



**HAL**  
open science

## Allele fixation in a dynamic metapopulation: Founder effects vs refuge effects

R. Aguilée, D. Claessen, A. Lambert

► **To cite this version:**

R. Aguilée, D. Claessen, A. Lambert. Allele fixation in a dynamic metapopulation: Founder effects vs refuge effects. *Theoretical Population Biology*, 2009, 76 (2), pp.105-117. 10.1016/j.tpb.2009.05.003 . hal-00447327

**HAL Id: hal-00447327**

**<https://hal.science/hal-00447327>**

Submitted on 10 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Allele fixation in a dynamic metapopulation: founder effects vs refuge effects

Robin Aguilée<sup>1,\*</sup>, David Claessen<sup>1</sup>, Amaury Lambert<sup>2</sup>

Published in *Theoretical Population Biology*, 2009, **76**(2): 105–117  
with doi: 10.1016/j.tpb.2009.05.003

<sup>1</sup> Laboratory of Ecology and Evolution (UMR 7625, University Paris 06, Ecole Normale Supérieure, AgroParisTech, CNRS), Unit of Eco-Evolutionary Mathematics, F-75005 Paris, France

<sup>2</sup> Laboratoire de Probabilités et Modèles Aléatoires, University Paris 06, F-75005 Paris, France

\* *Corresponding author*

Ecole Normale Supérieure, Laboratoire d'Ecologie (UMR 7625), 46 rue d'Ulm, F-75230 Paris Cedex 05, France.

Tel: +33 1 44 32 36 97; Fax: +33 1 44 32 38 85

E-mail: robin.aguilee@biologie.ens.fr

## Abstract

The fixation of mutant alleles has been studied with models assuming various spatial population structures. In these models, the structure of the metapopulation that we call the “landscape” (number, size and connectivity of subpopulations) is often static. However, natural populations are subject to repetitive population size variations, fragmentations and secondary contacts at different spatiotemporal scales due to geological, climatic and ecological processes. In this paper, we examine how such dynamic landscapes can alter mutant fixation probability and time to fixation. We consider three stochastic landscape dynamics: (i) the population is subject to repetitive bottlenecks, (ii) to the repeated alternation of fragmentation and fusion of demes with a constant population carrying capacity, (iii) idem with a variable carrying capacity. We show by deriving a variance, a coalescent and a harmonic mean population effective size, and with simulations that these landscape dynamics generate repetitive founder effects which counteract selection, thereby decreasing the fixation probability of an advantageous mutant but accelerate fixation when it occurs. For models (ii) and (iii), we also highlight an antagonistic “refuge effect” which can strongly delay mutant fixation. The predominance of either founder effects or refuge effects determines the time to fixation and mainly depends on the characteristic time scales of the landscape dynamics.

## Keywords:

allele fixation probability; time to fixation; subdivided population; dynamic landscape; bottleneck; fragmentation; fusion; founder effect; refuge effect; variance effective size; coalescent effective size

# 1 Introduction

The study of the fixation of novel alleles has known many developments since the beginning of population genetics (Fisher 1922; Haldane 1927; Wright 1931). Fixation probabilities and times to fixation are indeed important factors influencing, among others, the rate of evolution, the genetic load (Whitlock 2002; Theodorou and Couvet 2006), and the level of genetic diversity (Vuilleumier et al. 2008). The importance of understanding and characterizing allele fixation is linked to its practical implications: for example, conservation generally tries to restore genetic diversity in small and/or fragmented populations which risk extinction (Gao and Zhang 2005; Bohme et al. 2007); in public health, maintenance of resistance alleles to drugs is a major problem (Heinemann 1999; McLean 1995).

Most natural populations are subdivided into partially isolated demes (Hanski and Gaggiotti 2004). Following Keymer et al. (2000) we call the spatial structure of a subdivided population the “landscape”; we define it as the number, the size, and the connectivity of subpopulations. The landscape strongly affects how drift and selection act (Barton and Whitlock 1997; Colas et al. 2002; Roze and Rousset 2003; Whitlock 2004). It thus influences allele fixation probability and time. Understanding these influence is of great importance especially today because of intense landscape fragmentation due to human activities; many populations consist now of small demes poorly connected, leading to high local extinction risks (Wilcox and Murphy 1985; Hanski and Gaggiotti 2004).

