When sinks become sources: adaptive colonization in asexuals.

F Lavigne[a,b,c], G Martin[c], Y Anciaux[c,d], J Papaïx[a] and L Roques[a]

[a] BioSP, INRA, 84914, Avignon, France

[b] Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

[c] ISEM (UMR 5554), CNRS, 34095, Montpellier, France

[d] BIRC, Aarhus University, C.F. Møllers Allé 8, DK-8000 Aarhus C, Denmark

Abstract

The successful establishment of a population into a new empty habitat outside of its initial niche is a phenomenon akin to evolutionary rescue in the presence of immigration. It underlies a wide range of processes, such as biological invasions by alien organisms, host shifts in pathogens or the emergence of resistance to pesticides or antibiotics from untreated areas.

In this study, we derive an analytically tractable framework to describe the coupled evolutionary and demographic dynamics of asexual populations in a source-sink system. In particular, we analyze the influence of several factors — immigration rate, mutational parameters, and harshness of the stress induced by migration — on the establishment success in the sink, and on the time until establishment. To this aim, we use a classic phenotype-fitness landscape (Fisher's geometrical model in n dimensions) where source and sink habitats determine distinct phenotypic optima. The harshness of stress, in the sink, is determined by the distance between its optimum and that in the source. The dynamics of the full distribution of fitness and of population size in the sink are analytically predicted under a strong mutation strong immigration limit where the population is always polymorphic.

The resulting eco-evolutionary dynamics depend on mutation and immigration rates in a non straightforward way. Beyond some mutation rate threshold, establishment always occurs in the sink, following a typical three-phase trajectory of mean fitness. The waiting time to this establishment is independent of

immigration rate and decreases with the mutation rate. Beyond the mutation rate threshold, lethal mutagenesis impedes establishment and the sink population remains so, albeit with an equilibrium state that depends on the details of the fitness landscape. We use these results to get some insight into possible effects of several management strategies.

1 Introduction

Most natural populations are spread over a heterogeneous set of environments, to which local subpopulations may be more or less adapted. When these local populations exchange migrants we can define "source" and "sink" populations. Source populations, where the local genotypes have positive growth rate, are self-sustained and can send migrants to the rest of the system. They may be connected to sink populations, where local genotypes are so maladapted that they have negative growth rates (Pulliam, 1988). A recent review (Furrer and Pasinelli, 2016) showed that empirical examples of sources and sinks exist through the whole animal kingdom. In the absence of any plastic or evolutionary change, source-sink systems are stable, with the sources being close to their carrying capacity and the sinks being only maintained by incoming maladapted migrants from source environments. In the literature, different types of source-sink systems have been defined, depending on the presence of migration in and out of the sink (see Fig.1 in Sokurenko et al. (2006) or Table 1 in Loreau et al. (2013)): when emigration from the sink is negligible and only immigration into the sink is considered, one alludes to a "black-hole sink" (Gomulkiewicz et al., 1999), because this environment is a demographic dead-end, yet still avoids permanent extinction. These black-hole sinks, and their demographic and evolutionary dynamics, are the canonical model for studying the invasion of a new environment, outside of the initial species "niche" (Holt et al., 2003, 2004). Indeed, the assumption of an asymmetric migration from source to sink can be a realistic simplification as long as the sink remains much less populated than the source. In this article, we will only consider black-hole sinks and simply use the term sink, when in fact referring to black-hole sink populations.

As such, source-sink dynamics bear important implications, for example, in the definition of a species' niche or the management of endangered species (Pulliam, 1988). Several modeling approaches (*e.g.*, Drury et al., 2007; Garnier et al., 2012) detail the demographic aspects of these dynamics, neglecting the interaction with evolutionary forces, yielding adaptation or maladaptation. However, genetic adaptation can of course further change the situation, allowing the ultimate invasion of the sink, however harsh.

This phenomenon, akin to evolutionary rescue in the presence of immigration, has been dubbed "adaptive colonization" by some authors (Gomulkiewicz et al., 2010). It underlies a wide range of biological processes: biological invasions by alien organisms (Colautti et al., 2017), host shifts in pathogens or the emergence of resistance to pesticides or antibiotics, within treated areas or patients (discussed *e.g.* in Jansen et al. (2011) and Sokurenko et al. (2006)). Hence, a diversity of medical, ecological and agricultural issues depend on the likelihood and timescale of successful invasions (or establishment) of sinks from neighboring source populations. "Establishment" in a sink is generally considered successful when the population is self-sustaining in this new environment, even if immigration was to stop (*e.g.*, Blackburn et al., 2011, for a definition of this concept in the framework of biological invasions).

A rich literature has considered the joint effects of demography and evolution on populations facing a heterogeneous environment, in sexuals (*e.g.*, Kirkpatrick and Barton, 1997) as well as asexuals (*e.g.*, Débarre et al., 2013). The source-sink model is a sub-case of this general problem, that has received particular attention (for a review, see Holt et al., 2005). In general, establishment success may depend on factors that include characteristics of the species local populations (demography, mutation rate and effects, genetic system etc.), of the sink environment (*e.g.* carrying capacity, harshness of the environment, GxE effects, possibly its time-variation), or of the introduction event(s) (intensity and time dynamics of the immigration process, age-stage or phenotype-genotype distributions of the immigrant pool etc.). Of course, no model, be it mathematical or simulation-based, can tackle all these various factors together. The process of adaptive colonization has been mostly studied under simplified models (*e.g.* no age or stage structure, constant stress) and mostly in sexuals. A first class of models is "fitness based": fitness in the sink is additively determined by a set of freely recombining loci, and adaptation occurs by directional selection on fitness itself (Gomulkiewicz et al., 2010; Barton and Etheridge, 2017). However, the vast majority of models that consider evolution in source-sink systems are "trait-based": they assume that fitness is a concave function (quadratic, Gaussian) of a single additive quantitative trait, and that the source and sink environments correspond to alternative optima for this fitness function (*e.g.*, Holt et al., 2003, 2004). Note that Gomulkiewicz et al. (2010) explore both types of models. This second approach could be considered more realistic, because such nonlinear phenotype-fitness maps, with environment dependent optima, allow for epistasis and GxE interactions for fitness, both commonly observed empirically (discussed e.g. in Anciaux et al., 2018). However, these complications have also implied the necessity of some simplifying approximations in the analytical treatment of

the models, or to explore a limited set of parameter values in simulations. Most analytical treatments of the source-sink process rely on stationarity assumptions. Often what is described is an equilibrium state reached by the system after a long time. In asexuals, a mutation-selection-migration balance has been studied (Débarre et al., 2013). In sexuals, analytical models have assumed a constant genetic variance for the underlying trait (*e.g.*, Gomulkiewicz et al., 1999; Holt et al., 2004). Simulations often relax this assumption, but they are necessarily bound to study a limited set of parameter values.

Various key factors have been identified that may impede the ultimate adaptive colonization of a sink, in spite of constant immigration (reviewed in Whitlock and Gomulkiewicz, 2005): density-dependence or very high mutation rates, both in sexuals and asexuals (Gomulkiewicz et al., 1999; Débarre et al., 2013) and "gene swamping" in sexuals (*e.g.*, Kirkpatrick and Barton, 1997). The latter factor is a long-known and well-studied limit to adaptation to new environments: admixture with maladapted migrants can impede the invasion of a sink (Tufto, 2000; Barton and Etheridge, 2017) and stabilize species' ranges (Kirkpatrick and Barton, 1997). Comparatively, the focus, especially in analytical models, has less been on asexuals or on the transient dynamics of invasion (but see Barton and Etheridge (2017) for a treatment of establishment times in a sexual, fitness-based model). Nevertheless, we may also ask how long it will take an asexual population (impervious to gene swamping), in the presence of constant immigration, to invade a sink. The issue seems equally relevant and complementary to that of equilibrium outcomes. Indeed, a model can only approximate some "real" situation for a finite duration: conditions will ultimately change so much that the model will not apply any longer. If colonization is certain but would take thousands of generations, the prediction is less directly relevant than if it takes a few generations. We may also ask whether transient observable population dynamics could inform us on the future outcomes (early vs. late establishment vs. no establishment etc.). These issues are further complicated in asexuals, where classic approximations like the infinitesimal model or constant genetic variance assumptions cannot be used. Yet, many applied issues, especially regarding pathogens, depend on the possibility of sink invasion, by predominantly asexual populations, over relatively short timescales (before equilibrium).

In this paper, we model both the transient dynamics and equilibrium of a source-sink system for an asexual population, in a trait-based model. We assume constant immigration from a source population at mutation-selection balance, and ignore density-dependence in the sink and emigration from the sink, the latter being initially empty (invasion process). This may be deemed reasonable here, as we focus on the transient time during which establishment occurs or not, and thus when these factors may be

neglected *a priori*. We use the classic trait-based model but allow for multiple phenotypic traits to determine fitness, more precisely we use an isotropic version of Fisher's geometrical model (FGM) in n dimensions. As pointed above, trait-based models like the FGM reproduce several observed properties of mutation effects on fitness (reviewed in Tenaillon, 2014). In particular, the FGM has proven accurate in predicting empirical patterns of epistasis among multiple mutations (Martin et al., 2007) or the effect of stress on mutant fitness (Martin and Lenormand, 2006), including strong (antibiotic) stresses (MacLean et al., 2010; Trindade et al., 2012), such as may be encountered in sinks. To make analytical progress, we use a deterministic approximation (as in Martin and Roques, 2016) that neglects stochastic aspects of migration, mutation and genetic drift (we test this approximation by exact stochastic simulations). Under a weak selection strong mutation (WSSM) regime, when mutation rates are large compared to mutation effects, we further obtain an analytically tractable coupled partial-ordinary differential equation (PDE-ODE) model describing the evolutionary and demographic dynamics in the sink. This framework allows us to derive analytic formulae for the demographic dynamics and the full distribution of fitness, at all times, and to investigate the effect of demographic and evolutionary parameters on the establishment success, on the establishment time, and on the equilibrium mean fitness in the sink. In particular, we focus on the effects of the immigration rate, of the harshness of stress induced by the migration event, and of the mutational parameters.

2 Methods

Throughout this paper, we follow the dynamics of the fitness distribution of the individuals in the sink environment, under the joint action of mutation, selection and immigration from the source. The latter remains stable at mutation-selection balance, as migration is asymmetric in this black-hole sink. The absolute fitness of phenotypes that are optimal in the sink is denoted r_{max}. The absolute fitness r and the relative fitness m of any phenotype are connected by the relationship: $r = r_{max} + m$. The absolute and relative mean fitnesses in the sink environment at time t then satisfy:

$$\overline{r}(t) = r_{max} + \overline{m}(t).$$

We use a *deterministic approximation* which neglects variations among replicate populations. Under this approximation, we can equate $\overline{r}(t)$ and $\overline{m}(t)$, the mean (absolute and relative respectively) fitnesses within each population with their expected values among replicate populations. The main notations are summarized in Table 1.

