solution  $P_{\text{st}}(\xi) \equiv \lim_{t \rightarrow \infty} G(\xi, \tau|x, t)$  exists, it can be written explicitly in terms of the stationary density profile  $c_{\text{st}}(\xi)$ ,

$$P_{\text{st}}(\xi) \propto c_{\text{st}}^2(\xi) \exp(v\xi/D) \quad (6)$$

where a pre-factor is required to satisfy the normalization condition of  $P_{\text{st}}$ ,  $\int P_{\text{st}}(\xi, \tau|x, t) d\xi = 1$ . [48]

As shown below, the analytical expression Eq. (6) describes at least qualitatively the bell-like shapes found for the ancestral distribution function  $\langle P_i \rangle$  in our stochastic simulations. The exponential factor biases the fixation probability [49] towards the tip of the wave ( $\xi > 0$ ) and competes with the pre-factor controlled by the decaying density of individuals in the tip of the wave.

It is noteworthy that Eq. (6) not only applies to range expansions, but can be evaluated for any deterministic population dynamics, such as deterministic models of evolution [27, 28] (where however rare events might be crucial as found in Ref. [29]) and to source-sink populations [30, 31], a simple example of which is given in the Appendix B. If the spatial domain is unbounded, Eq. (6) yields finite results as long as  $c_{\text{st}}(\xi)$  decays faster than  $\exp[-v\xi/(2D)]$  as  $\xi \rightarrow \infty$ . This condition formally distinguishes the two classes of waves earlier denoted by

*pulled* and *pushed*. Within the mean-field description of such waves, the right hand side of Eq. (6) is normalizable only in the case of *pushed* waves [32]. The density of *pulled* waves, however, decays as  $\exp[-v\xi/(2D)]$  in the foot of the wave ( $\xi \rightarrow \infty$ ) as follows from a linearized mean field treatment [33]. The prime example of *pulled* waves, the mean-field Fisher wave, does not therefore allow for successful surfing,  $P_{\text{st}} \equiv 0$ , as is explicitly shown in Appendix C. This is in marked contrast to our simulations of *stochastic* Fisher waves (Fig. 2a). There, we found finite bell-like ancestral distributions up to deme sizes on the order of  $10^5$ . This striking discrepancy indicates that the classical Fisher equation is a poor approximation for the case of finite deme sizes (even if they are large). An improved deterministic equation with a modified reaction term has been proposed [34], which is able to reproduce the leading reduction in wave velocity due to the discreteness. A remarkable property of Eq. (6) is that it should be valid, *irrespective* of the actual form of the reaction term, if the demography is deterministic. By comparing Eq. (6) to simulations, it is possible to test the deterministic character of a population wave, i.e., whether or not a deterministic reaction-diffusion description might be appropriate.

For our simulations, such a test is given in Fig. 5, where we superimpose measured ancestral distribution functions  $\langle P_i \rangle$  with those predicted by Eq. (6) based on the measured wave velocity  $v$  and the occupation numbers  $\langle n_i \rangle$  (the discrete analog of the population density  $c_{\text{st}}(\xi)$ ). It is seen that systematic deviations occur in the *pulled* case ( $N_c = 0$ ), where the predicted distribution seems to be somewhat displaced towards the tip of the wave. The agreement of theory and simulation is much better for  $N_c = 10$  and further improves when  $N_c$  is increased. Altogether, our deterministic approximation Eq. (6) seems to apply best to *pushed* waves with a strong Allee effect ( $N_c \gg 1$ ), whereas significant deviations to Eq. (6) occur for *pulled* waves. An alternative test of theory and simulation, presented in Fig. 6, supports this conclusion and furthermore shows that the deterministic approximation applied to *pulled* waves improves slowly with increasing deme size. For reasonable system sizes, however, fluctuation effects in *pulled* waves are non-negligible [19].

## B. Decrease of genetic diversity

To measure how fast genetic diversity decreases at the wave front due to gene surfing, we also studied in our simulations  $H(t)$ , defined in Eq. (1) as the probability that two randomly sampled individuals carry different labels at a time  $t$  after a labeling event. To what extent are the decrease of  $H(t)$  and the shape of  $P(\xi)$  related? As before, a retrospective view on the problem simplifies the theoretical analysis. Imagine following the lineages of two randomly sampled individuals backward in time. They will drift and diffuse separately for a certain time  $t_c$

