Spatial Gene Frequency Waves Under Genotype-Dependent Dispersal

Sebastian Novak^{*,1} and Richard Kollár[†]

^{*}Institute of Science and Technology Austria, Klosterneuburg 3400, Austria and [†]Department of Applied Mathematics and Statistics, Comenius University, 842 48 Bratislava, Slovakia

ORCID ID: 0000-0002-2519-824X (S.N.)

ABSTRACT Dispersal is a crucial factor in natural evolution, since it determines the habitat experienced by any population and defines the spatial scale of interactions between individuals. There is compelling evidence for systematic differences in dispersal characteristics within the same population, *i.e.*, genotype-dependent dispersal. The consequences of genotype-dependent dispersal on other evolutionary phenomena, however, are poorly understood. In this article we investigate the effect of genotype-dependent dispersal on spatial gene frequency patterns, using a generalization of the classical diffusion model of selection and dispersal. Dispersal is characterized by the variance of dispersal (diffusion coefficient) and the mean displacement (directional advection term). We demonstrate that genotype-dependent dispersal may change the qualitative behavior of Fisher waves, which change from being “pulled” to being “pushed” wave fronts as the discrepancy in dispersal between genotypes increases. The speed of any wave is partitioned into components due to selection, genotype-dependent variance of dispersal, and genotype-dependent mean displacement. We apply our findings to wave fronts maintained by selection against heterozygotes. Furthermore, we identify a benefit of increased variance of dispersal, quantify its effect on the speed of the wave, and discuss the implications for the evolution of dispersal strategies.

KEYWORDS population genetics; selection-migration dynamics; Fisher wave; wave speed; dispersal evolution

NATURAL populations are distributed in space and their habitat is often much larger than the individual dispersal range. Thus, populations are not globally well mixed, but individuals interact randomly on a local scale. Dispersal makes the link between local and global-scale interactions. Most theoretical models in population genetics assume that dispersal is random with respect to genotype (Slatkin 1985; Lenormand 2002; Edelaar and Bolnick 2012). Such a random gene flow homogenizes populations, eroding genetic dissimilarities. However, as pointed out by Edelaar and Bolnick (2012), gene flow can play a much more complex role. Non-random genotype-dependent dispersal appears naturally in numerous organisms, *e.g.*, aquatic species (Lutscher *et al.* 2007; Bolnick *et al.* 2009), butterflies (Haag *et al.* 2005; Hanski *et al.* 2006), and plants that are polymorphic in flower shape and/or color (Stanton 1987). Thus, the question arises about the impact of genotype-dependent dispersal on

processes of natural selection in spatially extended populations. For example, habitat choice between two possible environments increases levels of local adaptation, thus reducing genetic loads and facilitating specialization relative to random dispersal (Ravigné *et al.* 2004, 2009; Bolnick and Otto 2013). This kind of adaptive dispersal enhances adaptation by allowing deleterious alleles to be removed more efficiently from the population, and it has been found to shorten waiting times for adaptive peak shifts in models with negative epistasis (Hadany *et al.* 2004). Biased movement along fitness gradients mitigates the deleterious effects of gene flow, favors spatial partitioning of populations, and keeps them geographically more constrained, compared to randomly dispersing species (Armstrong and Roughgarden 2005; Armstrong 2009).

Advantageous mutations tend to spread locally and propagate in the form of traveling wave fronts (“Fisher waves”, Fisher 1937). Gene frequency waves may move in space, but they have a stable shape that describes the transition between areas where different alleles are predominant. Stationary spatial changes in gene frequencies are commonly called clines (Haldane 1948). Gene frequency waves also emerge due to spatial heterogeneities in selection, when different alleles are favored in different places (Slatkin 1973; Nagylaki

1975), or as in the case of selection against hybrids, *e.g.*, in hybrid zones that result from secondary contact between populations (Barton 1979). Spatial gene frequency waves are frequently observed in many natural populations and provide insight into evolutionary patterns (*e.g.*, Haldane 1948; Endler 1977; Szymura and Barton 1986; Bridle *et al.* 2001; Whibley *et al.* 2006; Teeter *et al.* 2008). Depending on their mode of shifting, moving wave fronts are classified as “pulled” or “pushed” waves (*cf.* Stokes 1976). The speed of pulled waves is set by the dynamics of their leading edge. That is, the rate at which one genotype spreads into a foreign habitat and grows from low frequency determines its rate of advance. A classic example of pulled waves is Fisher waves. In contrast, a wave is called a pushed wave if its bulk determines its speed. For instance, this is the case for wave profiles emerging under selection for common alleles, where genes increase in frequency if they exceed a certain frequency threshold. Thus, the dynamics in areas where the focal genotype is common, *i.e.*, in the bulk of the wave, dictate the direction and speed of the wave.

Genotype-dependent dispersal and gene frequency waves in the form of stationary clines that are maintained by a spatial gradient in selection have been studied for various types of adaptive dispersal, *e.g.*, genotypes either biasing their direction of movement along fitness gradients, or basing their decision whether to disperse on their locally realized fitness. In such cases, clines become steeper relative to random dispersal since individuals tend to move out of the mixing zone between genotypes (Armsworth and Roughgarden 2008). More rigorous analytical treatments of traveling waves in continuous space under non-random dispersal exist for ecological models of interacting species (Holmes *et al.* 1994), *e.g.*, predator-prey dynamics (Dunbar 1983) and interspecific competition models (Shigesada *et al.* 1979; Lewis *et al.* 2002). While models in population genetics keep population densities fixed and follow the dynamics of gene frequencies, these ecological models consider the abundances of certain species subject to density-dependent regulation, and dispersal is assumed to depend on the species’ densities.