Notation	Description
n	Phenotype dimension
$\mathbf{z}$	Phenotype of an individual
$\mathbf{z}^*$	Optimal phenotype (source)
d	Immigration rate
U	Genomic mutation rate
λ	Mutational variance per trait
μ	$\sqrt{U\,\lambda}$
m	Relative fitness in the sink
m_D	Harshness of stress (fitness distance between source and sink optima)
m_{source}	Fitness of the migrants in the source
m_{migr}	Fitness of the migrants in the sink
p_{migr}	Probability density of m_{migr}
$r_{\max}$	Maximum absolute fitness (sink)
r	Absolute fitness $r = r_{\max} + m$
$N(t)$	Population size at time t
$\overline{m}(t)$	Mean relative fitness
$\overline{r}(t)$	Mean absolute fitness, $\overline{r}(t) = r_{\max} + \overline{m}(t)$
t_0	Establishment time
$C_t(z)$	Cumulant generating function of the relative fitness distribution in the sink

Table 1: **Main notations**

2.1 Demographic model and establishment time t_0

In our simple scenario without density-dependence, the dynamics of the mean absolute fitness alone are sufficient for the coupling of evolutionary and demographic dynamics. Indeed, we consider a sink population with current Malthusian fitness $\overline{r}(t)$ at time t, receiving on average d individuals per unit time by immigration. Under the deterministic approximation, the population size dynamics in the sink environment are therefore given by

$$N'(t) = \overline{r}(t)\, N(t) + d. \tag{1}$$

In the absence of adaptation, $\overline{r}$ is constant, leading to an equilibrium population size $N = d/(-\overline{r})$ when $\overline{r} < 0$, as mentioned in the Introduction. When genetic adaptation is taken into account, we need further assumptions to describe the dynamics of $\overline{r}(t)$ in the sink.

We always assume that the new environment is initially empty ($N(0) = 0$) and that the individuals from the source are, in average, unadapted in the sink. This implies that $\overline{r}(0) := \lim_{t \to 0} \overline{r}(t) < 0$. We define the establishment time t_0 as the first time when the absolute mean fitness in the sink becomes positive:

$$t_0 := \inf\{t > 0 \text{ s.t. } \overline{r}(t) > 0\}.$$

This means that, right after t_0, the population in the sink becomes self-sustaining in the absence of immigration and of further adaptation. By definition (assuming that $\overline{r}$ is continuous), t_0 satisfies $\overline{r}(t_0) = 0$. Depending on the behavior of $\overline{r}(t)$, t_0 may therefore be finite (successful establishment) or infinite (establishment failure).

2.2 Fisher's geometric model

We use Fisher's geometric model (FGM) to describe the relationships between phenotypes and fitnesses in each environment. This phenotype-fitness landscape model has the advantage of leading to realistic distributions of mutation effects on fitnesses (Trindade et al., 2012; Hietpas et al., 2013; Tenaillon, 2014) and to an emergent coupling between the stress, the standing variance (diversity of phenotypes in the source) and the mutation effects on fitnesses (Anciaux et al., 2018).

Phenotype-fitness relationships in the two environments. The FGM assumes that each genotype is characterized by a phenotype at n traits, namely a vector $\mathbf{z} \in \mathbb{R}^n$. Each environment (the source and the sink) is characterized by a distinct phenotypic

optimum. The distance between these optima determines the stress induced by a change of the environment. An optimal phenotype in the sink has maximal absolute fitness r_{max} (relative fitness $m = 0$) and sets the origin of phenotype space ($\mathbf{z} = 0$). Fitness decreases away from this optimum. Following the classic version of the FGM, Malthusian fitness is a quadratic function of the breeding value $r(\mathbf{z}) = r_{max} - \|\mathbf{z}\|^2/2$ and $m(\mathbf{z}) = -\|\mathbf{z}\|^2/2$.

In the source, due to a different phenotype optimum $\mathbf{z}^* \in \mathbb{R}^n$, the relative fitness is $m^*(\mathbf{z}) = -\|\mathbf{z}-\mathbf{z}^*\|^2/2$. As the population size is kept constant in the source (see below), only relative fitness matters in this environment. The harshness of stress $m_D > 0$ is the fitness distance between source and sink optima:

$$m_D = -m(\mathbf{z}^*) = \|\mathbf{z}^*\|^2/2. \tag{2}$$

The decay rate, in the sink, of an optimal phenotype from the source, is thus $r_D = r_{max} - m_D$.

Mutations. In the two environments, mutations occur at rate U and create independent and identically distributed (iid) random variations $\mathbf{dz}$ around the phenotype of the parent, for each trait. We assume here a standard Gaussian distribution of the mutation phenotypic effects (Kimura, 1965; Lande, 1980): $\mathbf{dz} \sim \mathcal{N}(0, \lambda I_n)$, where λ is the mutational variance at each trait, and I_n is the identity matrix in n dimensions. These assumptions induce a distribution of the mutation effects on fitness, given the relative fitness $m_p \leq 0$ of the parent. This distribution has stochastic representation (Martin, 2014) $s \sim -m_p - \frac{\lambda}{2}\chi^2_n\left(-\frac{2m_p}{\lambda}\right)$, where $\chi^2_n\left(-\frac{2m_p}{\lambda}\right)$ denotes the noncentral chi-square distribution with n degrees of freedom and noncentrality $-2m_p/\lambda$. Alternatively, this distribution can be characterized by its moment generating function:

$$\mathbb{E}[e^{s\,z}|m_p] = M_*(z)\, e^{\omega(z)\, m_p}, \tag{3}$$

with

$$M_*(z) = \frac{1}{(1+\lambda\, z)^{n/2}} \text{ and } \omega(z) = \frac{-\lambda\, z^2}{1+\lambda\, z}. \tag{4}$$

Migration events. Migration sends randomly sampled individuals from the source into the sink, at rate $d > 0$. Their relative fitness in the sink is $m_{migr}(\mathbf{z}) = -\|\mathbf{z}\|^2/2$, with $\mathbf{z}$ randomly sampled from the source's phenotype distribution.

2.3 Fitness distribution of the migrants

We assume that the distribution of phenotypes in the source is at mutation-selection balance. The resulting equilibrium distribution of phenotypes yields an equilibrium

fitness distribution in the source. Under a weak selection strong mutation (WSSM) regime, a simple expression for this equilibrium fitness distribution is (Martin and Roques, 2016, Appendix E): $m_{source} \sim -\Gamma(n/2, \mu)$, with $\mu := \sqrt{U\,\lambda}$. This WSSM regime can be quantitatively defined by the inequality $U > U_c := n^2\,\lambda/4$ (Martin and Roques, 2016).

To understand the dynamics of the fitness distribution in the sink, we need to compute the distribution of the relative fitness of the migrants m_{migr} when they arrive into the sink. More precisely, as we will see in the sequel, we need to know the moment generating function: $e^{\phi(z)} := \mathbb{E}[e^{m_{migr}\,z}]$, for any $z \geq 0$. Computations in Appendix A show that for any $z \geq 0$:

$$\phi(z) = -\frac{n}{2}\ln(1+\mu z) - m_D\,z + \frac{m_D\,\mu\,z^2}{1+\mu\,z}. \tag{5}$$

The corresponding distribution of m_{migr} (see Appendix A) is:

$$p_{migr}(m) = \begin{cases} \frac{1}{\mu}\,\left(\frac{|m|}{m_D}\right)^{\frac{1}{2}\left(\frac{n}{2}-1\right)}\,e^{\frac{m-m_D}{\mu}}\,I_{\frac{n}{2}-1}\left[\frac{2\sqrt{m_D|m|}}{\mu}\right], \text{ if } m<0 \\ 0, \text{ if } m \geq 0 \end{cases}, \tag{6}$$

where I_ν is the modified Bessel function of the first kind. The accuracy of this formula is illustrated in Fig. 1. We observe that the mean absolute fitness of the migrants, which coincides with $\overline{r}(0) = \lim \overline{r}(t)$ as $t \to 0$, is given by

$$\overline{r}(0) = r_{max} + \phi'(0) = r_{max} - m_D - \mu\,n/2, \tag{7}$$

with ϕ defined by (5).

The assumption that the individuals from the source are initially decaying ($\overline{r}(0) < 0$) can therefore be expressed by the inequality $r_{max} - \mu\,n/2 < m_D$.

2.4 Trajectories of fitness in the sink: a PDE approach

Assume that at time t, the population in the sink is made of the phenotypes $\{\mathbf{z}_i\}_{i=1,\ldots,N(t)}$ (with $N(t) \in \mathbb{N}$), with the corresponding values of relative fitness $\{m_i(t)\}_{i=1,\ldots,N(t)}$. In the absence of demography and immigration, the dynamics of the fitness distribution is traditionally investigated by a moment closure approximation (Burger, 1991; Gerrish and Sniegowski, 2012): the variations of the moment of order k depend on the moment of order $k+1$ through a linear ordinary differential equation, and the resulting system is solved by assuming that the moments vanish for k larger than some value. A way around this issue is the use of cumulant generating functions (CGFs), which handle

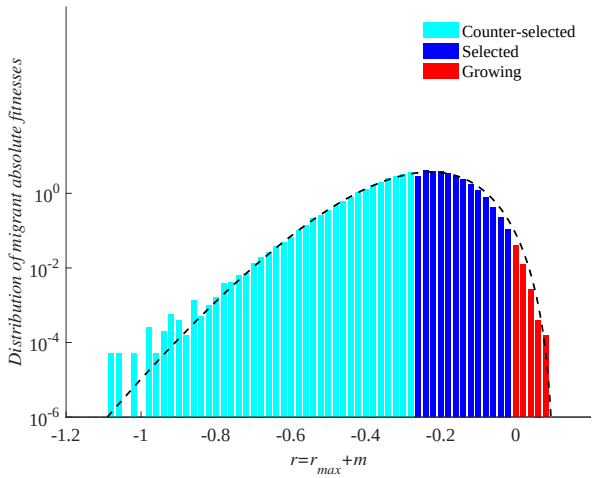


Figure 1: **Distribution of absolute fitness of the migrants in the sink.** The dashed line corresponds to the theoretical expected values of this distribution $p_{migr}(\cdot - r_{\max})$ given by formula (6). The histogram corresponds to the distribution of migrants obtained in exact stochastic simulations after reaching the mutation-selection balance in the source (see Section 2.5). When the sink is empty, individuals are counter-selected if their fitness is below the mean fitness $\overline{r}(0)$ given by (7), selected if their fitness is above $\overline{r}(0)$, and growing if their fitness is positive. The parameter values are $r_{\max} = 0.1$, $U = 0.1$, $m_D = 0.3$, $\lambda = 1/300$, $n = 6$ and $N = 10^6$.

all moments in a single function. In a relatively wide class of evolutionary models of mutation and selection, the CGF of the fitness distribution satisfies a partial differential equation (PDE) that can be solved without requiring a moment closure approximation (Martin and Roques, 2016). We follow this approach here. The empirical CGF of the relative fitness in a population of $N(t)$ individuals with fitnesses $m_1(t), \ldots, m_{N(t)}(t)$ is defined by

$$C_t(z) = \ln\left(\frac{1}{N(t)}\sum_{i=1}^{N(t)} e^{m_i(t)\,z}\right), \tag{8}$$

for all $z \geq 0$. The mean fitness and the variance in fitness in the sink can readily be derived from the CGF: $\overline{m}(t) = \partial_z C_t(0)$ (and $\overline{r}(t) = r_{\max} + \partial_z C_t(0)$), and $V(t) = \partial_{zz} C_t(0)$ (the variance in fitness). In the absence of demography and immigration, and under a weak selection strong mutation (WSSM) regime, Martin and Roques (2016) derived a deterministic nonlocal PDE for the dynamics of C_t. We extended this approach to take into account immigration effects and varying population sizes. This leads to the following PDE (derived in Appendix B):

$$\begin{aligned}\partial_t C_t(z) = \underbrace{\partial_z C_t(z) - \partial_z C_t(0)}_{selection} \; & - \underbrace{\mu^2\left(z^2\,\partial_z C_t(z) + \frac{n}{2}z\right)}_{mutation} \\ & + \underbrace{\frac{d}{N(t)}\left(e^{\phi(z)-C_t(z)} - 1\right)}_{migration,\ demography}, \; z \geq 0,\end{aligned} \tag{9}$$

where we recall that $\mu := \sqrt{U\,\lambda}$. The immigration term depends on $\phi(z)$, which is given by (5), and on $N(t)$, which satisfies the ODE (1), *i.e.* $N'(t) = (\partial_z C_t(0) + r_{max})\,N(t) + d$. This leads to a well-posed coupled system (1)-(9) which can be solved explicitly, as shown in Appendix C.