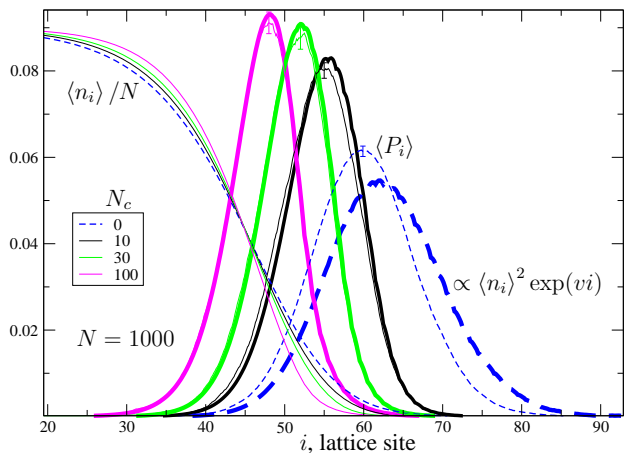


FIG. 5: The data from Fig. 2b (thin lines) superimposed with the prediction Eq. (6) (thick lines) for the ancestral distribution function  $\langle P_i \rangle$  based on the measured wave profile  $\langle n_i \rangle$  and velocity  $v$ . As explained in the text, the apparent systematic deviations in the *pulled* case ( $N_c = 0$ ) are caused by fluctuations in the tip of the wave.

until they *coalesce* in the most recent common ancestor. If the last labeling event occurred at an earlier time  $t > t_c$  (reversed time-direction) then both individuals must carry identical labels. If, on the other hand,  $t < t_c$  then these individuals will have different labels unless their different ancestors happen to be in the same deme at the labeling time. Up to a small error of the order of the inverse size of the simulation box, we may thus identify the probability  $H(t)$  of two individuals carrying different labels with the probability that their coalescence time  $t_c$  is larger than  $t$ .

In principle, the coalescence time distribution of two lineages can be explored by studying the simultaneous backward-diffusion process of two lineages conditional on having not coalesced before [35]. This process is described by a generalization of Eq. (2) augmented by a well-known sink term [35] that accounts for the probability of coalescence when lineages meet.

Here, we describe a (more tractable) approximation that estimates the behavior of  $H(t)$  from the distribution  $P(\xi)$ , analyzed in the previous section. It is based on the assumption that the coalescence rate of two lineages is so small that each lineage has enough time to equilibrate its spatial distribution before coalescence occurs. Under this quasi-static approximation, the behavior of  $H(t)$  is described by [36]

$$\partial_\tau H \approx -2H \int \frac{P^2(\xi, \tau)}{c(\xi, \tau)} d\xi, \quad (7)$$

when time is measured in units of generation times. The justification of Eq. (7) is as follows: The coalescence rate,  $-\partial_\tau H$ , at time  $\tau$  in the past is given by the probability that the two lineages have not coalesced earlier,  $H(\tau)$ , times the rate at which two separate lineages coalesce at time  $\tau$ . The latter is locally proportional to the product

of the probabilities that the two lineages meet at the same place,  $\propto P^2(\xi, \tau)$ , and that they meet in the same individual,  $\propto c^{-1}$ , given they are at the same place. Less obvious, unfortunately, is the numerical pre-factor “2” on the right-hand side, which is specific to the employed breeding scheme (Moran model[10]).

Eq. (7) yields the correct coalescence time distribution in the so-called *strong migration limit* [36, 37] of large population densities,  $c \rightarrow \infty$ , while the diffusivity  $D$  and the spatial extension of the habitat are held fixed. In our case, it serves as a simple approximation that tends to overestimate coalescence rates, because it neglects spatial anti-correlations between non-coalescing lineages: Lineages that have avoided coalescence will usually be found further apart than described by the product of (one-point) distribution functions in Eq. (7). Thus, their rate of coalescence will, typically, be smaller than in Eq. (7).

Equation (7) predicts exponential decay,  $H(t) \sim \exp[-2(\tau - t)/N_e]$ , in the deterministic approximation, Eq. (5), with a rate depending on a constant  $N_e$  given by [36]

$$N_e^{-1} = \int \frac{P_{\text{st}}^2(\xi)}{c_{\text{st}}(\xi)} d\xi. \quad (8)$$

In fact, a generalization of this argument to the coalescence process of a sample of  $n$  lineages shows that the *standard coalescent* [38] is obtained in the strong-migration limit with the parameter  $N_e$  interpreted as the *effective population size* [36]. In other words, the coalescence process in the strong-migration limit is identical, in every respect, to the coalescence of a well-mixed population of fixed size  $N_e$ .

The strong-migration approximation may be tested by comparing the effective population sizes measured in our simulations with the ones predicted by Eq. (8) based on the measured ancestral distribution  $\langle P_i \rangle$  and the number density profile  $\langle n_i \rangle$ . These inferred values are plotted in Fig. 4 as (red) triangles. For both *pushed* and *pulled* waves, the agreement between inferred and measured effective population sizes becomes excellent for large deme sizes,  $N > 100$ . For lower values of  $N$  the strong migration assumption overestimates genetic drift, presumably due to the neglect of correlations as mentioned above.