Any model of genotype-dependent dispersal and selection requires a connection between the dispersal trait and individual fitness. We choose the simplest approach by assuming that the two are completely linked, *i.e.*, a single locus under selection also controls the dispersal behavior of the individual. Thus, both the selected and the dispersal traits are determined by the genotype. The assumption that dispersal and viability are closely correlated is reasonable, since alleles affecting dispersal are likely to also have a pleiotropic effect on fitness. An intuitive example where this can be expected to be true is in flowering plants, whose flower shape both determines how successfully pollinators are attracted and what kind of pollinators to attract. The assumption is also justified if separate alleles determining fitness and dispersal are sufficiently closely linked, at least for a limited period of time. Generally, it has been argued that we should expect close correlations between loci affecting viability traits and indirectly selected life history traits in natural populations (Hawthorne and Via 2001).

A rigorous analysis of our differential-equation model is presented in Kollár *et al.* (2016) where we guarantee the existence of wave front solutions for the type of equations considered here, and study the speed of Fisher waves under genotype-dependent dispersal. In the present article, we discuss the implications of these mathematical results in the context of population genetics and extend them to biologically relevant questions. Studying the effect of genotype-dependent dispersal on spatial gene frequency waves, we show that genotype-dependent dispersal can affect both the shape of a wave profile and the speed at which it moves. Hence, the presence of different dispersal patterns in the same population can bias our inference made from the analysis of gene frequency waves. We partition the speed of wave fronts into components due to dispersiveness, directional bias, and selection, which allows for a qualitative understanding of the impact of systematic differences in each of those factors. Furthermore, we present analytic expressions for gene frequency waves under genotype-dependent dispersal in hybrid zones, thus generalizing classical results that have proven to be of empirical relevance (Barton and Turelli 2011).

Model

Throughout this article, we consider a spatially homogeneous environment with a (spatially and temporally) constant population density. The classical continuous model for selection and dispersal in unidimensional space assumes that all individuals follow the same spatially homogeneous dispersal pattern, which is characterized by a mean displacement (directional advection coefficient) M , and a variance of dispersal (diffusion coefficient) V . If the population consists of two genotypes with frequencies $p(x, t)$ and $1 - p(x, t)$ at location x and time t , the diffusion limit of the stochastic selection-dispersal dynamics for a large population is the well-studied Fisher–Kolmogorov–Petrovskii–Piskunov equation

$$\partial_t p = \frac{V}{2} \partial_{xx} p - M \partial_x p + F(x; p) \quad (1)$$

(Nagylaki 1975). Here, the function $F(x; p)$ describes selection: at position x , the frequency p grows at rate F relative to the frequency of the other genotype. If no selective force is active ($F = 0$), gene frequencies spread in space and equilibrate to a spatially constant value at a rate proportional to V . The parameter M adds a directional component that may emerge due to an active individual preference or due to the presence of a slope, wind, or a current of water in the habitat. Since this is a constant shift of the system, M disappears from the equation when transforming $x \mapsto x - Mt$.

Equation 1 can be generalized to different dispersal strategies for the two genotypes, including the possibility of conditional dispersal strategies that depend, *e.g.*, on space and time (Nagylaki and Moody 1980; Novak 2014). We briefly sketch a derivation of the dynamics here; the mathematical details are presented in Kollár *et al.* (2016). Assume that the two genotypes have mean displacements $M_1(x, t)$ and $M_2(x, t)$, and

variance of dispersal $V_1(x, t)$ and $V_2(x, t)$. If the abundances of the two genotypes are denoted by N_1 and N_2 , their reaction-diffusion dynamics can be written as:

$$\begin{aligned} \partial_t N_1(x, t) = & \frac{1}{2} \partial_{xx} [V_1(x, t) N_1(x, t)] - \partial_x [M_1(x, t) N_1(x, t)] \\ & + G_1(x; N_1, N_2), \end{aligned} \quad (2a)$$

$$\begin{aligned} \partial_t N_2(x, t) = & \frac{1}{2} \partial_{xx} [V_2(x, t) N_2(x, t)] - \partial_x [M_2(x, t) N_2(x, t)] \\ & + G_2(x; N_1, N_2), \end{aligned} \quad (2b)$$

where G_1 and G_2 are the growth rates of the two genotypes. If there was variation in the dispersal trait alone, the class of dispersal strategies that are unconditional with respect to space would be evolutionarily stable in homogeneous environments (Cantrell *et al.* 2010). To exclude the effects of selection on dispersal, we thus restrict our considerations to unconditional dispersal strategies defined by constant $M_i(x, t) \equiv M_i$ and $V_i(x, t) \equiv V_i$. Furthermore, since we assume that the environment is spatially homogeneous, the growth rates of the genotypes do not explicitly depend on x , hence $G_1(x; N_1, N_2) = G_1(N_1, N_2)$ and $G_2(x; N_1, N_2) = G_2(N_1, N_2)$. If the total population size $N_T = N_1 + N_2$ is regulated and remains constant in time and space, the system (2) can be reduced to a single equation for the variable $p = N_1/N_T$. Then, the gene frequency dynamics, Equation 1, for genotype-dependent dispersal generalizes to

$$\partial_t p = \frac{V(p)}{2} \partial_{xx} p - M(p) \partial_x p + F(p), \quad (3)$$

with $M(p) = p M_2 + (1-p) M_1$, $V(p) = p V_2 + (1-p) V_1$, and the selection function F being obtained from the individual growth rates G_1 and G_2 (for details, *cf.* Kollár *et al.* 2016). Since the growth rates do not depend on space explicitly (see above), $F(p)$ is also a function of the gene frequencies only, *i.e.*, selection acts homogeneously in space. It is often convenient to define the differences in mean displacement by $\Delta_M = M_2 - M_1$, and write $\Delta_V = V_2 - V_1$ and $V_1 = V > 0$. With this notation, Equation 3 can be written as

$$\partial_t p = \frac{V + p \Delta_V}{2} \partial_{xx} p - \left(\frac{M_1 + M_2}{2} - \frac{\Delta_M}{2} (1-2p) \right) \partial_x p + F(p). \quad (3')$$

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article.