2.5 Individual-based stochastic simulations

To check the validity of our approach, we used as a benchmark an individual-based, discrete time model of genetic drift, selection, mutation, reproduction and migration with non-overlapping generations.

Source population. A standard Wright-Fisher model with constant population size was used to compute the equilibrium distribution of phenotypes in the source. Our computations were carried out with $N^* = 10^6$ individuals in the source. Each individual $i = 1, \ldots, N^*$ has phenotype $\mathbf{z}_i \in \mathbb{R}^n$ and relative Malthusian fitness $m_i = -\|\mathbf{z}_i - \mathbf{z}^*\|^2/2$, with corresponding Darwinian fitness e^{m_i} (discrete time counterpart of the Malthusian

fitness). At each generation, N^* individuals are sampled with replacement proportionally to their Darwinian fitness. Mutations are simulated by randomly drawing, every generation and for each individual, a Poisson number of mutations, with rate U. Mutation acts additively on phenotype, with individual effects $d\mathbf{z}$ drawn into an isotropic multivariate Gaussian distribution with variance λ per trait (see Section 2.2). Simulations were started with a homogeneous population ($\mathbf{z}_i = \mathbf{z}^*$ for all i at initial time); the equilibrium was considered as being reached after $20/\sqrt{\mu}$ generations (Martin and Roques, 2016). An example of distribution of absolute fitness of the migrants in the sink (distribution of $r_{\max} - \|\mathbf{z}_i\|^2/2$) is presented in Fig. 1.

Sink population. We started with $N(0) = 0$ individuals in the sink. Then, the process to go from generation t to generation $t+1$ is divided into three steps: (i) migration: a Poisson number of migrants, with rate d, was randomly sampled from the equilibrium source population, and added to the population in the sink; (ii) reproduction, selection and drift: each individual produced a Poisson number of offspring with rate $\exp(r_i) = \exp(r_{max} + m_i)$ (absolute Darwinian fitness in the sink); (iii) mutation followed the same process as in the source population. The stopping criterion was reached when $N(t) > 1.5 \cdot 10^6$ individuals or $t > 5 \cdot 10^3$ to limit computation times.

3 Results

3.1 Trajectories of mean fitness

Dynamics of $\overline{r}(t)$ and $N(t)$. The system (1)-(9) leads to an expression for the mean absolute fitness (Appendix C):

$$\overline{r}(t) = \frac{f(t) - 1}{\int_0^t f(\tau)\, d\tau}, \text{ with } f(t) = \exp\left[\left(r_{max} - \mu\,\frac{n}{2}\right) t + \frac{m_D}{2\,\mu}(e^{-2\,\mu\,t} - 1)\right], \tag{10}$$

and for the population size thanks to $N'(t) = \overline{r}(t)\, N(t) + d$.

The good accuracy of eq. (10) is illustrated in Figs. 2-3, by comparing it with the results of individual-based stochastic simulations, under the WSSM assumption ($U > U_c := n^2\,\lambda/4$). Both the individual-based simulations and the analytic expressions show that sink invasion tends to follow three different phases, which are all the more pronounced as the harshness of stress m_D is large. *Phase 1:* After a brief initial increase, the mean fitness remains stable. *Phase 2:* Rapid increase in mean fitness. *Phase 3:* the mean

fitness stabilizes at some asymptotic value. In case of establishment failure (Fig. 3), the adaptation process remains in Phase 1 .

In all cases, formula (7) gives a very accurate estimate of the mean fitness of the migrants, as shown by the agreement between theoretical and simulated values of $\overline{r}(0)$.

Effect of the immigration rate. Unexpectedly, the value of $\overline{r}(t)$ in formula (10) does not depend on the immigration rate d. Thus, only the population size dynamics are influenced by the immigration rate, but not the dynamics of adaptation, in the WSSM and large d regime. To understand this phenomenon, we may divide the equation $N'(t) = \overline{r}(t)\,N(t) + d$ by d, leading to $P'(t) = \overline{r}(t)P(t) + 1$ with $P(t) = N(t)/d$. Then, we observe that the main system (1)-(9) can be written in terms of $P(t)$, independently of N and d. This means that the ratio $N(t)/d$ is not influenced by d. This yields the independence of the evolutionary dynamics of d, because the effect of migration on mean fitness in (9) only depends on $\frac{d}{N(t)}$.

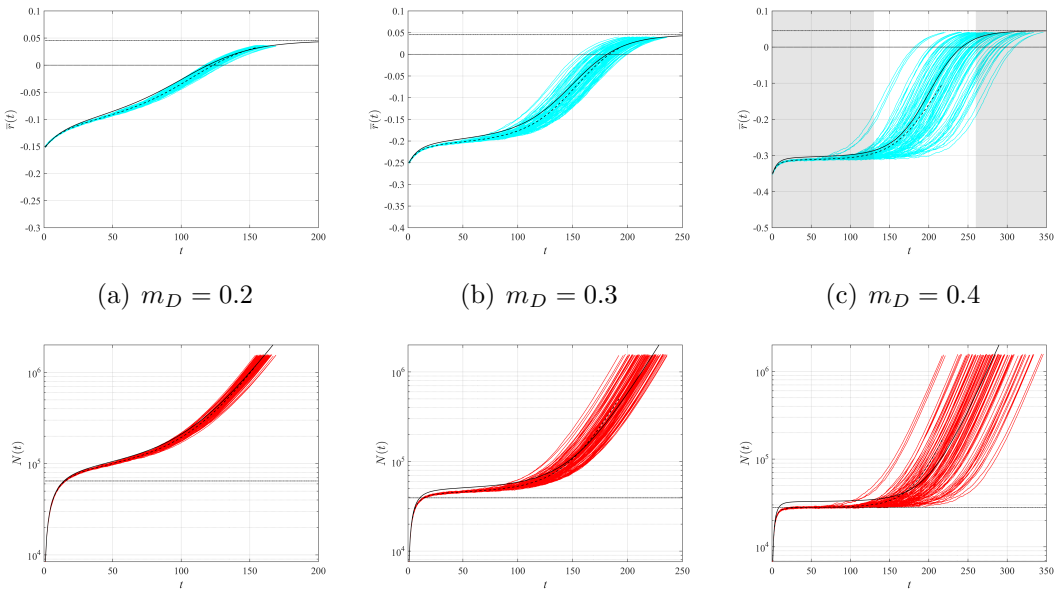
Large time behavior. As seen in Fig. 2, $\overline{r}(t)$ converges towards an asymptotic value $\overline{r}(\infty)$ at large times. The expression (10) shows that this value depends on r_{max}, μ and n. Interestingly, it becomes dependent on the harshness of stress m_D, only in case of establishment failure. More precisely, we get:

$$\begin{aligned}
&\text{if } r_{max} - \mu\, n/2 \geq 0 \text{ then } \overline{r}(\infty) = r_{max} - \mu\, n/2, \text{ and } N(\infty) = \infty \\
&\text{if } r_{max} - \mu\, n/2 < 0 \text{ then } \overline{r}(\infty) = r_{max} - \mu\, n/2 - \delta(m_D), \text{ and } N(\infty) = -d/\overline{r}(\infty),
\end{aligned} \tag{11}$$

for some function $\delta(m_D)$ such that $m_D > \delta(m_D) > m_D/8$ for μ large enough (the inequality $\delta(m_D) > m_D/8$ is true whatever the phenotype dimension n. When n is large enough, sharper lower bounds can be obtained, e.g. $\delta(m_D) > 3\, m_D/8$) for $n \geq 6$), see Appendix D.

These asymptotic results can be interpreted as follows. Below some threshold ($U < U_{lethal} := 4r_{max}^2/(\lambda\, n^2)$, or equivalently $\mu < \mu_{lethal} := 2r_{max}/n$), establishment is always successful and the sink population ultimately explodes (as we ignore density-dependence in the sink). As $d/N(\infty) = 0$, the demographic and evolutionary effects of migrants thus become negligible (being diluted in an effectively infinite population). The sink population thus reaches mutation-selection balance, with a mutation load $\mu\, n/2$, as if it was isolated. It ultimately grows exponentially at rate $r_{max} - \mu\, n/2$ as illustrated in Fig. 2.

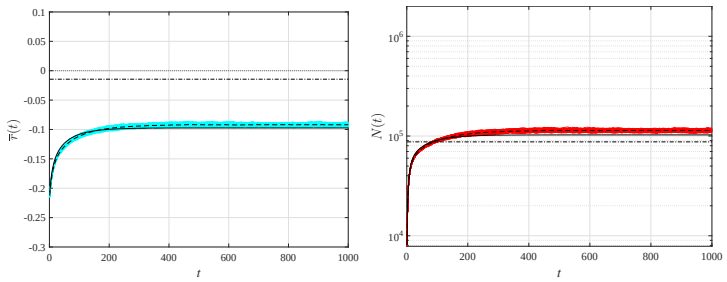
On the contrary, large mutation rates ($U \geq U_{lethal}$ or equivalently $\mu \geq \mu_{lethal}$) lead to establishment failure, which is a form of lethal mutagenesis (see Bull et al. (2007)



(a) $m_D = 0.2$ (b) $m_D = 0.3$ (c) $m_D = 0.4$

(d) $m_D = 0.2$ (e) $m_D = 0.3$ (f) $m_D = 0.4$

Figure 2: **Trajectories of mean fitnesses and population sizes in a WSSM regime, depending on the harshness of stress.** Plain lines: theory given by formulae (1) and (10) vs 100 trajectories obtained by individual-based simulations (blue curves for $\overline{r}(t)$ and red curves for $N(t)$; dashed lines: mean values averaged over the 100 populations). Dash-dot lines: theoretical value of $\overline{r}(\infty) = r_{max} - \mu\, n/2$ (upper panels) and equilibrium population size $-d/\overline{r}(0)$ in the absence of adaptation (lower panels). The three phases of invasion (Phases 1-3, see main text) are illustrated by distinct shaded areas on panel (c) where they are the most obvious. The parameter values are $U = 0.1$ (thus, $U > U_c = 0.03$, which is consistent with the WSSM regime), $r_{max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$. Note that the individual-based computations were stopped when a population size $N(t) = 1.5 \cdot 10^6$ was reached; the mean values could therefore not be computed over the full time span.



(a) $U = 0.44 > U_{lethal}$ (b) $U = 0.44 > U_{lethal}$

Figure 3: **Trajectories of mean fitnesses and population sizes, lethal mutagenesis regime.** Same legend as in Fig. 2. Other parameter values are $m_D = 0.2$, $r_{max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$, leading to a theoretical threshold value for lethal mutagenesis $U_{lethal} = 4r_{max}^2/(\lambda\, n^2) = 0.33$. The panel (a) clearly illustrates the bifurcation in the behavior of the equilibrium mean fitness as $r_{max} - \mu\, n/2$ becomes negative.

for viruses and Bull and Wilke (2008) for bacteria) illustrated in Fig. 3. In this regime, the mutation load $\mu\, n/2$ is larger than the absolute maximal fitness $r_{\max}$ in the sink. Therefore, at mutation-selection balance and even in the absence of any migration, the population could never show positive growth: establishment is impossible because the fitness peak is too low, given the mutation rate and effect. We further identify a "jump" of amplitude $\delta(m_D)$ in the equilibrium mean fitness, as μ increases beyond the lethal mutagenesis threshold (illustrated in Fig. 4). Then, the population ultimately reaches a stable size determined by an immigration - decay equilibrium: a migration load can build up at equilibrium ($\delta(m_D)$) together with the mutation load ($\mu\, n/2$). This migration load is produced by the constant inflow of maladapted genotypes from the source and does depend on the harshness of stress m_D. It is this migration load that creates the "phase transition" in equilibrium fitness as μ crosses beyond μ_{lethal}, the lethal mutagenesis threshold (Fig. 4). Note, however, that contrary to sexual models, migrants do not further breed with and genetically "pollute" locally adapted genotypes. Note also that, in this lethal mutagenesis regime, the sink population does establish to a stable size, that may be higher than that expected in the absence of mutation and adaptation. However, this is not an establishment in that the population would still get extinct if migration was to be stopped.