### III. DISCUSSION

We have studied the impact of a range expansion on the genetic diversity of a population by means of simulations and analytical techniques. The one dimensional case treated in this article applies to populations following a (possibly curved) line, like a migration route, coast line, river or railway track. We have further simplified our analysis by neglecting habitat boundaries, which is appropriate for describing the colonization period, as long as the wave front is sufficiently far away from the boundaries.



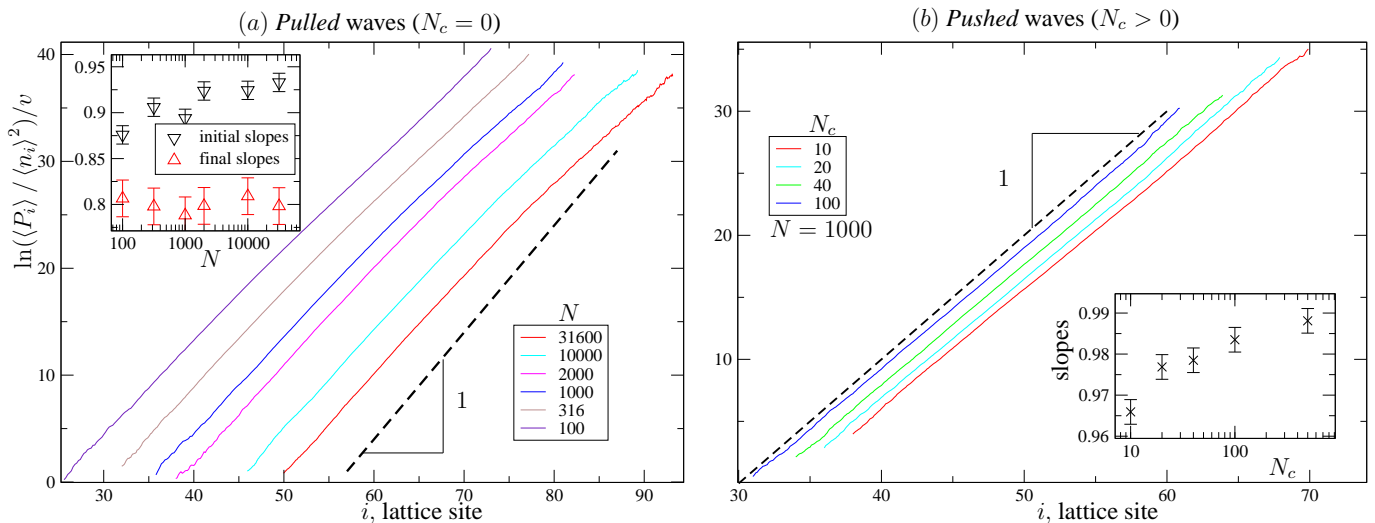


FIG. 6: The quantity  $\ln(\langle P_i \rangle / \langle n_i \rangle^2) / v$  as a function of lattice site  $i$  in the co-moving frame, which should have slope 1 according to the deterministic approximation, Eq. (6). (a) Results for *pulled* waves ( $N_c = 0$ ) with various deme sizes  $N$ ; the values for the “initial slope” and “final slope” have been obtained by fitting a straight line to the lower and upper half of each shown curve, respectively. (b) Results for various *pushed* waves ( $N_c > 0$ ) with fixed deme size,  $N = 1000$ . (The domain of each curve is restricted to a region, in which the bell-like distributions  $\langle P_i \rangle$  has enough support to sample a sufficient amount of data points. This region roughly covers 98% of all successful surfing events.)

Our findings suggest the following general scenario. Suppose, an initially well-mixed population increases its range from a smaller to a larger habitat, and that mutations may be neglected on the time scale of the range expansion. Our simulations show that the heterozygosity at the moving colonization front decays, due to genetic drift, exponentially in time with a rate  $2/N_e$ , depending on an *effective* population size  $N_e$ . Upon combining this rate with the velocity  $v$  (per generation) of the colonization front, we obtain a length  $\lambda = vN_e/2$  that characterizes the pattern of genetic diversity generated by the colonization process: As the wave front moves along it leaves behind saturated demes with heterozygosities given by the value of the front heterozygosity as the wave front passes through. The transient colonization process therefore engraves a spatially decreasing profile of heterozygosity into the newly founded habitat. This profile decays exponentially in space on the characteristic length  $\lambda$  and serves as an initial condition for the succeeding period of demographic equilibration, which may be described by traditional models of population genetics [10].

Of critical importance for the interpretations of gene frequency clines in natural populations is the question as to which parameters chiefly control  $\lambda$ . Our computer simulations have revealed, that the effective population size, and thus  $\lambda$ , only grows *sub-linearly* with increasing deme size, in contrast to the naive expectation that the effective population size should roughly be given by the characteristic number of individuals contained in the wave front (deme size times the width of the wave). On the other hand, we found that the population “bottleneck” at the wave front was significantly widened when we implemented an Allee effect [18] into our model, by

which growth rates for small population densities are decreased. As a consequence, the region of major growth shifted away from the frontier into the bulk of the wave and, thus, the effective size of the gene pool for the further colonization was increased.