Results

Existence of wave front solutions

As a first step, we discuss the existence of traveling wave solutions $\tilde{p}(x, t) = P(x - ct) = P(z)$ with $0 \leq P \leq 1$ to Equa-

tion 3 for a wide range of local selection functions F . We focus on solutions that connect $P(-\infty) = 1$ with $P(+\infty) = 0$ and move with speed c (positive c means movement to the right, negative c to the left). The analysis of the inverse case, $P(-\infty) = 0$ and $P(+\infty) = 1$, follows by symmetry.

We consider continuous selection functions F that fulfill $F(0) = F(1) = 0$, and $F'(0) \neq 0$ and $F'(1) \neq 0$, see Figure 1. This contains all types of selection regimes without mutation where the stability of the equilibria $p(x, t) \equiv 0$ and $p(x, t) \equiv 1$ can be determined by first-order stability analysis. Depending on the selection function, one of the following is the case: (i) *If F does not change sign on $[0, 1]$* (Figure 1A), *a continuum of wave solutions exists*. In particular, if $F > 0$ on the interval $(0, 1)$, there is a minimal wave speed c^* such that there exists a wave solution $P(x - ct)$ for every $c \geq c^*$. Analogously, if $F < 0$ on $(0, 1)$, there is a maximal wave speed c^* such that there exists a wave solution $P(x - ct)$ for every $c \leq c^*$. (ii) *If F changes sign in $(0, 1)$ exactly once* (Figure 1B), *there is a unique wave speed c^* permitting a wave solution $P(x - c^*t)$* . (iii) *If F changes sign in $(0, 1)$ more than once* (Figure 1C), *wave solutions may or may not exist, depending on F itself, and on $M(p)$ and $V(p)$* .

Furthermore, wave solutions connect $p = 0$ with $p = 1$ monotonically. We present a rigorous proof of these statements in Kollár *et al.* (2016), also discussing the conditions on the existence of traveling waves under case (iii) in more detail. Similar results for nonlinear diffusion equations with a single dispersal strategy have been derived earlier, for example, see classical work by McKean (1975), Fife and McLeod (1977), Aronson and Weinberger (1978). These authors also consider the asymptotic stability of wave front solutions, which is mathematically nontrivial. In this article, however, we only discuss existence of the wave front solutions, as the nonlinearities in the diffusion and advection terms significantly complicate even the spectral stability analysis of the traveling waves and hence their asymptotic stability.

In our existence statement above, case (i) encompasses Fisher waves, where F is a quadratic polynomial, see below. Cubic selection functions that give rise to bistable waves are contained in case (ii) and will be discussed in more detail in *Genotype-dependent dispersal in hybrid zones*. Both types of wave fronts are well understood for the genotype-independent case, yet genotype-dependent dispersal makes our analysis nontrivial.

Fisher waves under genotype-dependent dispersal

As an illustrative example, consider Fisher waves that model the spread of an advantageous mutation through a spatially distributed population with $F(p) = k p(1-p)$ for some $k > 0$. This function satisfies case (i) in *Existence of wave front solutions*, see also Figure 1A. It corresponds to the choice $G_j(N_1, N_2) = r_j N_j$ ($j = 1, 2$) in Equation 2; then, $k = r_1 - r_2$. The resulting equation for genotype-independent dispersal was analyzed by Fisher (1937). *Existence of wave front solutions* implies that, for every wave speed c above a threshold c^* (which depends on the parameters), there is a wave solution with $P(-\infty) = 1$ and $P(+\infty) = 0$ to the genotype-dependent dispersal Equation 3 that moves at speed c . For $V_1 = V_2 = V$,

Figure 1 Illustration of different choices of $F(p)$. (A) F does not change its sign on $[0, 1]$. (B) F changes sign in $(0, 1)$ exactly once. (C) F changes sign in $(0, 1)$ more than once.

it is possible to explicitly determine the lower bound to the wave speed (critical wave speed) c^* . Using the notation $\Delta_M = M_2 - M_1$, we have

$$c^* = \begin{cases} M_1 + \sqrt{2Vk} & \text{for } \Delta_M \leq \sqrt{2Vk} \\ \frac{M_1 + M_2}{2} + \frac{Vk}{\Delta_M} & \text{for } \Delta_M > \sqrt{2Vk} \end{cases} \quad (4)$$

(the mathematical proof is presented in Kollár *et al.* 2016). Note that the speed of the wave, Equation 4, is always positive in the reference frame traveling with speed M_1 . Hence, the advantageous mutation advances into the range of the wild type at a rate of at least c^* , which is at least $M_1 + \sqrt{2Vk}$, the value found by Fisher (1937) under genotype-independent dispersal. Figure 2 shows the critical wave speed c^* as a function of Δ_M for $M_1 = 0$. We see that for large Δ_M , the threshold wave speed c^* approaches a linear dependence on Δ_M with slope $1/2$. It is interesting to observe that $c^* = M_1 + \sqrt{2Vk}$ is valid for Δ_M in an interval around zero, $\Delta_M \in (-\infty, \sqrt{2Vk}]$. Hence, small differences in the mean displacement between the genotypes do not change the threshold c^* . It can be shown that this phenomenon is due to a transition between a pulled and a pushed wave (Kollár *et al.* 2016). For genotype-independent dispersal, we have a pulled wave, *i.e.*, the speed of the critical wave with the smallest speed is determined by its leading edge (where $p \approx 0$). However, when Δ_M exceeds the critical value of $\sqrt{2Vk}$, the drift of the bulk of the wave (where $p \approx 1$) becomes faster than the pulling speed of the leading edge. Hence, the wave becomes pushed, *i.e.*, its speed is determined by the speed of its bulk. This example demonstrates that genotype-dependent dispersal can have a structurally significant impact on the behavior of gene frequency waves.