Behavior outside of the WSSM regime. For $U < U_c$, the WSSM approximations should lose accuracy. We compared the approximation of $\overline{r}(t)$ given by (10) (and

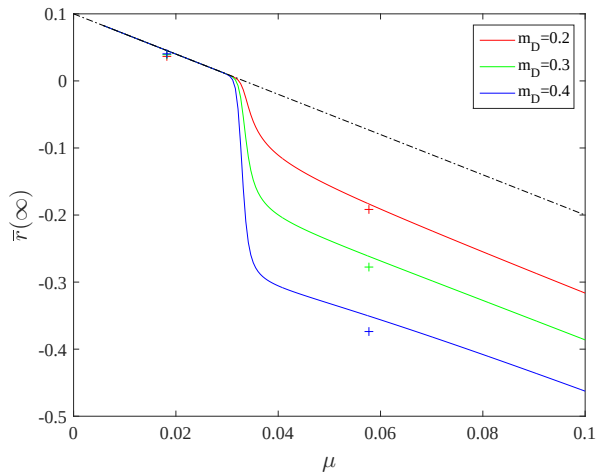
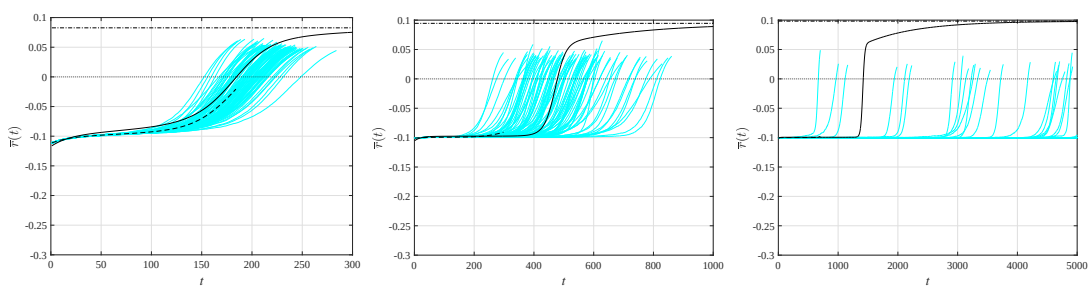
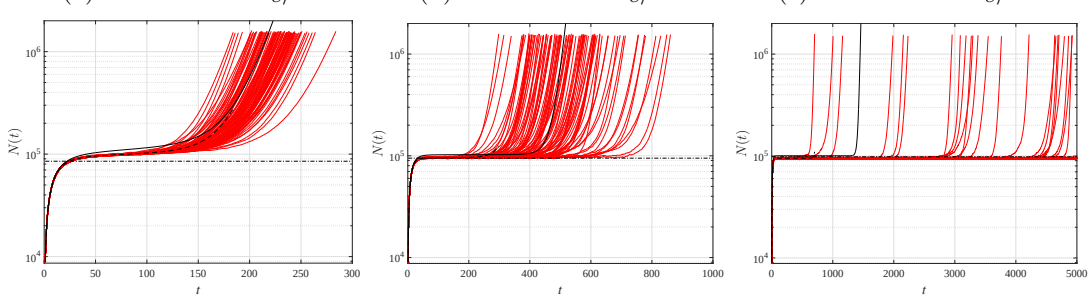


Figure 4: **Mean fitness at large times, dependence with μ and m_D.** The plain lines are the values given by formula (11). The crosses correspond to the result of individual-based simulations. The dash-dot line corresponds to $r_{max} - \mu\, n/2$; the gap between the dash-dot line and the plain lines represents the amplitude of the jump $\delta(m_D)$. Parameter values: $r_{\max} = 0.1$, $n = 6$.



(a) $U = 10^{-2} = U_c/3$ (b) $U = 10^{-3} = U_c/30$ (c) $U = 10^{-4} = U_c/300$



(d) $U = 10^{-2} = U_c/3$ (e) $U = 10^{-3} = U_c/30$ (f) $U = 10^{-4} = U_c/300$

Figure 5: **Trajectories of mean fitnesses and population sizes, low mutation rates.** Same legend as in Fig. 2. Other parameter values are $m_D = 0.2$, $r_{max} = 0.1$, $\lambda = 1/300$, $n = 6$ and $d = 10^4$, leading to $U_c = 0.03$.

the corresponding value of $N(t)$) with individual-based simulations, for mutation rates $U < U_c$ (but keeping a large d). The results are depicted in Fig. 5. We observe a highly variable length of "Phase 1" in the numerical simulations, which increases as U is decreased. Of course, our deterministic approach cannot capture this variability. Still, the global shape of the theoretical trajectories, and the value of the average establishment time, remain close to that given by the numerical simulations when U is not too small, even far from the WSSM regime ($U = U_c/3$ or $U = U_c/30$).

3.2 Establishment time t_0

Of critical importance is the waiting time until the sink becomes a source, when this happens, namely the time t_0 at which $\overline{r}(t)$ becomes positive. This section is devoted to the analysis of this time.

Derivation of an analytical expression. Using the expression (10), we can solve the equation $\overline{r}(t_0) = 0$. We recall that, due to our assumptions, $t_0 > 0$, *i.e.* $\overline{r}(0) = r_{\max} - \mu\, n/2 - m_D < 0$.

The result in (11) shows that $t_0 = \infty$ if $r_{\max} - \mu\, n/2 \leq 0$ (establishment failure). In the case of successful establishment ($m_D > \overline{r}(\infty) = r_{\max} - \mu\, n/2 > 0$), the waiting time to this establishment is

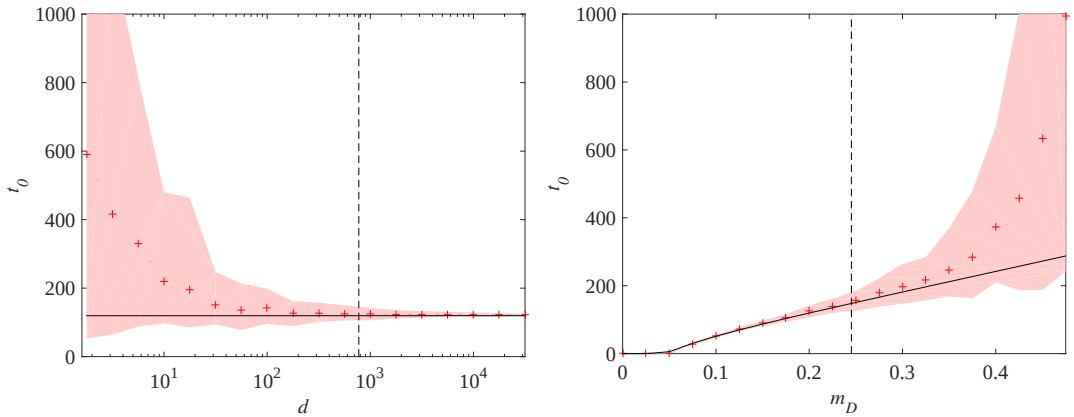
$$t_0 = \frac{1}{2\mu}\left[c + W_0\left(-c\, e^{-c}\right)\right],\ c = \frac{m_D}{r_{\max} - \mu\, n/2}, \tag{12}$$

with W_0 the principal branch of the Lambert-W function (see Appendix E). The waiting time shows a transition (around $c = 1$) from $t_0 \approx c/2\mu$ for small c to $t_0 \approx c/\mu$ for large c, so the establishment time always increases close to linearly with the harshness of stress m_D. Far below the lethal mutagenesis threshold ($\mu << \mu_{lethal} : c \approx m_D/r_{\max}$), the establishment time also increases proportionally to $1/r_{\max}$, $1/\sqrt{U}$ and $1/\sqrt{\lambda}$ as the maximal growth rate $r_{\max}$, the mutation rate U or the variance in mutation effects on phenotypes λ decrease (resp.). Most of these effects are fairly intuitive: it takes more time to establish from a more maladapted source (m_D), with a smaller mutational variance ($U\lambda$), although their particularly simple quantitative effect on t_0 was somewhat unexpected. The effect of $r_{\max}$, although quantitatively simple, is less straightforward to interpret because $r_{\max}$ affects various parameters of the establishment process, all else being equal: it decreases the initial rate of decay ($\bar{r}(0) = r_{\max} - m_D - \mu\, n/2$) and increases the proportion of migrants that are resistant to the sink environment (fitness peak height) or the ultimate exponential growth rate of the population ($\bar{r}(\infty) =$

$r_{\max} - \mu\, n/2$). The latter effect is likely irrelevant to t_0, however, as this growth phase occurs after the establishment time.

We checked the accuracy of this expression for a wide range of parameter values, see Fig. 6 and Appendix F. The accuracy of our approximation increases as the immigration rate d and the mutation rate U increase, because stochasticity (in migration, mutation and reproduction) becomes limited, as assumed in the model. In particular, for d large enough, our result on the independence of t_0 with respect to d is confirmed in simulations (panel a). More precisely, the early population size in the sink is $N(t) \approx -d/\overline{r}(0)$, with $\overline{r}(0) = r_{\max} - \mu n/2 - m_D$, thus whenever $dU >> -\overline{r}(0)$, the mutant input $N(t)U$ in the sink population quickly reaches a large value $N(t)U \approx -dU/\overline{r}(0) >> 1$. Adaptive evolution can then take place within the sink, in a way that is accurately captured by our deterministic approximation (see the dotted lines in Fig. 6). Note that this threshold $d >> |\bar{r}(0)|/U$ is conservative. In fact, even when d is below this level, adaptive evolution yields higher population sizes in the sink, thus limiting stochastic effects later on. On the other hand, when d is smaller and/or m_D is larger, the early population size in the sink is small, so that the deterministic approximation does not apply anymore. In this case, we see that the time t_0 is much more variable, and increases on average with smaller d and larger m_D. This behavior is expected in a migrant-limited regime where colonization relies on invasion by a rare migrant already resistant to the stress in the sink. The waiting time to sample such a migrant will then be more variable, and, on average proportional, to $1/d$. Such relationship is indeed suggested in the lower values of panel (a).