Finally, we have developed a theoretical framework to study the backward-dynamics of neutral genetic markers for a given non-equilibrium population dynamics, which is summarized by the Fokker-Planck equation (3) and its generalization in Appendix A. In stationary populations, it leads to a simple expression, Eq. (6), for the long-time probability distribution of the common ancestors, which can be used, in the strong-migration limit to determine the effective population size, via Eq. (8). Comparison with our stochastic simulations reveals that the simple deterministic results are good approximations to stochastic simulations in the case of *pushed* waves (strong Allee effect), but that significant deviations occur for *pulled* waves due to the fluctuations in the frontier of the population wave. For sufficiently large deme sizes, the effective population size could, in both cases, be inferred with remarkable accuracy from Eq. (8).

## Acknowledgments

This research was supported by the German Research Foundation through grant no. Ha 5163/1 (OH). It is a pleasure to acknowledge conversations with Michael Brenner, Michael Desai and John Wakeley.

## APPENDIX A: BASIC FORMALISM

In this appendix, we describe the derivation of the Fokker–Planck equation, Eq. (3), and its generalization to heterogeneous migration and higher dimensions.

Central to our analysis is the assumption that the total population consists of statistically identical entities, such that two different individuals at a given time and location behave in the same way. In particular, we assume that migration as well as reproduction of an individual are *exchangeable* [39] random processes, i.e., independent of any label that might be assigned to it.

Let us indeed imagine a subpopulation labeled by a neutral marker, and ask: What is the dynamics of the labeled individuals for a *given* dynamics of the total population, as described by Eq. (2)? It is clear that the density  $c^*(x, t)$  of labels obeys a reaction-diffusion equation with coefficients  $D^*$ ,  $v^*$  and  $K^*$  that are closely related to the one for total population by means of the exchangeability assumption. Firstly, labeled and unlabeled individuals should migrate statistically in the same way, which is measured by diffusion and drift coefficients, i.e., we have  $D^* = D$  and  $v^* = v$ . Secondly, the labeled subpopulation must carry the number fluctuations of the total population, encoded in  $K$ , in proportion to its reduced size:  $K^* = Kc^*/c$ . However, these statements are true only *on average*: The discreteness of the particle numbers lead to fluctuations, for instance, in the quantity  $K^* - (c^*/c)K$ . To illustrate this point, imagine that at a given time, the term  $K$  dictates that an individual dies in some small spatio-temporal region, then this individual has to be sampled from the labeled subpopulation with probability  $c^*/c$ . The fluctuations of  $c^*$  due to this sampling procedure represents random genetic drift and must have zero mean according to the exchangeability assumption. Similar fluctuations affect the migration currents of the labeled subpopulation. Thus, only upon averaging over this source of stochasticity, we may formulate a reaction-diffusion equation of the form

$$\partial_t \bar{c}^* = D \partial_x^2 \bar{c}^* + v \partial_x \bar{c}^* + K \frac{\bar{c}^*}{c} \quad (\text{A1})$$

for the average density  $\bar{c}^*(x, t)$  of labeled individuals. In Eq. (A1), only the effect of the genetic drift of labeled individuals within the total population has been averaged out, the number fluctuations affecting the total population density are retained through the fluctuating reaction term  $K$ . In other words, Eq. (A1) describes the behavior of labeled individuals, if we average over many realizations conditional on a given fixed evolution of the total population, described by Eq. (2). Note that, this averaging “works” because Eq. (A1) is linear in  $c^*$ , such that a noise term with zero mean added to the right-hand-side to account for the genetic drift can be averaged out without generating higher moments.

Next, we use

$$K(x, t) = \partial_t c(x, t) - D \partial_x^2 c(x, t) - v \partial_x c(x, t)$$

from Eq. (2) to substitute the reaction term  $K$  in Eq. (A1). After rewriting the average density  $\bar{c}^*(x, t) \equiv \bar{p}(x, t)c(x, t)$  of labeled individuals in terms of their average frequency (or ratio)  $\bar{p}$ , we obtain

$$\partial_t \bar{p}(x, t) = D \partial_x^2 \bar{p} + \{v + 2D \partial_x \ln[c(x, \tau)]\} \partial_x \bar{p}. \quad (\text{A2})$$

Notice that,  $\bar{p} = \text{const.}$  is a (steady state) solution of Eq. (A2). Relations formally equivalent to Eq. (A2) have been formulated in Refs. [40, 41] in different contexts, in which, unlike the present case, random genetic drift could not be averaged out due to non-linearities, but had to be disregarded, instead. As a transport equation for *deterministic* gene frequencies, an equation similar to Eq. (A2) was also recently obtained in Ref. [25].