Measuring the steepness of a wave profile by its maximal slope (Endler 1977), the analysis in Kollár *et al.* (2016), furthermore reveals that the pushed waves become steeper and steeper in the transition region between $p \approx 1$ and $p \approx 0$ as Δ_M increases. Therefore the wave profile becomes narrower. Genotype-dependent dispersal thus has an influence both on the speed and the shape of gene frequency waves.

The speed of a wave

For an arbitrary selection function F satisfying $F(0) = F(1) = 0$, $F'(0) \neq 0$, and $F'(1) \neq 0$, assume that

$P(z) = P(x - ct)$ is a traveling wave solution to Equation 3' that moves with wave speed c . Inserting P into Equation 3' produces

$$-c P' = \frac{V + P\Delta_V}{2} P'' - \left[\frac{M_1 + M_2}{2} - \frac{\Delta_M}{2} (1 - 2P) \right] P' + F(P).$$

Multiply this equation by $P' = dP/dz$ and integrate over $z \in (-\infty, \infty)$. We assume that the admissible solutions P have vanishing derivatives for $z \rightarrow \pm\infty$. Then, we obtain upon integration by parts

$$c \int_{-\infty}^{\infty} (P')^2 dz = \frac{\Delta_V}{4} \int_{-\infty}^{\infty} (P')^3 dz + \frac{M_1 + M_2}{2} \int_{-\infty}^{\infty} (P')^2 dz - \frac{\Delta_M}{2} \int_{-\infty}^{\infty} (1 - 2P) (P')^2 dz - \int_{-\infty}^{\infty} F(P) P' dz.$$

With $P(-\infty) = 1$ and $P(+\infty) = 0$, the rightmost integral is transformed into $-\int_0^1 F(P) dP$. Thus, we write the speed of the wave front as

$$c = c_S + c_V + c_M, \quad (5)$$

where the three components of this expression are given by

$$c_S = \frac{\int_0^1 F(P) dP}{\int_{-\infty}^{\infty} (P')^2 dz}, \quad c_V = \frac{\Delta_V}{4} \frac{\int_{-\infty}^{\infty} (P')^3 dz}{\int_{-\infty}^{\infty} (P')^2 dz},$$

and

$$c_M = \frac{M_1 + M_2}{2} - \frac{\Delta_M}{2} \frac{\int_{-\infty}^{\infty} (P')^2 (1 - 2P) dz}{\int_{-\infty}^{\infty} (P')^2 dz}.$$

This formula is implicit in the solution P , hence the contributions of selection, mean displacement, and variance of dispersal are not purely additive as suggested by the decomposition of c into c_S , c_V , and c_M . Rather, this representation is chosen so that the respective summand disappears if any of these characteristics—selection, variance of dispersal, or mean displacement—is neutral with respect to the genotypes.

The contribution of selection, c_S : If the two genotypes have different fitnesses, selective pressure acts on them. The term c_S

Figure 2 The critical wave speed c^* as a function of Δ_M . The bold line shows c^* as given by Equation 4 for $k = 1$, $V_1 = V_2 = V = 1$, and $M_1 = 0$. The value $\Delta_M = \sqrt{2Vk}$ marks the transition from a pulled to a pushed wave. The wave speed equals the pulled wave speed for $\Delta_M \leq \sqrt{2Vk}$ and then monotonically increases, approaching a linear function with slope $1/2$.

is already known (Fife 1979) and can be made explicit, e.g., in the case of disruptive selection (see below). In the absence of genotype-dependent dispersal, the direction of movement of the wave front is determined by the sign of the average selection $\int_0^1 F(P)dP$.

The contribution of the variance of dispersal, c_V : If the two genotypes have different variances of dispersal, only their difference Δ_V has a direct effect on the speed of the wave. Since P' does not change its sign (see *Existence of wave front solutions*), the wave front shifts toward the genotype with smaller variance of dispersal, i.e., more mobile individuals gradually push back less mobile ones.

The contribution of the mean displacement, c_M : The mean displacements enter c_M in two ways, first via the arithmetic mean $(M_1 + M_2)/2$. This expression already appears in the genotype-independent case ($M_1 = M_2$) and can be scaled away by shifting the coordinate system accordingly. Second, there is a term proportional to the difference $M_2 - M_1 = \Delta_M$ that is difficult to interpret. If the wave solution P is symmetric to the point $[z_0, P(z_0)]$, where z_0 is the position of the half height of the wave defined by $P(z_0) = 1/2$, the term evaluates to zero. This applies, for example, if the wave front has a sigmoid shape, i.e., is of the form $\{1 + \exp[\gamma(x - ct)]\}^{-1}$, as is the case for genotype-independent dispersal and disruptive selection (see below). However, the sigmoid shape of the wave can be perturbed by breaking the symmetry of forces maintaining it, e.g., by genotype-dependent variance of dispersal.