There is an abundant literature about mutant fixation in subdivided populations (see *e.g.* the review of Charlesworth et al. 2003; Patwa and Wahl 2008). Many spatial structures have been analyzed, in particular island, stepping-stone, spatially continuous, source-sink, and extinction-recolonization models. For populations of constant size such that migration does not change allele frequencies in the whole population, spatial structure does not affect allele fixation probability. Other spatial structures generally decrease the fixation probability of advantageous mutants.

The landscape described by most of these models is static, or at most only one component of the landscape is varying. First, the number of patches is constant over time. Second, the size of demes is often considered as constant. Many authors analyzed population size variations (one size change, exponential/logistic growth or decline, size fluctuation), but only for one isolated population (see for example Ewens 1967; Kimura and Ohta 1974; Otto and Whitlock 1997; Barton and Whitlock 1997; Wahl and Gerrish 2001; Iizuka 2001; Iizuka et al. 2002; Heffernan and Wahl 2002; Lambert 2006). Note that extinction-recolonization models could be considered as models with population size variations since each deme can become extinct. Third, the connectivity of subpopulations via migration is assumed constant over time, except in Whitlock and Barton (1997) and Whitlock (2003).

However, all components of the landscape are dynamic simultaneously in natural populations. For example, external factors can cause variations of connections between demes, to the point where connectivity either falls to its minimum (unconnected demes, *e.g.* vicariance) or rises to its maximum (fusion of demes, *e.g.* postglacial secondary contacts) (Young et al. 2002). Climatic variations as well as volcanic events can cause sea level changes resulting in separations and fusions of islands (Cook 2008). Repeated changes of the water level causing fragmentations and fusions of lakes are known in the Great African Lakes (Owen et al. 1990; Delvaux 1999; Galis and Metz 1998; Stiassny and Meyer 1999). At a different spatiotemporal scale, the number and size of populations can vary because of dispersal and recolonization events (establishment of new colonies and their later fusion) (DeHeer and Kamble 2008; Vasquez and Silverman 2008). All aspects of the spatial structure of a population can change because of new ecological interactions, *e.g.* the emergence or extinction of a predator or parasite (Batzli 1992). Contemporary frag-

mentation of habitat due to human action is also always changing the landscape (Davies et al. 2006).

These spatial processes cause, repeatedly, bottlenecks and fragmentations of subpopulations. These two phenomena are well known, but have been studied separately and, most of the time, when occurring only once. Their association and their repetition have no simple outcome regarding allele fixation: bottlenecks and fragmentations are expected to decrease fixation probability of a beneficial allele (Otto and Whitlock 1997; Wahl and Gerrish 2001; Whitlock 2003), but they can increase or decrease the time to fixation, in particular depending on the effective size of the population (Whitlock 2003). Moreover, to keep constant the number of demes of a fragmenting population, models generally assume repetitive extinctions. However, the spatial processes listed above do not necessarily lead to repetitive local extinctions. They can also lead, repeatedly, to the fusion of entire subpopulations. To our knowledge, such periodic fusions (repetitive secondary contacts) have not yet been studied regarding allele fixation, except in Jesus et al. (2006).

In this paper, we examine how such dynamic landscapes can alter fixation probability and time to fixation of a mutant allele, with or without selection. We consider three landscape dynamics: a population subject to repetitive bottlenecks (Model 1) and a population subject to the repeated alternation of fragmentation and fusion of demes (Model 2), that is, alternatively divided into two demes or undivided, with population size variations but a constant carrying capacity (Model 2a) and with a variable carrying capacity (Model 2b). Note that Wahl and Gerrish (2001) examined the effects of cyclic bottlenecks in experimental conditions, *i.e.* regular and severe bottlenecks. In contrast, we take into account the stochasticity of the occurrence of bottlenecks and any intensity of bottlenecks. We derive diffusion approximations based on the assumption of a large population. Depending on the characteristic time scales of the landscape dynamics, our models can mimic each of the spatial processes listed above. Our results constitute a first step to analyze the rate of evolution, and then speciation, in dynamic landscapes.