Effect of an intermediate sink. The numerical simulations predict a threshold value for the harshness of stress, beyond which establishment does not occur (or occurs at very large times), see Appendix F. We see in Fig. 6 that as m_D gets close to this threshold, the dependence between t_0 and m_D shifts from linear to superlinear (convex). We argue that this type of dependence has important implications for the potential effect of an intermediate habitat (a milder sink), with phenotype optimum $\mathbf{z}^{**}$ in between $\mathbf{z}^*$ (optimum in the source) and 0 (optimum in the sink), connected by a stepping stone migration process. A natural question is then whether the waiting time to establish in the original sink, via the intermediate one is different from that to establish in the original sink, directly from the source population. In that respect, assume that the overall harshness of stress is the same with and without the intermediate optimum: schematically, $m_D = m_D(\mathbf{z}^* \to 0) = m_D(\mathbf{z}^* \to \mathbf{z}^{**}) + m_D(\mathbf{z}^{**} \to 0)$. When m_D is low, t_0 is roughly linear with m_D so that it may take a similar time to establish



(a) (b)

Figure 6: **Establishment time t_0, dependence with the harshness of stress m_D and the immigration rate d.** Theoretical value of t_0 (black curve) vs value obtained with individual-based simulations (red crosses) and 95% confidence intervals, with fixed $m_D = 0.2$ (panel a) and fixed $d = 10^3$ (panel b). The vertical dotted lines correspond to the values of d ($d = 774$, panel a) and m_D ($m_D = 0.25$, panel b) such that $-dU/\overline{r}(0) = 500$. In all cases, the parameter values are $U = 0.1$, $r_{max} = 0.1$, $\lambda = 1/300$, $n = 6$.

in two step and in one (the sum of intermediate establishment times would be the same as that to establish in a single jump). However, for harsher stress levels where t_0 is superlinear with m_D, the intermediate habitat could provide a springboard to invade the final sink, if both intermediate jumps are much faster than the leap from source to final sink. To illustrate this effect, take $d = 10^3$, $U = 0.1$, $m_D = 0.5$. The individual-based simulations show that $5 \cdot 10^3 \approx t_0(m_D) >> 2\, t_0(m_D/2) \approx 300$, thus, the establishment time in the sink is dramatically reduced in the presence of an intermediate sink habitat. When m_D takes smaller values (in the linear part in Fig. 6, b), *e.g.* $m_D = 0.3$, we get $200 \approx t_0(m_D)$ which is of the same order as $2\, t_0(m_D/2) \approx 160$, thus a smaller effect of the intermediate sink.

4 Discussion

We derived an analytically tractable PDE-ODE framework describing evolutionary and demographic dynamics of asexuals in a source-sink system. Comparison with individual-based stochastic simulations shows that the approach is accurate in the WSSM regime (large mutation rates compared to mutation effects) and with a large migration rate,

and seems robust to small deviations from this regime. This approach reveals the typical shape of the trajectories of mean fitness and population sizes in a sink: (1) in the case of establishment failure, after a brief increase, the mean fitness remains stable at some negative level which depends on the harshness of stress; (2) in the case of successful establishment, this "plateau" is followed by a sudden increase in mean fitness up to the point where it becomes positive and the sink becomes a source. Note that here, we ignored density dependent effects in the sink, so that mean fitness ultimately converges towards an equilibrium that is independent of any migration effect, the latter being diluted into an exploding population.

The same three phases predicted by the model, for the case of successful establishment, are qualitatively observed in Dennehy et al. (2010), an experimental study of invasion of a black-hole sink (an asexual bacteriophage shifting to a new bacterial host). The "host shift" scenario in their Fig. 3 corresponds roughly to our scenario with a population evolved on the native host sending migrants to a new host. The conditions may differ however as the population may not be initially at equilibrium in the native host at the onset of migration. Yet, the dynamics are strikingly similar to those in our fig. $\bar{r}(t)$. An extension of the present work could be to allow for non-equilibrium source populations, which can readily be handled by the PDE (9) (reformulating $\phi(z) = \phi(z,t)$). However, our analytical result on t_0 does rely on an equilibrium source population. Note also that the three phases identified here are observed, in simulations, even in the low d or low U regimes where our analytical derivations break down. Therefore, while the model may provide qualitatively robust insight, quantitative analyses are necessary to really test its predictions. This would ideally include associated measures of m_D and $r_{\max}$, at least, with a known immigration rate d.

Quite unexpectedly, the evolutionary dynamics (especially the waiting time t_0 to establishment) do not depend on the immigration rate. This emerges mathematically from the fact that the evolutionary dynamics only depend on the population size through the ratio $N(t)/d$ between the current population size and the immigration rate, this ratio itself remaining independent of d. This is confirmed by stochastic individual-based simulations: establishment time roughly decreases as $1/d$ when d is small but indeed stabilizes as d becomes larger. More precisely, the result on the independence of t_0 with respect to d should always hold when $NU >> 1$, with $N = d/(-\overline{r}(0))$ the equilibrium population size in the absence of adaptation, which also depends on the harshness of stress m_D and on $r_{\max}$. In this case, the mutant input in the sink population is always large enough to enable our deterministic framework to accurately capture the evolution in the sink. Note that, if density-dependence had been taken into account

with a standard logistic ODE $N'(t) = \overline{r}(t)\, N(t)(1 - N(t)/K) + d$, with K the carrying capacity, the ratio $N(t)/d$ would not be independent of d anymore. This would lead to a dependence of t_0 with respect to the immigration rate d. However, we argue that, provided that K is large enough, the demographic and evolutionary dynamics until the establishment (*i.e.*, for $t \leq t_0$) should be accurately approached by our framework. The mathematical analysis of the resulting PDE-ODE system should be more involved. In a black-hole sink experiment Perron et al. (2007) studied the evolution of resistance to two lethal doses of antibiotics and their combinations in the bacterium *Pseudomonas aeruginosa* (also asexual). Their experiment differs from our scenario in that the sink populations were initially filled with many "naive" individuals ($N_0 >> 1$, amounting to an initial large single immigration event). The authors did notice that immigration rate d affected population densities, but this is not directly a test of our model: even in the large d regime, we do expect that $N(t)$ depends on d, only the time to establishment does not.

Some insight into possible effects of management strategies, *e.g.* quarantine (d), lethal mutagenesis (U), prophylaxis (m_D and r_{max}), can be developed from the results presented here. Migration (propagule pressure) is considered an important determinant of the success of biological invasions in ecology (Von Holle and Simberloff, 2005; Lockwood et al., 2005). However, our results suggest that only a drastic reduction in propagule can be effective (Lockwood et al., 2005) in increasing the establishment time as the immigration rate must be at least below $d = -\overline{r}(0)/U$ to impact significantly the establishment time. Host shift or cross-species transmission cause severe issues in agriculture (emergence of crop pathogens) as well as human health (infections from wild reservoir). For example, it has been shown that the factors increasing potential contacts between human populations and an established animal pathogen or its host tend to increase the risk of emergence of infectious diseases (Morse, 2001). Our study indicates that infection control procedures (*e.g.*, quarantine) that reduce contact between infected animals (source) and recipient host (sink) could be effective to limit the risk of cross-species transmission only if they allow to reach sufficiently low contact levels.

The use of a chemical mutagen to avoid the adaptation of a microbial pathogen and the breakdown of drugs is based on lethal mutagenesis (Bull et al., 2007; Bull and Wilke, 2008). Our approach successfully captures the occurrence of this phenomenon: the establishment fails when the mutation rate U exceeds a certain threshold, which depends on $r_{\max}$, on the mutational variance λ and on the dimensionality of the phenotypic space. Additionally, once this threshold is reached, the equilibrium mean fitness ceases

to depend linearly on the mutational parameter ($\mu = \sqrt{U\,\lambda}$), but rapidly decays. The existence of this negative "jump" in the equilibrium mean fitness, whose magnitude depends on the harshness of stress, leaves no room for evolutionary rescue. Conversely, our approach also reveals that below the lethal mutagenesis threshold, increasing the mutation rate decreases the establishment time as $1/\sqrt{U}$. Hence, the use of a mutagen may be a double-edged sword since it can both hamper and increase the potential for adaptation in the sink.

As expected, the establishment time t_0 increases with the harshness of stress m_D; the population simply needs more time to adapt to more stressful environmental conditions. Increasing m_D or decreasing $r_{\max}$, whenever possible, are probably the safest ways to reduce the risks of biological invasions through adaptive processes or cross-species transmissions of pathogens (in both low and high d regimes). The precise dependence of t_0 with respect to m_D brings us further valuable information. As long as our approach is valid (not too large stresses, leading to finite establishment times), a linear dependence emerges. It suggests that, in a more complex environment with a source and several neighbouring sinks connected by a stepping stone model of migrations, the exact pathway before establishment occurs in a given sink does not really matter. Only the sum of the stresses due to habitat shifts has an effect on the overall time needed to establish in the whole system. Conversely, for larger stress values our analytical approach is not valid, and the numerical simulations indicate a convex (surlinear) dependence of t_0 with respect to m_D. In such case, for a fixed value of the cumulated stress, the establishment time in the sink could be drastically reduced by the presence of intermediate sink habitat. This result, which needs to be confirmed by more realistic modelling approaches, might have applications in the understanding of the role of so-called "preadaptation" in biological invasions. Recent adaptation to one or more facets of the environment within the native range has been proposed as a factor facilitating invasions to similar environments (*e.g.* Hufbauer et al., 2012, anthropogenically induced adaptation to invade). Our results indicate that preadaptation reduces the overall time (*i.e.*, taking the preadapation period into account) before the invasion only when this invasion is associated with a large stress.

The effect of a given environmental challenge, and thus their joint effects when combined (Rex Consortium, 2013), might be modelled in various ways in a fitness landscape framework (see also discussions in Harmand et al., 2017; Anciaux et al., 2018). The first natural option is to consider that multiple stresses tend to pull the optimum further away, and possibly lower the fitness peak $r_{\max}$. In the simplified isotropic model studied here, a larger shift in optimum amounts to increasing m_D. However, a more

realistic anisotropic version, with some directions favored by mutation, might lead to directional effects (where two optima at the same distance are not equally easy to reach) and be particularly relevant to multiple stress scenarios. Another view might consider that different traits are affected by different stresses, leading to anisotropic fitness functions e.g., $m(\mathbf{z}) = -\alpha_1\, z_1^2 - \ldots - \alpha_n\, z_n^2$. In this case, each single stress may increase one or several α_i's, depending on the number of affected traits.

Clearly, many developments are possible and could prove useful to understand how qualitative and quantitative aspects of environmental stresses may affect rescue and invasion. The present isotropic approach provides a simple, tractable null model for the latter, where all environmental effects are summarized by their measurable effects on m_D, $U\,\lambda$ and $r_{\max}$.

Acknowledgments

This work was supported by the French Ministry of Higher Education, Research and Innovation (MESRI allocation doctorale to Y.A.), and the French Agence Nationale de la Recherche (ANR-13-ADAP-0016 "Silentadapt" to G.M. and ANR-13-ADAP-0006 "MeCC" and ANR-14-CE25-0013 "NONLOCAL" to L.R.). This work was fostered by stimulating discussions with Ophélie Ronce and François Hamel.

References

Anciaux, Y., L.-M. Chevin, O. Ronce, and G. Martin (2018). Evolutionary rescue over a fitness landscape. *Genetics*, Early Online.

Barton, N. and A. Etheridge (2017). Establishment in a new habitat by polygenic adaptation. *Theoretical Population Biology*.

Blackburn, T. M., P. Pyšek, S. Bacher, J. T. Carlton, R. P. Duncan, V. Jarošík, J. R. Wilson, and D. M. Richardson (2011). A proposed unified framework for biological invasions. *Trends in Ecology & Evolution 26*(7), 333–339.

Bull, J. J., R. Sanjuan, and C. O. Wilke (2007). Theory of lethal mutagenesis for viruses. *Journal of Virology 81*(6), 2930–2939.

Bull, J. J. and C. O. Wilke (2008). Lethal mutagenesis of bacteria. *Genetics 180*(2), 1061–1070.

Burger, R. (1991). Moments, cumulants, and polygenic dynamics. *Journal of Mathematical Biology 30*(2), 199–213.

Colautti, R. I., J. M. Alexander, K. M. Dlugosch, S. R. Keller, and S. E. Sultan (2017). Invasions and extinctions through the looking glass of evolutionary ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences 372*(1712), 20160031.

Débarre, F., O. Ronce, and S. Gandon (2013). Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology 26*(6), 1185–1202.

Drury, K. L. S., J. M. Drake, D. M. Lodge, and G. Dwyer (2007). Immigration events dispersed in space and time: Factors affecting invasion success. *Ecological Modelling 206*, 63–78.

Furrer, R. D. and G. Pasinelli (2016). Empirical evidence for source–sink populations: a review on occurrence, assessments and implications. *Biological Reviews 91*(3), 782–795.