Genotype-dependent dispersal in hybrid zones

Consider two alleles subject to selection against heterozygotes, such that common alleles have an advantage over rare alleles. Call the first allele the A allele with frequency p , and the second allele the a allele. Assume that selection is homogeneous in space and that the fitness values of the three possible genotypes AA , Aa , and aa are $1 + 2s(1 - 2\hat{p})$, $1 - 2\hat{p}s$, and 1 , re-

spectively. With this notation, $s > 0$ measures the strength of selection and $\hat{p} \in (0, 1)$ is a threshold frequency such that selection increases the frequency of the A allele if it is above $\hat{p}$, and reduces it from below that value. If we assume that fitness differences are small ($s \ll 1$), we can reformulate the local selection terms and obtain $F(p) = 2sp(1 - p)(p - \hat{p})$. With $0 < \hat{p} < 1$, this function satisfies case (ii) in *Existence of wave front solutions*, see also Figure 1B. Our result from *Existence of wave front solutions* implies that Equation 3 with this specification of F has a unique wave front solution $\tilde{p}(x, t) = P(x - ct)$ with $P(-\infty) = 1$ and $P(+\infty) = 0$. Clearly, there is another unique wave front solution with the opposite configuration, $P(-\infty) = 0$ and $P(+\infty) = 1$, and with a generally different wave speed.

Genotype-dependent mean displacement: Let $\Delta_V = 0$, such that $V_1 = V_2 = V$. With genotype-dependent mean displacements $M_1 \neq M_2$, Equation 3 has two sigmoidal wave front solutions given by

$$\tilde{p}^{\pm}(x, t) = \{1 + \exp[\zeta^{\pm}s(x - c^{\pm}t)]\}^{-1}, \quad (6)$$

where

$$\zeta^{\pm} = 4 \left(-\Delta_M \pm \sqrt{\Delta_M^2 + 8Vs} \right)^{-1},$$

and the speed of the wave front is

$$c^{\pm} = \frac{M_1 + M_2}{2} + \frac{1 - 2\hat{p}}{\zeta^{\pm}} \quad (7a)$$

$$= \frac{M_1 + M_2}{2} + \frac{2Vs}{\Delta_M} \frac{1 - 2\hat{p}}{1 \pm \sqrt{1 + \frac{8Vs}{\Delta_M^2}}}. \quad (7b)$$

For each set of parameters (i.e., V , s , M_1 , and M_2), $\tilde{p}^+$ is monotonically increasing and connects $\tilde{p}^+(-\infty) = 0$ with $\tilde{p}^+(+\infty) = 1$. Conversely, $\tilde{p}^-$ is monotonically decreasing, $\tilde{p}^-(-\infty) = 1$ and $\tilde{p}^- (+\infty) = 0$. These solutions hence correspond to the two possible configurations of either genotype being absent on one end and fixed on the other end of the habitat. As discussed above in *Existence of wave front solutions*, these two solutions are the only traveling wave solutions.

The two components of the wave speed (Equation 7a) have been identified in Equation 5. The summand $(1 - 2\hat{p})/\zeta^{\pm}$ brings in the effect of selection. It is scaled relative to dispersal by the parameter $\zeta^{\pm}$, which leads to a nonadditive dependence of the speed of the wave front on selection and dispersal. If dispersal is independent of the genotypes ($\Delta_M = 0$), the term $(M_1 + M_2)/2$ becomes a shift of the wave due to the displacement of the population as a whole. Then, the second summand can be disposed of by the rescaling described above, $x \rightarrow x - Mt$, such that we recover known results (Bazykin 1969; Barton 1979). Note that the wave fronts in Equation 6 emerging from cubic selection functions are pushed waves (Stokes 1976). Correspondingly, Equation 7b has a structure similar to Equation 4 for $\Delta_M > \sqrt{2Vk}$.

Genotype-dependent variance of dispersal: Let $\Delta_M = 0$. If the variance of dispersal is different for the two genotypes, the wave front solution becomes asymmetric, which impedes an explicit solution of Equation 3. However, if the values of V_1 and V_2 are sufficiently similar, $|\Delta_V| \ll 1$, we may assume that the shape of the wave front does not differ from the wave solution under genotype-independent dispersal. Then, from Equation 5, we find that the contribution to the speed of the wave due to the genotype-dependent variance of dispersal is

$$c_V = \frac{\Delta_V}{4} \frac{\int_{-\infty}^{\infty} (P')^3 dz}{\int_{-\infty}^{\infty} (P')^2 dz} \approx -\frac{\Delta_V}{10} \sqrt{\frac{s}{2V}}. \quad (8)$$

Thus, the wave front moves toward the genotype with smaller variance of dispersal [$P(-\infty) = 1$ and $P(+\infty) = 0$]. In Figure 3, the accuracy of the approximation leading to Equation 8 is confirmed numerically. If $|\Delta_V|$ is small relative to V , simulated wave speeds are very precisely predicted by Equation 8. For higher values of $|\Delta_V|$, the induced wave speed increases more slowly than predicted. However, the fit is very close up to considerable deviations in the variances of dispersal; prediction and simulation start to diverge by more than 5% only around $|\Delta_V| \approx 0.1 V$ for the parameters studied here.