## 2 The Models

### 2.1 Within-deme population dynamics

We use a population genetics haploid model with two types, mutants and residents, representing individuals carrying two different alleles, respectively. This model, referred to as the Moran model or Moran process (Moran 1962), is embedded into a model of landscape dynamics, specified below. The Moran process is similar to the Wright-Fisher model (Wright 1931), but in continuous time (overlapping generations). It is a stochastic process which describes a finite population of constant size and based on the following mechanism: during an infinitesimal time  $dt$ , a birth or death event can occur or not; if it does, the population at time  $t + dt$  is updated from that of time  $t$  by randomly selecting an individual to reproduce and then, independently, randomly selecting an individual to be removed. Each individual with birth rate  $b$  has a probability  $b dt$  to reproduce during  $dt$ .

Each resident reproduces at rate  $b = 1$  and each mutant at rate  $b = 1 + s$  where  $s$  is its selective advantage (see Table 1 for a summary of the notation). For an undivided isolated population whose allele frequency fluctuates via a Moran process, classical results and approximations are known for fixation probability and time to fixation and will be used as reference results of unstructured populations in a static landscape (Wright 1931; Kimura 1962; Kimura and Ohta 1969; Ewens 2004).

Table 1: Notation and range of numerical values

Variables:		
$X_t$	Overall number of mutants at time $t$	
$x_t$	Overall frequency of mutants at time $t$	
Parameters:		
$s$	Selective advantage of mutants	Numerical values used:
$x_0$	Initial frequency of mutants	From -0.25 to 0.25
$g$	Bottleneck rate	From 0.0001 to 10
$d$	Intensity of bottlenecks	From 0 to 0.99
$f$	Fragmentation rate	From 0.0001 to 10
$c$	Fusion rate	From 0.0001 to 10
$p$	“Asymmetry parameter”	From 0.5 to 0.99
		(symmetrical to $p \in ]0; 0.5]$ )
$N$	Carrying capacity at state 1 (undivided)	From 50 to 1000
Outputs:		
$U$	Fixation probability of a mutant allele	
$T$	Time to fixation of a mutant allele, conditional on its fixation	

## 2.2 Model 1: repeated bottlenecks

Model 1 consists of a population which undergoes repeated decreases in population size (Figure 1). We are mostly interested in bottlenecks, that is, severe reductions in population size. Bottlenecks occur stochastically at exponential rate  $g$ . The higher  $g$  is, the more often a bottleneck is likely to occur. The intensity of bottlenecks is characterized by  $d$ : during a bottleneck, each individual has a probability  $d$  to die; the number of surviving individuals is thus drawn from a binomial distribution. Note that we assume that the selective advantage of mutants does not convey resistance to bottlenecks:  $d$  is identical for residents and mutants.

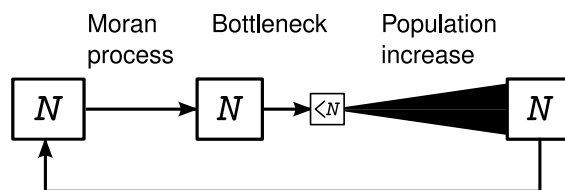


Figure 1: Model 1, repeated bottlenecks. Model 1 describes landscape dynamics which consist of repeated bottlenecks. Bottlenecks occurs at rate  $g$ . Each individual dies with probability  $d$  during a bottleneck. The size of the population is indicated at each step. After each bottleneck, the population reaches its carrying capacity  $N$  via a pure birth process. Between bottlenecks, mutant allele frequencies fluctuate via a Moran process.