For a given realization of the time-evolution of the total population, the quantity  $\bar{p}(x, t)$  can be interpreted as the probability that an individual sampled from  $(x, t)$  is labeled, i.e., that it has descended from the initially labeled population. If the above tracer experiment for a *given* dynamics of  $c(x, t)$  is repeated multiple times,  $\bar{p}(x, t)$  represents the histogram of the number of times a individual sampled from  $(x, t)$  is labeled.

By choosing proper initial conditions, this allows us to study “where individuals come from”: Suppose that, at time  $\tau$ , the labeled population contains all individuals within a small interval around the position  $\xi$ . The solution of Eq. (A2) for later times will then tell us the probability that a individual at  $(x, t)$  has descended from an ancestor sampled from that narrow region around  $\xi$  at time  $\tau$ .

Hence, the probability density  $G(\xi, \tau|x, t)$ , introduced above Eq. (3), that an individual at  $(x, t)$  has an ancestor who lived at  $(\xi, \tau)$ , is just the solution of Eq. (A2) for the initial condition

$$G(\xi, 0|x, 0) = \delta(x - \xi),$$

where  $\delta(x)$  is the Dirac delta function. It is straightforward [24] to show that this Green’s function of Eq. (A2) also obeys the Fokker–Planck equation (3), in which time is measured in the backward direction; Eq. (A2) is usually called the (Kolmogorov–) backward equation [24] associated with the Fokker-Planck equation (3).

The content of the Fokker-Planck equation (3) may be further illustrated by a physical analog. The dynamics of a single ancestral line backward in time conditional on a particular demographic history, as described by Eq. (2), is equivalent to the Brownian motion of a particle (with  $k_B T = 1$ ) in a potential  $U(\xi, \tau)$ , whose negative gradient is given by the drift coefficient  $v + 2D \partial_\xi \ln(c)$  in Eq. (3). The form of this time-dependent potential,

$$U(\xi, \tau) = -v\xi - 2D \ln[c(\xi, \tau)], \quad (\text{A3})$$

suggests an interpretation in terms of a fluctuating free energy in which the first and second term are the energetic and entropic contribution, respectively. If the spatial domain is unbounded, the long-time distribution function of the fluctuating particle will be non-trivial only

if the potential Eq. (A3) has the form of a well, which is able to “trap” the fluctuating particle on long times. Otherwise, for instance if the potential is half-open as in the case of a Fisher wave, the ancestral probability distribution will decay to zero on long times.

So far, our analysis was restricted to a one-dimensional reaction diffusion model, in which drift and diffusion are space-independent. The generalization of Eq. (2) to heterogeneous migration and higher dimensions reads

$$\partial_t c = \frac{\partial}{\partial x_i} \left( D_{ij} \frac{\partial c}{\partial x_j} \right) + \frac{\partial}{\partial x_i} (v_i c) + K. \quad (\text{A4})$$

The conservative dynamics is now described by a matrix  $D_{ij}(\vec{x}, t)$  of diffusivities and a velocity vector  $v_i(\vec{x}, t)$ , which may both depend on space and time. Again, the reaction term  $K(\vec{x}, t)$  accounts for deterministic and stochastic fluctuations in the number of individuals due to birth and death processes.

Repeating the above arguments for the dynamics of neutral labels under the more general population dynamics Eq. (A4) yields a multi-dimensional Fokker–Planck equation for the probability density  $G(\vec{\xi}, \tau | \vec{x}, t)$  that an individual at  $(\vec{x}, t)$  has descended from an ancestor who lived at  $(\vec{\xi}, \tau)$ , which is given by

$$\begin{aligned} \partial_\tau G(\vec{\xi}, \tau | \vec{x}, t) &= -\frac{\partial}{\partial \xi_i} J_i(\vec{\xi}, \tau | \vec{x}, t) \\ J_i(\vec{\xi}, \tau | \vec{x}, t) &\equiv -\frac{\partial}{\partial \xi_j} (D_{ij} G) + \\ &\quad \left[ v_j - \frac{\partial D_{ij}}{\partial \xi_j} + \frac{2}{c} \frac{\partial (D_{ij} c)}{\partial \xi_j} \right] G. \end{aligned} \quad (\text{A5})$$

Here, summation over identical indices is implied and time again increases in the backward direction. The natural requirement that there is no probability flux  $\vec{J}$  out of the region  $S$  of non-vanishing population density leads to reflecting boundary condition,  $J = 0$ , on the boundary  $\partial S$  of  $S$ . Note that our stochastic description of the backward dynamics of a single lineage, Eq. (A5), is fully determined by the demographic history  $c(x, t)$ . A knowledge of the actual form of the reaction term  $K(x, t)$  in the reaction-diffusion equation (A4) is not necessary.