Discussion

In nature, dispersal characteristics may vary between individuals even of the same population. There are documented cases in which either the capacity to disperse is variable itself, or systematic differences in a dispersal direction are observable (Edelaar and Bolnick 2012). Classical mathematical models of evolution in spatially extended populations, however, predominantly assume homogeneous dispersal characteristics (we refer to classical work by Fisher 1937, Bazykin 1969, Slatkin 1973, Nagylaki 1975, and Barton 1979 for the type of models treated in this article). Genotype-dependent variation in dispersal may be conditional on environmental factors, for example, variation in fitness (Armsworth and Roughgarden 2008; Ravn et al. 2009) or the intensity of competition (Holmes et al. 1994). This article focuses on unconditional genotype-dependent dispersal, where each genotype i has a fixed dispersal strategy given by constant mean displacement M_i and variance of dispersal V_i . We consider a population consisting of two different dispersal types that may also differ in fitness. The assumption of complete linkage is justified if either a single locus pleiotropically determines fitness and dispersal at the same time, or if recombination between the fitness and the dispersal genes is sufficiently unlikely. The other extreme, a dispersal trait that is completely unlinked from any fitness-related genes, is typically studied in the context of dispersal evolution (Ronce 2007). We employ a continuous diffusion model, Equation 3, to investigate the effect of genotype-dependent dispersal on gene frequency waves. We consider two genotypes that may differ in their fitness, their variance of dispersal (dispersal propensity), and their mean displacement (directional dispersal bias).

Figure 3 Wave speed induced by differences in the variance of dispersal. We numerically solved Equation 3 with $M_1 = M_2 = \Delta_M = 0$, $s = 1$, $F(p) = 2sp(1-p)(p-\hat{p})$, and $\hat{p} = 1/2$ using standard integration routines of the software Mathematica. The simulation was initialized at $t = 0$ with the solution of Equation 3 for genotype-independent dispersal, i.e., by $\tilde{p}(x, 0)$ as in Equation 6 for $\Delta_M = 0$, and stopped at $t = 60$. Assuming that the solution closely approaches a stable wave form within the first 50 time units (confirmed numerically), the displacement of the solution in the last 10 time units (between $t = 50$ and $t = 60$) was used to calculate its wave speed c_V induced by Δ_V . The relative deviation of the simulated wave speed c_V^{sim} from the predicted c_V in Equation 8 is evaluated for various values of Δ_V/V ; the data coincide for positive and negative Δ_V/V of the same absolute value. The reference variance of dispersal is $V = 1$ ($\otimes$), $V = 1/2$ ($+$), and $V = 2$ ($\times$). Small values of $|\Delta_V/V|$ lead to a good fit between the simulation and the analytic prediction.

By mathematical analysis detailed elsewhere (Kollr et al. 2016), we prove the existence of wave solutions to Equation 3 for a general class of selection functions, in particular for quadratic and cubic selection functions. These types of functions are commonly used in population biology. Quadratic selection functions are applied to describe the spread of an advantageous mutation in a population (Fisher 1937). With $0 < \hat{p} < 1$ (see *Genotype-dependent dispersal in hybrid zones*), cubic selection functions model positive, frequency-dependent selection that leads to the maintenance of wave fronts in hybrid zones (Barton 1979). In *The speed of a wave*, we derive a general formula for the wave speed, Equation 5, given implicitly in terms of the wave solution. It allows us to decompose the total speed c of the wave front into individual components due to selection (c_S), differences in the variance of dispersal (c_V), and differences in the mean displacement (c_M). While the term c_S is well known, the components c_V and c_M are new and quantify the relative contributions of selection, variance of dispersal, and mean displacement to the speed of a wave.

Genotype-dependent mean displacement alone, $V_1 = V_2 = V$, as introduced in this model, does not break the asymmetry in the solutions of the model in Equation 3. Thus, the equations can still be solved relatively easily whenever the choice of the local selection function F permits. For instance, if $F(p) \equiv 0$, it is straightforward to check that

$$\tilde{p}(x, t) = \left\{ 1 + \exp \left[\frac{\Delta_M}{V} (x - ct) \right] \right\}^{-1},$$

where $c = c_M = (M_1 + M_2)/2$ is a solution to Equation 3. Hence, one should not simply assume that natural selection is always responsible for adaptive divergence between populations (Bolnick and Otto 2013); even without selection, there may be stable wave fronts that are maintained solely by genotype-dependent dispersal. This is not surprising; if both genotypes have a movement bias in the direction of the environment where they are predominantly present, they only mix to a certain degree in the region where they meet. Similar patterns may emerge if individuals have an active habitat preference, which has been found to exhibit genetic variation and adaptive potential (Jaenike and Holt 1991; Bolnick *et al.* 2009).

Genotype-dependent dispersal may drastically change the behavior of gene frequency waves maintained by selection, as we show by the example of Fisher waves. The minimal velocity at which an advantageous mutation advances through space is known to be $\sqrt{2Vk}$ for genotype-independent dispersal. The wave front is a pulled wave, *i.e.*, the dynamics in its leading edge determines its speed. This is also the case for moderate differences in mean displacement between the genotypes, $\Delta_M = M_2 - M_1 \leq \sqrt{2Vk}$. If Δ_M exceeds its critical value of $\sqrt{2Vk}$, however, the bulk of the wave moves faster than the leading edge would afford. Hence, the wave turns into a pushed wave as the bulk of the wave determines its speed. Thus, genotype-dependent dispersal may not only influence the speed of the wave, but also its dynamic properties.

Our results can be used to detect the significance of genotype-dependent dispersal strategies from experimental data. In the simplest case of two genotypes with the same variance of dispersal, $V_1 = V_2$, but genotype-dependent mean displacement, $M_1 \neq M_2$, modeled by Equation 3, it is enough to compare the asymptotic rates of decay of the wave profile $P(z)$ as $z \rightarrow \infty$ and $z \rightarrow -\infty$ that we denote λ_0 and λ_1 , respectively. These quantities can be read off as the slopes of the data in a logarithmic plot. Here, we assume that the traveling wave has the critical speed c^* given by Equation 4. A short calculation shows that the case $\Delta_M \leq \sqrt{2Vk}$ corresponds to the condition $\lambda_0 + \lambda_1 \leq 0$ and that $\Delta_M/V = \lambda_1 - (\lambda_0 + \lambda_1)^2/(2\lambda_1)$. Thus, one can estimate Δ_M/V directly from the data without any *a priori* knowledge of k , V , M_1 , M_2 , and c . On the other hand, the case $\Delta_M \geq \sqrt{2Vk}$ corresponds to the condition $\lambda_0 + \lambda_1 \geq 0$, and then $\Delta_M/V = \lambda_1$. Note that the condition $\lambda_0 + \lambda_1 > 0$ directly implies $\Delta_M \geq \sqrt{2Vk}$ and thus it is sufficient for the significant difference in the mean displacement. In the general case, $V_1 \neq V_2$, some knowledge of parameters of the model (k , Δ_M , ΔV , or c) is needed to use an analogous criterion.