Garnier, J., L. Roques, and F. Hamel (2012). Success rate of a biological invasion in terms of the spatial distribution of the founding population. *Bulletin of Mathematical Biology 74*, 453–473.

Gerrish, P. J. and P. D. Sniegowski (2012). Real time forecasting of near-future evolution. *Journal of the Royal Society Interface 9*(74), 2268–2278.

Gomulkiewicz, R., R. D. Holt, and M. Barfield (1999). The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. *Theoretical Population Biology 55*(3), 283–296.

Gomulkiewicz, R., R. D. Holt, M. Barfield, and S. L. Nuismer (2010). Genetics, adaptation, and invasion in harsh environments. *Evolutionary Applications 3*(2), 97–108.

Harmand, N., R. Gallet, R. Jabbour-Zahab, G. Martin, and T. Lenormand (2017). Fisher's geometrical model and the mutational patterns of antibiotic resistance across dose gradients. *Evolution 71*(1), 23–37.

Hietpas, R. T., C. Bank, J. D. Jensen, and D. N. Bolon (2013). Shifting fitness landscapes in response to altered environments. *Evolution 67*(12), 3512–3522.

Holt, R. D., M. Barfield, and R. Gomulkiewicz (2004). Temporal variation can facilitate niche evolution in harsh sink environments. *The American Naturalist 164*(2), 187–200.

Holt, R. D., M. Barfield, and R. Gomulkiewicz (2005). *Theories of niche conservatism and evolution: could exotic species be potential tests*, pp. 259–290. Sinauer Associates Sunderland, MA.

Holt, R. D., R. Gomulkiewicz, and M. Barfield (2003). The phenomenology of niche evolution via quantitative traits in a "black-hole" sink. *Proceedings of the Royal Society of London B: Biological Sciences 270*(1511), 215–224.

Hufbauer, R. A., R. Facon, V. Ravigné, J. Turgeon, J. Foucaud, C. E. Lee, O. Rey, and A. Estoup (2012). Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications 5*(1), 89–101.

Jansen, M., A. Coors, R. Stoks, and L. De Meester (2011). Evolutionary ecotoxicology of pesticide resistance: a case study in Daphnia. *Ecotoxicology 20*(3), 543–551.

Kimura, M. (1965). A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences 54*(3), 731–736.

Kirkpatrick, M. and N. H. Barton (1997). Evolution of a species'range. *The American Naturalist 150*, 1–23.

Lande, R. (1980). The genetic covariance between characters maintained by pleiotropic mutations. *Genetics 94*(1), 203–215.

Lockwood, J. L., P. Cassey, and T. Blackburn (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution 20*(5), 223–228.

Loreau, M., A. Daufresne, Tand Gonzalez, D. Gravel, F. Guichard, S. J. Leroux, N. Loeuille, F. Massol, and N. Mouquet (2013). Unifying sources and sinks in ecology and earth sciences. *Biological Reviews 88*(2), 365–379.

MacLean, R. C., A. R. Hall, G. G. Perron, and A. Buckling (2010). The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nature Reviews Genetics 11*(6), 405.

Martin, G. (2014). Fisher's geometrical model emerges as a property of complex integrated phenotypic networks. *Genetics 197*(1), 237–255.

Martin, G., S. F. Elena, and T. Lenormand (2007). Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nature Genetics 39*(4), 555.

Martin, G. and T. Lenormand (2006). The fitness effect of mutations across environments: a survey in light of fitness landscape models. *Evolution 60*(12), 2413–2427.

Martin, G. and T. Lenormand (2015). The fitness effect of mutations across environments: Fisher's geometrical model with multiple optima. *Evolution 69*(6), 1433–1447.

Martin, G. and L. Roques (2016). The non-stationary dynamics of fitness distributions: Asexual model with epistasis and standing variation. *Genetics 204*(4), 1541–1558.

Morse, S. S. (2001). Factors in the emergence of infectious diseases. In *Plagues and politics*, pp. 8–26. Springer.

Perron, G. G., A. Gonzalez, and A. Buckling (2007). Source-sink dynamics shape the evolution of antibiotic resistance and its pleiotropic fitness cost. *Proceedings of the Royal Society of London B: Biological Sciences 274*(1623), 2351–2356.

Pulliam, H. R. (1988). Sources, sinks, and population regulation. *The American Naturalist 132*(5), 652–661.

Rex Consortium (2013). Heterogeneity of selection and the evolution of resistance. *Trends in Ecology & Evolution 28*(2), 110–118.

Sokurenko, E. V., R. Gomulkiewicz, and D. E. Dykhuizen (2006). Source-sink dynamics of virulence evolution. *Nature Reviews Microbiology 4*(7), 548.

Tenaillon, O. (2014). The utility of Fisher's geometric model in evolutionary genetics. *Annual Review of Ecology, Evolution, and Systematics 45*, 179–201.

Trindade, S., A. Sousa, and I. Gordo (2012). Antibiotic resistance and stress in the light of Fisher's model. *Evolution: International Journal of Organic Evolution 66*(12), 3815–3824.

Tufto, J. (2000). Quantitative genetic models for the balance between migration and stabilizing selection. *Genetics Research 76*(3), 285–293.

Von Holle, B. and D. Simberloff (2005). Ecological resistance to biological invasion overwhelmed by propagule pressure. *Ecology 86*(12), 3212–3218.

Whitlock, M. C. and R. Gomulkiewicz (2005). Probability of fixation in a heterogeneous environment. *Genetics 171*(3), 1407–1417.

A Fitness distribution of the migrants: derivation of formulae (5) and (6)

Consider an individual with phenotype $\mathbf{z}$. Its fitness in the source is $m_{source} = -\|\mathbf{z} - \mathbf{z}^\star\|^2/2$, where $\mathbf{z}^\star$ is the optimal phenotype in the source, whereas its fitness in the sink is $m_{migr} = -\|\mathbf{z}\|^2/2$. We observe that

$$\begin{aligned} m_{migr} &= -\frac{\|\mathbf{z} - \mathbf{z}^\star + \mathbf{z}^\star\|^2}{2} \\ &= -\frac{\|\mathbf{z} - \mathbf{z}^\star\|^2 + \|\mathbf{z}^\star\|^2 + 2(\mathbf{z} - \mathbf{z}^\star)\cdot\mathbf{z}^\star}{2} \\ &= m_{source} - \frac{\|\mathbf{z}^\star\|^2}{2} - \|\mathbf{z} - \mathbf{z}^\star\|\ \|\mathbf{z}^\star\| u \\ &= m_{source} - m_D - 2\sqrt{m_D|m_{source}|}u, \end{aligned} \tag{13}$$

with $m_D = \|\mathbf{z}^\star\|^2/2$ and a constant $u \in [-1, 1]$. As the source is assumed to be at the mutation-selection equilibrium, the distribution of fitness in the source satisfies $m_{source} \sim -\Gamma(n/2, \mu)$ (Martin and Roques, 2016) and the corresponding moment generating function is $M_{m_{source}}(z) = (1 + \mu z)^{-n/2}$. The results in (Martin and Lenormand, 2015) show that u is a random variable with moment generating function:

$$M_u(z) := \mathbb{E}[e^{uz}] = {}_0F_1(n/2,\ z^2/4),$$

with ${}_0F_1$ the hypergeometric function, defined by ${}_0F_1(\theta, z) = \sum_{k=0}^{\infty} \frac{1}{\theta(\theta+1)\cdots(\theta+k-1)} \frac{z^k}{k!}$. Let us first compute the moment generating function $M_{migr}(z) := \mathbb{E}[e^{m_{migr}z}]$. We have

$$M_{migr}(z) = \mathbb{E}_{m_{source}}[\mathbb{E}[e^{m_{migr}z}|m_{source}]],$$

and using (13),

$$\begin{aligned} M_{migr}(z) &= \mathbb{E}_{m_{source}}\left[e^{m_{source}z} M_u\left(-2\sqrt{m_D|m_{source}|}z\right)\right] e^{-m_D z} \\ &= \mathbb{E}_{m_{source}}\left[e^{m_{source}z}\ {}_0F_1\left(n/2,\ -m_D m_{source} z^2\right)\right] e^{-m_D z}. \end{aligned}$$

Thanks to the definition of the hypergeometric function ${}_0F_1(n/2, z)$, we get:

$$\begin{aligned} M_{migr}(z) &= \sum_{k=0}^{\infty} \frac{(-m_D)^k}{n/2(n/2+1)\cdots(n/2+k-1)} \frac{z^{2k}}{k!} \mathbb{E}_{m_{source}}[e^{m_{source}z} m_{source}^k] e^{-m_D z} \\ &= \sum_{k=0}^{\infty} \frac{(-m_D)^k}{n/2(n/2+1)\cdots(n/2+k-1)} \frac{z^{2k}}{k!} M_{m_{source}}^{(k)}(z) e^{-m_D z}, \end{aligned}$$

with $M^{(k)}_{m_{source}}(z)$ the kth derivative of $M_{m_{source}}(z)$ with respect to z. Thus,

$$\begin{aligned} M_{migr}(z) &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{m_D \mu z^2}{1+\mu z} \right)^k (1+\mu z)^{-n/2} e^{-m_D z} \\ &= \frac{1}{(1+\mu z)^{n/2}} \cdot \exp\left[-m_D z + \frac{m_D \mu z^2}{1+\mu z} \right]. \end{aligned}$$

Setting $\phi(z) = \ln\left(M_{migr}(z)\right)$, we obtain formula (5).

Let us now show that the distribution of the migrants in the sink satisfies (6). Let p_{migr} be defined by (6). We just have to check that the moment generating function of p_{migr} is M_{migr}:

$$\begin{aligned} \int_{-\infty}^{0} e^{zx} p_{migr}(x) dx &= \int_{-\infty}^{0} e^{zx} \frac{1}{\mu} \left(\frac{|x|}{m_D} \right)^{\frac{n/2-1}{2}} e^{\frac{x-m_D}{\mu}} I_{\frac{n}{2}-1} \left[\frac{2\sqrt{m_D |x|}}{\mu} \right] dx \\ &= e^{-m_D/\mu} \int_{-\infty}^{0} \sum_{p=0}^{\infty} e^{(z+1/\mu)x} \frac{m_D^p}{\mu^{2p+n/2}} \cdot \frac{1}{p!} \cdot \frac{|x|^{p+n/2-1}}{\mathbf{\Gamma}(p+n/2)} dx \\ &= e^{-m_D/\mu} \sum_{p=0}^{\infty} \frac{m_D^p}{\mu^{2p+n/2}} \cdot \frac{1}{p!} \cdot \frac{1}{\mathbf{\Gamma}(p+n/2)} \int_{-\infty}^{0} e^{(z+1/\mu)x} |x|^{p+n/2-1} dx, \end{aligned}$$

where I_ν is the modified Bessel function of the first kind and $\mathbf{\Gamma}$ the gamma function.

Now, for all positive numbers a and b, we have:

$$\int_{-\infty}^{0} e^{ax} |x|^{b-1} dx = \frac{1}{a^b} \int_{0}^{\infty} e^{-x} |x|^{b-1} dx = \frac{\mathbf{\Gamma}(b)}{a^b}.$$

Therefore, we get, for $z > -1/\mu$:

$$\begin{aligned} \int_{-\infty}^{0} e^{zx} p_{migr}(x) dx &= e^{-m_D/\mu} \sum_{p=0}^{\infty} \frac{m_D^p}{\mu^{2p+n/2}} \cdot \frac{1}{p!} \cdot \frac{1}{\mathbf{\Gamma}(p+n/2)} \frac{\mathbf{\Gamma}(p+n/2)}{(z+1/\mu)^{p+n/2}} \\ &= \frac{e^{-m_D/\mu}}{(1+\mu z)^{n/2}} \sum_{p=0}^{\infty} \left(\frac{m_D/\mu}{1+\mu z} \right)^p \cdot \frac{1}{p!} \\ &= \frac{e^{-m_D/\mu}}{(1+\mu z)^{n/2}} \exp\left(\frac{m_D/\mu}{1+\mu z} \right) \\ &= \frac{1}{(1+\mu z)^{n/2}} \exp\left(-\frac{m_D z}{1+\mu z} \right). \end{aligned}$$

This is consistent with formula (5), which proves that the expression (6) is correct.