## APPENDIX B: SOURCE-SINK POPULATIONS

For purely conservative populations [42] of neutral individuals, subject only to diffusion and drift, it is well-known that the fixation probability is the same for all individuals,  $u = \text{const.} = 1/N$ , and that the effective population size equals the total population,  $N_e = N$ . However, when reaction terms are important, individuals become privileged or handicapped depending on where they linger. As a telling example, let us consider the case of an “oasis” [31] (or source [43]) with a carrying capacity  $c_o$  in equilibrium with a “desert” (a sink) of smaller carrying capacity  $c_d$ . Sufficiently far away [50] from the contact zone, the population densities will be saturated at

their respective carrying capacities. According to Eq. (6), the ancestral distribution function of a stationary non-moving population will be locally proportional to the square of the density,  $P_{\text{st}}(\xi) \propto c_{\text{st}}^2(\xi)$ . As noted in footnote 5 of the main text, the fixation probability of a mutation occurring in a single individual is given by the ratio of ancestral distribution function and population density,  $u(\xi) \propto P_{\text{st}}(\xi)/c_{\text{st}}(\xi)$ . Thus, the probability that a neutral mutation fixes will be larger if it arises deep in the oasis,  $u_o$ , than if it arises in the desert,  $u_d$ . The ratio of both fixation probabilities is given by the ratio of the respective densities,

$$\frac{u_o}{u_d} = \frac{c_o}{c_d}. \quad (\text{B1})$$

Apart from its simplicity, the relation Eq. (B1) is remarkable because it is independent of the diffusion constants and the details of the particular logistic interaction between individuals.

If the combined system of oasis and desert is closed, the effective population size, Eq. (8), evaluates to

$$N_e = \frac{(c_d^2 L_d + c_o^2 L_o)^2}{c_d^3 L_d + c_o^3 L_o} < c_d L_d + c_o L_o \quad (\text{B2})$$

to leading order in the linear sizes  $L_o$  and  $L_d$  of oasis and desert. As mentioned earlier, our reasoning regarding coalescence times only applies to the strong-migration limit, in which the fixation time  $\sim N_e T$ , where  $T$  is the generation time, is much smaller than the longest relaxation time of the Fokker–Planck equation. For the present case, the latter may be estimated by the time needed for lineage to cross the habitat,  $\sim L^2/D$ .

## APPENDIX C: SURFING ON A FISHER WAVE

Here, we apply our theory of gene surfing to the Fisher equation [23],

$$\partial_t c(x, t) = D \partial_x^2 c(x, t) + s \left[ 1 - \frac{c(x, t)}{c_\infty} \right] c(x, t), \quad (\text{C1})$$

which is the prime example of a *pulled* front. This equation was originally proposed as a mean-field model for the spread of a dominant gene with selective fitness advantage  $s$  through a population with constant density  $c_\infty$  (carrying capacity). Equation (C1) has also been useful as a description of an expanding population, for which  $s$  is the difference between the linear birth and death rates, and the term  $-sc^2/c_\infty$  represents some self-limiting process, roughly proportional to the number of pairs of individuals at position  $x$ .

There are two spatially homogeneous fixed points: an unstable fixed point at  $c(x) = 0$ , in which there is no population at all, and a stable fixed point at  $c(x) = c_\infty$ , where the population saturates to the carrying capacity  $c_\infty$  of the environment.

Non-negative initial configurations evolve smoothly toward the stable fixed point; analysis of the time development of spatial fluctuations in this model reveals that equilibrium can be reached via traveling soliton-like solutions  $c_{\text{st}}(x - vt)$ , referred to as Fisher waves [23], which represent steady state solutions of Eq. (C1). In the wave front, where the population density is much smaller than the carrying capacity, the non-linear logistic term  $\propto c^2$  in Eq. (C1) may be neglected. The steady state solution of the remaining linear equation is exponentially decaying,

$$c_{\text{st}} \sim \exp(-x/\lambda), \quad (\text{C2})$$

for  $x \rightarrow \infty$ , where decay length  $\lambda$  and velocity  $v$  are related by

$$0 = \frac{D}{\lambda^2} - \frac{v}{\lambda} + s. \quad (\text{C3})$$

The population density to be nonnegative requires real solutions  $\lambda > 0$  of Eq. (C3), which do not exist unless

$$v \geq 2\sqrt{Ds}. \quad (\text{C4})$$

It can be shown that the solution corresponding to the lowest velocity  $v = 2\sqrt{Ds}$  and decay length  $\lambda = \sqrt{D/s}$  is approached for any initial conditions with compact support, and, thus, is the solution most relevant for biological applications [33].