In hybrid zones, when $F(p) = 2sp(1-p)(p-\hat{p})$, selection against heterozygotes is known to maintain gene frequency waves (Bazykin 1969). These wave fronts—at least in the simplest cases—can be calculated analytically and have proven to be of practical importance to predict rates of spatial spread, local introduction numbers necessary to initialize spatial spread, and sufficient environmental conditions that interrupt spatial spread in biocontrol applications (Barton

and Turelli 2011). For the case that only the mean displacement is genotype dependent, we generalize the known wave solution to genotype-dependent dispersal (Equation 6). Genotype-dependent variance of dispersal disrupts the symmetry in the solutions of the selection-dispersal model (Equation 3) and consequently makes analytic solutions intractable. Thus, we employ a perturbation argument assuming that the difference in the variance of dispersal values (Δ_V) is small. This allows us to derive a prediction for the wave speed induced by Δ_V , Equation 8, which should be accurate at least for small values of $|\Delta_V|$. Numerical simulations, however, show that the validity of the approximation is very broad; for the studied parameter ranges, noticeable deviations from the predicted wave speed only start to appear if $|\Delta_V/V|$ exceeds 10%, see Figure 3.

The variance of dispersal can be interpreted as the mobility of a genotype. Thus, the result above—and more generally the formula for c_V in Equation 5 for arbitrary selection function F —shows that more mobile genotypes push back less mobile ones. It is interesting to interpret this in the light of the evolution of dispersal strategies, as we have a case in which increased dispersal spreads through the population even in the absence of the classic factors of dispersal evolution, *i.e.*, any kind of spatio-temporal habitat variability, explicit relatedness structure, or inbreeding effect. Instead, selection maintains heterogeneities in the spatial genotype frequency profile, which has been shown to be sufficient to create selection for elevated dispersal (Novak 2014). In the framework of that article, the rate of increase of a dispersal type with elevated variance of dispersal is $|\Delta_V/2| \times \int_{\mathbb{R}} (\partial_x p)^2 dx$ (equation 7 in Novak 2014). Inserting the unperturbed wave front solution for genotype-independent dispersal and cubic $F(p) = 2sp(1-p)(p-\hat{p})$ yields $c_V \approx |\Delta_V| \sqrt{s/2V}/6$. Hence, the order of magnitude of c_V and its dependence on the parameters s and V agree with our Equation 8, even though the two models are different.

Throughout this article, we assume that population density is spatially and temporally constant. This is a common assumption in population genetics, but in the case of genotype-dependent dispersal it is even more of an abstraction. Areas in which more mobile genotypes are located will experience an excess of emigration and hence will be underpopulated relative to areas where slow dispersers are abundant. Hence, population density is bound to fluctuate in time, heterogeneously over space, as a consequence of genotype-dependent dispersal. However, constant population densities can be justified if the differences in dispersal strategies between genotypes are small (Novak 2014) or if there is an external population regulation mechanism, for example, a limited number of nesting places available to a large progeny each generation. To incorporate a dynamic response of population density to the heterogeneities in dispersal strategies of the population, one has to include an ecological layer into the model. Traveling wave fronts are known to speed up when moving down population density gradients and thus can be trapped in population sinks (Barton 1979; Barton and Turelli 2011). With genotype-dependent dispersal, however, the population density changes in time with the changing genetic composition of the population.

Hence, simple qualitative predictions will be hard to establish for a joint ecological–evolutionary selection–dispersal model.

Acknowledgments

We thank Nick Barton, Katarína Bod'ová, Srđan Sarikas, Tiago Paixão, and three anonymous reviewers for constructive feedback and support. This project has received funding from the European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement 618091 Speed of Adaptation in Population Genetics and Evolutionary Computation and the European Research Council grant number 250152 (S.N.); as well as by the Scientific Grant Agency of the Slovak Republic under the grant 1/0459/13; and by the Slovak Research and Development Agency under the contract number APVV-14-0378 (R.K.).