B PDE satisfied by the CGF of the fitness distribution

In the WSSM regime, and in the absence of immigration, Martin and Roques (2016) have shown that the CGF of the fitness distribution satisfies the following equation:

$$\partial_t C_t(z) = \partial_z C_t(z) - \partial_z C_t(0) - \mu^2 \left(z^2 \, \partial_z C_t(z) + \frac{n}{2} z \right), \; z \geq 0.$$

We derive here the additional term in (9), which describes the effect of immigration on the CGF.

In that respect, we consider a discrete population of size $N(t) \in \mathbb{N}$ at time t, and the corresponding fitnesses $(m_1(t), \ldots, m_{N(t)}(t))$. We define the "empirical" moment generating function

$$M_t(z) := \frac{1}{N(t)} \sum_{i=1}^{N(t)} e^{m_i(t)\, z}.$$

Assuming a Poisson number of immigration events, with rate d per unit time (see Section 2.5), for Δt small enough, the probability that a single immigration events occurs during $(t, t{+}\Delta t)$ is approximately $d\,\Delta t$. The probability that several immigration events occur during this time interval is close to 0. Therefore, the expected change in the moment generating function during Δt, conditionally on the fitness m_{migr} of the unique migrant, is:

$$\begin{aligned}
\Delta M_t(z|m_{migr}) &= d\ \Delta t \left[\frac{1}{N(t)+1} \left(\sum_{i=1}^{N(t)} e^{m_i(t)\ z} + e^{m_{migr}\ z} \right) - \frac{1}{N(t)} \sum_{i=1}^{N(t)} e^{m_i(t)\ z} \right] \\
&= d\ \Delta t \left[\frac{e^{m_{migr}\ z}}{N(t)+1} - \frac{M_t(z)}{N(t)+1} \right].
\end{aligned}$$

Taking expectation over the distribution of m_{migr} (see Appendix A for more details on the distribution of m_{migr}), we get

$$\Delta M_t(z) = \frac{d\ \Delta t}{N(t)+1} \left(e^{\phi(z)} - M_t(z) \right),$$

with $\phi(z) = \ln\left(\mathbb{E}\left[e^{m_{migr}\ z}\right]\right)$. The corresponding change in the CGF $C_t(z) = \ln M_t(z)$ is $\Delta C_t(z) \approx \Delta M_t(z)/M_t(z)$. Thus,

$$\Delta C_t(z) \approx \frac{d\ \Delta t}{N(t)} \left(e^{\phi(z) - C_t(z)} - 1 \right).$$

Finally, dividing the above expression by Δt and passing to the limit $\Delta t \to 0$, we obtain the last term in (9), which describes the effect of immigration on the CGF:

$$\frac{d}{N(t)} \left(e^{\phi(z) - C_t(z)} - 1 \right). \tag{14}$$

C Solution of the system (1)-(9)

We rewrite the system (1)-(9) in the following form:

$$\begin{cases} \partial_t C_t(z) = \alpha(z) \partial_z C_t(z) - \overline{m}(t) + \beta(z) + \frac{d}{N(t)} \left(e^{\phi(z) - C_t(z)} - 1 \right), \\ N'(t) = N(t) \left(r_{\max} + \overline{m}(t) \right) + d, \\ C_t(0) = 0, \\ N(0) = N_0, \end{cases} \tag{15}$$

with $t > 0$ and $z \geq 0$, and where $d \geq 0$, $\alpha(z) := 1 - \mu^2 \, z^2$, $\beta(z) := -\mu \, n \, z/2$. Set $D_t(z) = C_t(y(z))$, with $y(z) = \tanh(\mu z)/\mu$ which satisfies:

$$\begin{cases} y'(z) = \alpha(y(z)), \\ y(0) = 0, \end{cases}$$

so that

$$\partial_t D_t(z) = \partial_t C_t(y(z)) \;\text{ and }\; \partial_z D_t(z) = \alpha(y(z)) \partial_z C_t(y(z)).$$

Thus, $D_t(z)$ satisfies the simpler equation

$$\partial_t D_t(z) = \partial_z D_t(z) - \overline{m}(t) + \beta(y(z)) + \frac{d}{N(t)} \left(e^{\phi(y(z)) - D_t(z)} - 1 \right),$$

with $\overline{m}(t) = \partial_z D_t(0)$ (and $\overline{r}(t) = r_{\max} + \overline{m}(t)$).

Using the method of characteristics, we derive an analytic expression for $D_t(z)$. Fix $z \geq 0$ and denote for all $z \geq t \geq 0$:

$$v(t) = \exp(D_t(z - t)).$$

The function $v \in C^1([0, z])$ satisfies for all $t \in (0, z)$:

$$\begin{aligned} v'(t) &= \left(\partial_t D_t(z - t) - \partial_z D_t(z - t) \right) v(t) \\ &= \left[\beta(y(z - t)) - \overline{m}(t) - \frac{d}{N(t)} \right] v(t) + \frac{d}{N(t)} \, e^{\phi(y(z-t))} \\ &= \left[\beta(y(z - t)) - \frac{N'(t)}{N(t)} + r_{max} \right] v(t) + \frac{d}{N(t)} \, e^{\phi(y(z-t))}, \end{aligned}$$

thanks to $N'(t) = (r_{\max} + \overline{m}(t))N(t) + d$. Thus, for all $t > 0$, $v(t)$ is given by:

$$v(t) = \exp\left[\int_0^t \left(\beta(y(z-\tau)) - \frac{N'(\tau)}{N(\tau)} + r_{max}\right) d\tau\right]$$
$$\left[e^{C_0(y(z))} + \int_0^t \frac{d\ e^{\phi(y(z-\tau))}}{N(\tau)} \exp\left(-\int_0^\tau \left(\beta(y(z-s)) - \frac{N'(s)}{N(s)} + r_{max}\right) ds\right) d\tau\right].$$

Using the fact that the initial population in the sink is $N(0) = 0$, the above expression can be simplified to:

$$v(t) = d\int_0^t e^{\phi(y(z-\tau)) - \int_0^\tau (\beta(y(z-s)) + r_{max})ds} d\tau \cdot \exp\left[-\ln N(t) + \int_0^t (\beta(y(z-\tau)) + r_{max})\, d\tau\right]$$

Hence, by reversing the characteristics, we get:

$$D_t(z) = \int_0^t \beta(y(z+\tau))d\tau - \ln(N(t)) + r_{max}t$$
$$+ \ln\left[d\int_0^t e^{\phi(y(z+\tau)) - r_{max}(t-\tau) - \int_\tau^t \beta(y(z+s))ds} d\tau\right].$$

This leads to an explicit but complex formula for $C_t(z)$ thanks to the relation

$$C_t(z) = D_t\left(\frac{1}{\mu}\arctan(\mu z)\right).$$

Additionally, we have:

$$\partial_z D_t(z) = \beta(y(z+t)) - \beta(y(z)) + \frac{\int_0^t \partial_z\, g(t,z,\tau)\, d\tau}{\int_0^t g(t,z,\tau)\, d\tau},$$

with $g(t,z,\tau) = \exp\left[\phi(y(z+\tau)) + r_{max}(\tau - t) - \int_\tau^t \beta(y(z+s))ds\right]$. Using the fact that $\overline{m}(t) = \partial_z D_t(0)$, $y(0) = 0$ and $\beta(0) = 0$, we get:

$$\begin{aligned}
\overline{m}(t) =& \beta(y(t)) + \frac{\int_0^t g(t,0,\tau)\ [y'(\tau)\phi'(y(\tau)) + \beta(y(\tau)) - \beta(y(t))]\ d\tau}{\int_0^t g(t,0,\tau)\, d\tau},\\
=& \frac{\int_0^t g(t,0,\tau)\ [y'(\tau)\phi'(y(\tau)) + \beta(y(\tau))]\ d\tau}{\int_0^t g(t,0,\tau)\, d\tau},\\
=& \frac{\int_0^t g(t,0,\tau)\ [y'(\tau)\phi'(y(\tau)) + \beta(y(\tau)) + r_{\max}]\ d\tau}{\int_0^t g(t,0,\tau)\, d\tau} - r_{\max},\\
=& \frac{\int_0^t g(t,0,\tau)\, \partial_\tau g(t,0,\tau)\, d\tau}{\int_0^t g(t,0,\tau)\, d\tau} - r_{\max},\\
=& \frac{g(t,0,t) - g(t,0,0)}{\int_0^t g(t,0,\tau)\, d\tau} - r_{\max}.
\end{aligned}$$

Using the expression $g(t,0,\tau) = \exp\left[\phi(y(\tau)) + r_{max}(\tau - t) - \int_\tau^t \beta(y(s))ds\right]$, the formula (5) for ϕ and $y(z) = \tanh(\mu z)/\mu$, we finally get:

$$\overline{r}(t) = \overline{m}(t) + r_{\max} = \frac{\exp\left[(r_{\max} - \mu\,\frac{n}{2})t + \frac{r_D}{2\mu}\left(e^{-2\mu t} - 1\right)\right] - 1}{\int_0^t \exp\left[(r_{\max} - \frac{n}{2}\mu)\tau + \frac{r_D}{2\mu}\left(e^{-2\mu\tau} - 1\right)\right] d\tau},$$

which implies formula (10) in the main text. As we have an explicit formula for $\overline{r}(t)$, we can also solve the ODE $N'(t) = \overline{r}(t)\,N(t) + d$.

D Large time behavior of $\overline{r}(t)$

We recall that, according to formula (10),

$$\overline{r}(t) = \frac{f(t) - 1}{\int_0^t f(\tau)\,d\tau},$$

with $f(t) = \exp\left[\left(r_{max} - \mu\,\frac{n}{2}\right)t + \frac{m_D}{2\,\mu}(e^{-2\,\mu\,t} - 1)\right]$.

We first show that $\overline{r}(t)$ is an increasing function of t. First, we can check that

$$f'(t) = f(t)\,\left(r_{\max} - \frac{\mu n}{2} - m_D e^{-2\mu t}\right).$$

Second, we have

$$\begin{aligned}\overline{r}'(t) &= \frac{f'(t)}{\int_0^t f(\tau)d\tau} - \frac{f(t) - 1}{\left(\int_0^t f(\tau)d\tau\right)^2} f(t) \\ &= \frac{f(t)}{\left(\int_0^t f(\tau)d\tau\right)^2}\left[\left(r_{\max} - \frac{\mu n}{2} - m_D e^{-2\mu t}\right)\int_0^t f(\tau)d\tau - (f(t) - 1)\right].\end{aligned}$$

Let $h(t) = \left(r_{\max} - \frac{\mu n}{2} - m_D e^{-2\mu t}\right)\int_0^t f(\tau)d\tau - (f(t) - 1)$. Thus we see that

$$h'(t) = 2\,\mu\,m_D\,e^{-2\mu t}\int_0^t f(\tau)d\tau \geq 0.$$

Therefore for all $t > 0$, $h(t) > h(0) = 0$, which shows that $\overline{r}$ is increasing.

Since $\overline{r}(0) = r_{max} - \mu\,n/2 - m_D$, this implies that $\overline{r}(t) > r_{max} - \mu\,n/2 - m_D$ for all $t > 0$. In particular, $\overline{r}(\infty) \geq r_{max} - \mu\,n/2 - m_D$ which implies that $\delta(m_D) < m_D$ in (11).