Now, when we evaluate the surfing probability, Eq. (6) for this standard model of a spreading wave, we find zero - a somewhat surprising result in light of the finite bell-like curves obtained from our simulations of stochastic Fisher waves (Fig. 2a). The exponential decrease of the population density at the wave front, Eqs. (C2, C4), is simply not fast enough to render the function  $c_{\text{st}}^2(\xi) \exp(v\xi/D)$  normalizable - not even for the lowest velocity for which it approaches a positive constant as  $\xi \rightarrow \infty$ . Hence, a non-zero stationary distribution function of the common ancestor does not exist, even though the total population is in a steady state.

A closer look to the *dynamics* of a lineage, as described by the Fokker-Planck equation (3), reveals how the distribution of the common ancestor decays to zero with time. As we evolve the probability distribution  $G(\xi, \tau|x, t = 0)$  that a lineage diffuses from a location  $x$  to  $\xi$  back through time, it spreads out, due to diffusion, and is subject to a drift of strength  $v + 2D\partial_\xi \log c_{\text{st}}$ . If a lineage starts out deep in the wake of the wave,  $x \ll -1$ , where  $\partial_\xi \log c_{\text{st}} \rightarrow 0$ , it experiences a drift pushing it towards the wave front. After a time of the order of  $|x|/v$ , the probability cloud of the single lineage reaches the front and, when the inflection point is passed, the drift has decreased appreciably because  $-\partial_\xi \log c_{\text{st}} = 0(1)$ . In the tip of the wave, the density profile is exponential, Eq. (C2), such that the drift,  $v + 2D\partial_\xi \log(c_{\text{st}})$ , saturates at a *non-negative* value  $w \equiv (v - 2D/\lambda) \geq 0$ , which tends to move the lineage even further away from the bulk into the tip of the wave. For large times, the distribution  $G$  assumes the form of a bell-like curve of width  $\sim \sqrt{D\tau}$  moving with a velocity  $w$ . For any fixed  $\xi$  in the foot of the wave and  $w > 0$ , the distribution function  $G$  decays exponentially to zero for times, when the probability “cloud” has passed  $\xi$ . Only for  $w = 0$ , which corresponds to the lowest allowed velocity  $v = 2\sqrt{sD}$ , drift is absent in the tip of the wave, such that the decay is much slower,  $G \rightarrow \tau^{-1/2}$ . In any case, however,  $G$  decays to zero for any location  $\xi$ , which means that, no matter which individual we choose, the fixation probability is zero. Thus, “successful surfing” is not possible in the case of the deterministic Fisher wave, as was mentioned already in the main text. We expect that, in a *stochastic* simulation of a finite number of particles, the time-dependent features discussed above for the mean-field Fisher equation are merely *transient* and only visible for times smaller than the longest relaxation time of the Fokker-Planck equation, Eq. (3). This conjecture can be supported by employing an approximation scheme due to Brunet-Derrida [34] to take into account leading order effects of discreteness.

- 
- [1] Templeton, A. (2002) *Nature* **416**, 45-51.
  - [2] Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A., and Feldman, M. W. (2002) *Science* **298** (5602), 2381 - 2385.
  - [3] Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., and Cavalli-Sforza, L. L. (2005) *Proc. Natl. Acad. Sci. USA*, 15942-15947.
  - [4] Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. (1993) *Science* **259**, 639-646.
  - [5] Phillips, B. L., Brown, G. B., Webb, J. K., and Shine, R. (2006) *Nature* **439** (7078), 803.
  - [6] Hewitt, G. M., (2000) *Nature* **405**, 907-913.
  - [7] Currat, M., Excoffier, L., Maddison, W., Otto, S., Ray, N., Whitlock, M. C., and Yeaman, S. (2006) *Science* **313**, 172a.
  - [8] Hewitt, G. M., (1996) *Biol. J. Linn. Soc.* **58**(3), 247-276.
  - [9] Rousset, F., (2001) *Handbook of Statistical Genetics* (John Wiley & Sons, London), 239-269.
  - [10] Hartl, D. L. and Clark, A. G. (1997) *Principles of population genetics* Sinauer Associates Sunderland, Mass..
  - [11] Nichols, R. A. and Hewitt, G. M. (1994) *Heredity* **72**, 312-317.
  - [12] Austerlitz, F., and Garnier-Gere, P. H. (2003) *Heredity* **90**, 282-290.
  - [13] Edmonds, C. A., Lillie, A. S., and Cavalli-Sforza, L. L. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 975-979.
  - [14] Klopstein, S., Currat, M. and Excoffier, L. (2006) *Mol. Biol. Evol.* **23**, 482-490.
  - [15] Liu, H., Prugnolle, F., Manica, A., and Balloux, F. (2006) *Am. J. Hum. Genet.* **79**, 230-237.