Literature Cited

- Armstrong, P. R., 2009 Conditional dispersal, clines, and the evolution of dispersiveness. *Theor. Ecol.* 2: 105–117.
- Armstrong, P. R., and J. E. Roughgarden, 2005 The impact of directed vs. random movement on population dynamics and biodiversity patterns. *Am. Nat.* 165: 449–465.
- Armstrong, P. R., and J. E. Roughgarden, 2008 The structure of clines with fitness-dependent dispersal. *Am. Nat.* 172: 648–657.
- Aronson, D. G., and H. F. Weinberger, 1978 Multidimensional nonlinear diffusion arising in population genetics. *Adv. Math.* 30: 33–76.
- Barton, N. H., 1979 The dynamics of hybrid zones. *Heredity* 43: 341–359.
- Barton, N. H., and M. Turelli, 2011 Spatial waves of advance with bistable dynamics: cytoplasmic and genetic analogues of Allee effects. *Am. Nat.* 178: E48–E75.
- Bazykin, A., 1969 Hypothetical mechanism of speciation. *Evolution* 23: 685–687.
- Bolnick, D. I., and S. P. Otto, 2013 The magnitude of local adaptation under genotype-dependent dispersal. *Ecol. Evol.* 3: 4722–4735.
- Bolnick, D. I., L. K. Snowberg, C. Patenia, W. E. Stutz, T. Ingram *et al.*, 2009 Phenotype-dependent native habitat preference facilitates divergence between parapatric lake and stream stickleback. *Evolution* 63: 2004–2016.
- Bridle, J. R., S. J. Baird, and R. K. Butlin, 2001 Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution* 55: 1832–1843.
- Cantrell, R. S., C. Cosner, and Y. Lou, 2010 Evolution of dispersal and the ideal free distribution. *Math. Biosci. Eng.* 7: 17–36.
- Dunbar, S. R., 1983 Travelling wave solutions of diffusive Lotka–Volterra equations. *J. Math. Biol.* 17: 11–32.
- Edelaar, P., and D. I. Bolnick, 2012 Non-random gene flow: an underappreciated force in evolution and ecology. *Trends Ecol. Evol.* 27: 659–665.
- Endler, J. A., 1977 *Genetic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ.
- Fife, P. C., 1979 *Mathematical Aspects of Reacting and Diffusing Systems, Volume 28 of Lecture Notes in Biomathematics*. Springer-Verlag, Berlin.
- Fife, P. C., and J. B. McLeod, 1977 The approach of solutions of nonlinear diffusion equations to travelling front solutions. *Arch. Ration. Mech. Anal.* 65: 335–361.
- Fisher, R. A., 1937 The wave of advance of advantageous genes. *Ann. Hum. Genet.* 7: 355–369.
- Haag, C. R., M. Saastamoinen, J. H. Marden, and I. Hanski, 2005 A candidate locus for variation in dispersal rate in a butterfly metapopulation. *Proc. Biol. Sci.* 272: 2449–2456.
- Hadany, L., I. Eshel, and U. Motro, 2004 No place like home: competition, dispersal and complex adaptation. *J. Evol. Biol.* 17: 1328–1336.
- Haldane, J. B. S., 1948 The theory of a cline. *J. Genet.* 48: 277–284.
- Hanski, I., M. Saastamoinen, and O. Ovaskainen, 2006 Dispersal-related life-history trade-offs in a butterfly metapopulation. *J. Anim. Ecol.* 75: 91–100.
- Hawthorne, D. J., and S. Via, 2001 Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412: 904–907.
- Holmes, E. E., and R. Durrett, 1994 Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology* 75: 17–29.
- Jaenike, J., and R. D. Holt, 1991 Genetic variation for habitat preference: evidence and explanations. *Am. Nat.* 137: 67–90.
- Kollár, R., and S. Novak, 2016 Existence of traveling waves for the generalized FKPP equation, arXiv:1607.00944 math.AP.
- Lenormand, T., 2002 Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17: 183–189.
- Lewis, M. A., B. Li, and H. F. Weinberger, 2002 Spreading speed and linear determinacy for two-species competition models. *J. Math. Biol.* 45: 219–233.
- Lutscher, F., E. McCauley, and M. A. Lewis, 2007 Spatial patterns and coexistence mechanisms in systems with unidirectional flow. *Theor. Popul. Biol.* 71: 267–277.
- McKean, H. P., 1975 Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piskunov. *Commun. Pure Appl. Math.* 28: 323–331.
- Nagylaki, T., 1975 Conditions for the existence of clines. *Genetics* 80: 595–615.
- Nagylaki, T., and M. Moody, 1980 Diffusion model for genotype-dependent migration. *Proc. Natl. Acad. Sci. USA* 77: 4842–4846.
- Novak, S., 2014 Habitat heterogeneities vs. spatial type frequency variances as driving forces of dispersal evolution. *Ecol. Evol.* 4: 4589–4597.
- Ravigné, V., I. Olivieri, and U. Dieckmann, 2004 Implications of habitat choice for protected polymorphisms. *Evol. Ecol. Res.* 6: 125–145.
- Ravigné, V., U. Dieckmann, and I. Olivieri, 2009 Live where you thrive: joint evolution of habitat choice and local adaptation facilitates specialization and promotes diversity. *Am. Nat.* 174: E141–E169.
- Ronce, O., 2007 How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annu. Rev. Ecol. Evol. Syst.* 38: 231–253.
- Shigesada, N., K. Kawasaki, and E. Teramoto, 1979 Spatial segregation of interacting species. *J. Theor. Biol.* 79: 83–99.
- Slatkin, M., 1973 Gene flow and selection in a cline. *Genetics* 75: 733–756.
- Slatkin, M., 1985 Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* 16: 393–430.
- Stanton, M. L., 1987 Reproductive biology of petal color variants in wild populations of *Raphanus sativus*: I. Pollinator response to color morphs. *Am. J. Bot.* 74: 178–187.
- Stokes, A. N., 1976 On two types of moving front in quasilinear diffusion. *Math. Biosci.* 31: 307–315.
- Szymura, J. M., and N. H. Barton, 1986 Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina orientalis* and *B. variegata*, near Cracow in southern Poland. *Evolution* 40: 1141–1159.
- Teeter, K. C., B. A. Payseur, L. W. Harris, M. A. Bakewell, L. M. Thibodeau *et al.*, 2008 Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Res.* 18: 67–76.
- Whibley, A. C., N. B. Langlade, C. Andalo, A. I. Hanna, A. Bangham *et al.*, 2006 Evolutionary paths underlying flower color variation in *Antirrhinum*. *Science* 313: 963–966.

Communicating editor: G. Coop