Next, we compute the limit of $\overline{r}(t)$ as $t \to \infty$.

Case (i): we assume that $r_{\max} - \mu n/2 > 0$. Then, $f(t) \sim e^{-\frac{m_D}{2\mu}} e^{(r_{\max} - \mu n/2)t}$ and

$$\int_0^t f(\tau)\, d\tau \sim e^{-\frac{m_D}{2\mu}} \frac{e^{(r_{\max} - \mu n/2)t}}{r_{\max} - \mu n/2}, \text{ as } t \to \infty.$$

Thus,

$$\overline{r}(t) \to r_{\max} - \mu n/2 \text{ as } t \to \infty.$$

Case (ii): we assume that $r_{\max} - \mu n/2 = 0$. Then $f(t) = \exp\left[\frac{m_D}{2\mu}\left(e^{-2\mu t} - 1\right)\right]$ and $\int_0^t f(\tau)\, d\tau \sim t\, e^{-m_D/(2\mu)}$ as $t \to \infty$. Thus,

$$\overline{r}(t) \sim \frac{e^{-m_D/(2\mu)} - 1}{e^{-m_D/(2\mu)} t} \to 0 \text{ as } t \to \infty.$$

Case (iii): we assume that $r_{\max} - \mu n/2 < 0$. Consider an arbitrary constant $\alpha \in (0, 2)$. We can check that, for all $t < T_\alpha := \frac{1}{2\mu} \ln \frac{2}{\alpha}$, we have:

$$e^{-2\mu t} < 1 - \alpha \mu t.$$

In the sequel, we denote $X := r_{\max} - \mu n/2$. We get:

$$\begin{aligned}
\int_0^\infty f(t) dt &= \int_0^{T_\alpha} f(t) dt + \int_{T_\alpha}^\infty f(t) dt \\
&\leq \int_0^{T_\alpha} \exp\left((X - m_D \alpha/2)t\right) dt \\
&+ \int_{T_\alpha}^\infty \exp\left[Xt + \frac{m_D}{2\mu}\left(e^{-2\mu T_\alpha} - 1\right)\right] dt.
\end{aligned}$$

Using the assumption $X = r_{\max} - \mu n/2 < 0$, we obtain:

$$\int_0^\infty f(t) dt \leq \frac{e^{(X - m_D \alpha/2)T_\alpha} - 1}{X - m_D\, \alpha/2} - \exp\left[\frac{m_D}{2\mu}\left(e^{-2\mu T_\alpha} - 1\right)\right] \frac{e^{X T_\alpha}}{X},$$

and using the definition of $T_\alpha = \frac{1}{2\mu} \ln \frac{2}{\alpha}$, we obtain

$$\int_0^\infty f(t) dt \leq -\left(\frac{\alpha}{2}\right)^{\frac{-X}{2\mu}} \left[\frac{\gamma}{X - \alpha\, m_D/2} + \frac{\rho}{X}\right],$$

with $\gamma := \left(\frac{\alpha}{2}\right)^{\frac{X}{2\mu}} - \left(\frac{\alpha}{2}\right)^{\alpha m_D/(4\mu)}$ and $\rho = \exp\left[\frac{m_D}{2\mu}\left(\frac{\alpha}{2}-1\right)\right]$. This leads to the following inequality:

$$\overline{r}(\infty) = -\frac{1}{\int_0^\infty f(t)dt} \leq \left(\frac{\alpha}{2}\right)^{\frac{X}{2\mu}} \frac{X - \alpha m_D/2}{\gamma + \rho\left(1 - \frac{\alpha m_D}{2X}\right)},$$

which can be rewritten:

$$\overline{r}(\infty) \leq \frac{X - \alpha m_D/2}{1+\varepsilon},$$

with

$$\varepsilon := -\left(\frac{\alpha}{2}\right)^{\frac{\alpha\, m_D}{4\,\mu} - \frac{X}{2\,\mu}} + \rho\left(\frac{\alpha}{2}\right)^{-\frac{X}{2\,\mu}}\left(1 - \frac{\alpha\, m_D}{2\,X}\right).$$

A series expansion as $\mu \to \infty$ shows that

$$\varepsilon = \frac{\kappa}{\mu} + O\left(\frac{1}{\mu^2}\right),$$

and κ has the same sign as $\ln(\alpha/2) - 1 + 2\left(\frac{1}{\alpha} - \frac{2}{n}\right)$. Finally, $\varepsilon < 0$ for μ large enough, if and only if

$$n > \frac{4}{\ln(\alpha/2) - 1 + 2/\alpha}.$$

Taking $\alpha = 1/2$, we obtain that $\varepsilon < 0$ for all $n \geq 3$, and consequently, $\overline{r}(\infty) \leq X - \frac{m_D}{4}$ for μ large enough. Thus $\delta(m_D) \geq \frac{m_D}{4}$ in (11), for μ large enough. Similarly, taking $\alpha = 1/4$, we obtain that $\overline{r}(\infty) \leq X - \frac{m_D}{8}$ and $\delta(m_D) \geq \frac{m_D}{8}$ for μ large enough and for all $n \geq 1$. In our numerical computations ($n = 6$), we can use $\alpha = 3/4$, which leads to $\overline{r}(\infty) \leq X - \frac{3\,m_D}{8}$ and $\delta(m_D) \geq \frac{3\,m_D}{8}$ for large μ.

E Establishment time t_0: formula (12)

We recall that t_0 is defined as the first zero of $\overline{r}(t)$. We note that, since $\overline{r}(t)$ is increasing, it admits at most one zero.

First, if $r_{\max} - \mu n/2 \leq 0$, as $\overline{r}(t)$ is increasing and $\overline{r}(\infty) < r_{\max} - \mu\, n/2$ (see (11) and Appendix D), we have $\overline{r}(t) < 0$ for all $t \geq 0$. This implies that $t_0 = \infty$.

Second we assume that $r_{\max} - \mu n/2 > 0$. In this case, $\overline{r}(\infty) = r_{\max} - \mu n/2 > 0$ and the time t_0 is finite (and positive). Therefore, we can solve the equation $\overline{r}(t) = 0$, which is equivalent to:

$$(r_{\max} - \mu n/2)t + \frac{m_D}{2\mu}\left(e^{-2\mu t} - 1\right) = 0.$$

This equation has a unique positive real solution, given by

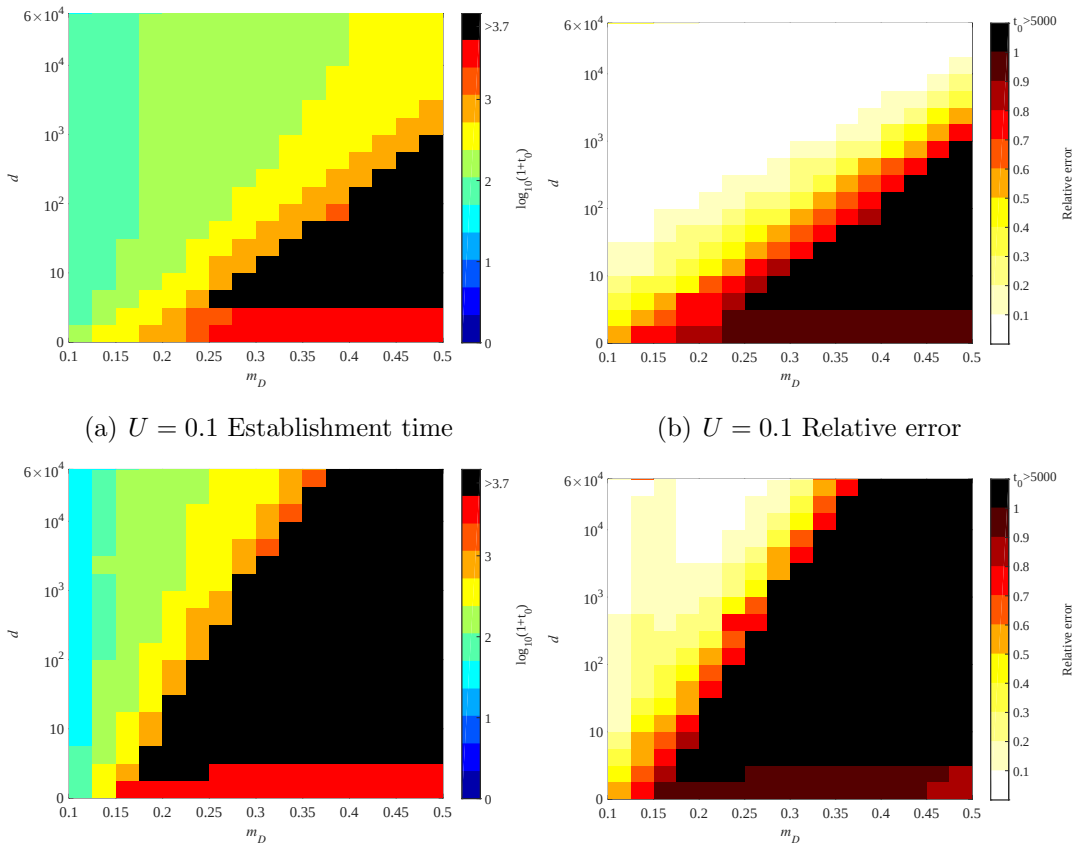
$$t_0 = \frac{1}{2\mu}\left(c + W_0(-ce^{-c})\right),$$

with $c := m_D/(r_{\max} - \mu\, n/2)$ and W_0 the principal branch of the Lambert-W function.

F Establishment time t_0: dependence with the harshness of stress m_D and the immigration rate d

Using the stochastic individual-based model of Section 2.5, we analysed the dependence of the establishment time t_0 with respect to m_D and d for a wide range of parameter values. Namely, taking $U = 0.1$, $r_{max} = 0.1$, $\lambda = 1/300$ and $n = 6$ as in Fig. 6, m_D was varied between 0.1 and 0.5. The results are presented in Fig. 7 (a). It shows that, for each value of m_D, there is a threshold value of the immigration rate above which the establishment time t_0 becomes almost independent of d. This threshold tends to increase as the harshness of stress m_D takes higher values. Additionally, we measured the relative error between the theoretical value of t_0 given by formula (12) and the value given by individual-based simulations; see Fig. 7 (b). As soon as the parameters are far from the black region in Fig. 7, (a,b), the approximation is accurate (relative error < 0.1). This black region corresponds to values of $t_0 > 5000$, for which individual-based simulations were stopped before establishment, and where we can expect that the final outcome is establishment failure. This means that there is only a narrow region where formula (12) is not accurate; it is located close to the region where establishment fails, and describes a rapid increase in t_0 which is not captured by our analytical approach.

Fig. 7 (c,d) depicts comparable simulations, with $U = U_c/3 = 0.01$, *i.e.*, outside of the WSSM regime. The conclusions are similar to the case $U = 0.1$, but with a larger region corresponding to establishment failure, and a lower accuracy (panel d).



(a) $U = 0.1$ Establishment time

(b) $U = 0.1$ Relative error

(c) $U = 0.01$, establishment time

(d) $U = 0.01$, relative error

Figure 7: **Establishment time t_0, dependence with the harshness of stress m_D and the immigration rate d.** (a,c): Average value of t_0 over 100 individual-based simulations. The color legend corresponds to $\log_{10}(1 + t_0)$. (b,d): relative error between the theoretical value of t_0 given by formula (12) and the average value obtained by individual-based simulations. The black regions correspond to parameter values for which at least one simulation led to $t_0 > 5000$; in that case, the average value of t_0 was not computed numerically. In all cases, the parameter values are $r_{max} = 0.1$, $\lambda = 1/300$, $n = 6$.