- [16] Austerlitz, F. , Jung-Muller, B., Godelle, B., and Gouyon, P. H. (1997) *Theor. Pop. Biol.* **51**, 148-164.
- [17] Le Corre, V. and Kremer, A. (1998) *J. Evol. Biol.* **11**, 495-512.
- [18] Allee, W. C. (1931) *Am. J. Soc.* **37** (3), 386-398.
- [19] van Saarloos, W. (2003) *Phys. Rep.* **386**, 29-222.
- [20] Panja, D. (2004) *Phys. Rep.* **393**, 87-174.
- [21] Courchamp, F., Clutton-Brock, T., and Grenfell, B. (1999) *Trends Ecol. Evol.* **14**, 405-410.
- [22] Doering, C. R., Mueller, C., and Smereka, P. (2003) *Physica A*, **325**(1-2), 243-259.
- [23] Fisher, R. A. (1937) *Ann. Eugenics* **7**, 355-369.
- [24] van Kampen, N. G. (2001) *Stochastic processes in physics and chemistry*. (Elsevier, Amsterdam).
- [25] Vlad, M. O., Cavalli-Sforza, L. L., and Ross, J. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 10249-10253.
- [26] Barton, N. H. and Wilson, I. (1995) *Philos. Trans. R. Soc. London, Ser. B* **349**(1327), 49-59.
- [27] Tsimring, L. S., Levine, H., and Kessler, D. A. (1996) *Phys. Rev. Lett.* **76**, 4440-4443.
- [28] Rouzine, I. M., Wakeley, J., and Coffin, J. M. (2003) *Proc. Natl. Acad. Sci. USA* **100** (2), 587-592.
- [29] Brunet, E., Derrida B., Mueller A. H., and Munier, S. (2006) *Europhys. Lett.* **76**, 1-7.
- [30] Pulliam, H. R. (1988) *Am. Nat.* **132** (5), 652-661.
- [31] Dahmen, K. A., Nelson, D. R. and Shnerb, N. M. (2000) *J. Math. Biol.* **41** (1), 1-23.
- [32] Ebert, U. and van Saarloos, W. (2000) *Physica D* **146**, 1-99.
- [33] Murray, J. D. (2004) *Mathematical Biology* (Springer, New York).
- [34] Brunet, E. and Derrida, B. (1997) *Phys. Rev. E* **56**, 2597-2604.
- [35] Wilkins, J. F. and Wakeley, J. (2002) *Genetics* **161** (2), 873-888.
- [36] Nagylaki, T. (1980) *J. Math. Biol.* **9**, 101-114.
- [37] Notohara, M. (1993) *J. Math. Biol.* **31**(2), 115-122.
- [38] Kingman, J. F. C. (1982) *Stoch. Proc. Appl.* **13**, 235-248.
- [39] Aldous, D. J. (1985) *Lecture notes in mathematics* **1117**, 1-198.
- [40] Nagylaki, T. (1975) *Genetics* **80**, 595-615.
- [41] Pease, C. M., Lande, R. and Bull, J.J. (1989) *Ecology* **70** (6), 1657-1664.
- [42] Nagylaki, T. (2000) *J. Math. Biol.* **41**, 123-142.
- [43] Gaggiotti, O. E. (1996) *Theor. Pop. Biol.* **50**(2), 178-208.
- [44] The simulation box is forced to move with the wave front such that it always contains less than a given large number  $M$  of individuals. We usually set  $M = 45 \times N$  individuals, where  $N$  is the deme size. The size  $n$  of the simulation box had to be chosen so that it contained the entire front up to the foremost individual. One hundred lattice sites usually were sufficient.
- [45] If the average wave velocity is  $v$ , then the heterozygosity at position  $x$  (in the non-moving frame) as the wave front passes through is given by  $H(x/v)$ .
- [46] For the smallest deme sizes  $N = 30$ , we carried out 10 realizations each measuring  $10^5$  fixation processes. We have less statistics for larger deme sizes because of the larger number of degrees of freedom and fixation times. For  $N = 36100$ , we ran 30 simulations and measured 500 fixation processes.
- [47] In a deterministic analysis of the simulations in Ref. [13], the Kolmogorov-backward equation [24] associated with Eq. (3) has recently been obtained by Vlad et al. [25].
- [48] In fact, strict stationarity, Eq. (5), is not necessary for Eq. (6) to hold, rather the drift coefficient in the Fokker-Planck equation (2) has to be time-independent. This condition is satisfied whenever the density profile is separable, i.e.,  $c(x, t) = g(t)h(x)$ , for two functions  $g(t)$  and  $h(x)$ .
- [49] The actual fixation probability [10]  $u(\xi)$  of a mutation occurring in a *single* individual at  $\xi$  is obtained from  $P_{\text{st}}$  after dividing by the population density,  $u(\xi) = P_{\text{st}}(\xi)/c_{\text{st}}(\xi)$ . This quantity measures the probability of ultimate evolutionary success of a neutral genetic marker in a single individual at location  $\xi$  in the wave front.
- [50] The transition zone from oasis to desert has a characteristic width  $\sqrt{D/s}$  that depends on a diffusion constant and the effective growth rate  $s$ .