Just after a bottleneck, we assume that the population reaches its carrying capacity (size  $N$ ) instantaneously. Indeed, an initial population of  $N(1 - d)$  individuals will reach its carrying capacity  $N$  in about  $-\log(1 - d)$  time units, which is much smaller than the characteristic timescale of allele frequency change (about  $N$  time units). We model this growth phase using a stochastic pure birth process. Between bottlenecks, the number of mutants,  $X_t$ , (and the mutant allele frequency  $x_t$ ) fluctuates through a Moran process. We evaluate fixation probability  $U$  and time to fixation  $T$  of a mutant allele (conditional

on its fixation) initially at frequency  $x_0$  using simulations and diffusion approximations. Note that a bottleneck can lead to population extinction when the bottleneck intensity  $d$  is high since the number of surviving individuals is stochastic. Therefore, we evaluate  $U$  and  $T$  conditional on the persistence of the population until the fixation of either the mutant allele or the resident allele.

Changing values of  $g$  and  $d$  allows us to model very different spatial (and ecological) processes. For example, a high value of  $g$  ( $g > 1$ ) with a small value of  $d$  ( $d < 0.2$ ) simulates frequent weak bottlenecks, which can correspond to periodic oscillations observed in consumer-resource systems (Turchin 2003). In contrast, a small value of  $g$  ( $g < 0.01$ ) with a high value of  $d$  ( $d > 0.8$ ) simulate rare but strong bottlenecks: it can correspond to rare and violent climatic events such as severe fires (Malhi et al. 2008).

### 2.3 Model 2a: alternations of fragmentation and fusion, constant carrying capacity

Landscape dynamics of Model 2a consist of an oscillation of the population between 1 and 2 demes (Figure 2): (i) the population consists of one deme of  $N$  individuals (state 1), (ii) the population splits into two demes (fragmentation), (iii) both demes reach their ecological equilibrium (size  $N/2$ ), (iv) the population consists of two independent isolated demes (no migration between them) of  $N/2$  individuals (state 2), (v) the two demes merge and then form one deme of  $N$  individuals (fusion, return to state 1). The fragmentation-fusion dynamics can be interpreted in a second way: the number of demes is always 2, and they are either connected via enough migration to consider the two subpopulations as only one (state 1), or they are isolated from each other (no migration, state 2). Note that in both interpretations of the dynamics, no explicit spatial structure is assumed.

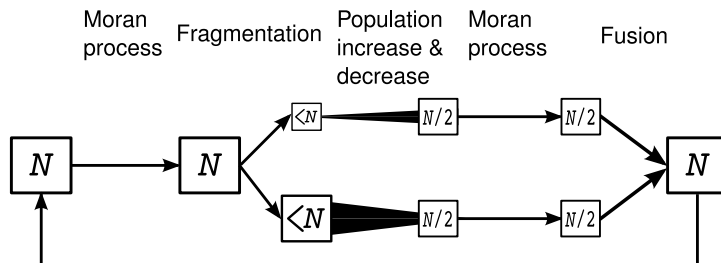


Figure 2: Model 2a, alternation of fragmentations and fusions with a constant overall carrying capacity. Model 2a describes landscape dynamics which consist of the repeated alternations of fragmentation and fusion of subpopulations. Fragmentations occurs at rate  $f$ , fusions at rate  $c$ . The size of each deme is indicated in each square depicting a (sub-)population. The parameter  $p$  defines the asymmetry of fragmentations. The overall carrying capacity of the population is kept constant. Between fragmentation and fusion events, mutant allele frequencies fluctuate via a Moran process. Just after a fragmentation, each deme reaches its carrying capacity via a pure birth or death process.

Fragmentations and fusions are modelled as stochastic events occurring respectively at exponential rates  $f$  and  $c$ . The higher  $f$  (respectively  $c$ ), the more often a fragmentation (respectively fusion) is likely to occur. At a fragmentation event, we assume that each individual has a probability  $p$  to be in the “left-hand” deme. The number of individuals in the “left-hand” deme is hence drawn from a binomial distribution. The closer  $p$  is to 0.5, the more the two demes are likely to have the same population size just after

fragmentation:  $p$  thus characterizes the asymmetry of fragmentations and will be referred to as the “asymmetry parameter”.

As in Model 1, between landscape changes (here, fragmentation and fusion events), the number of mutants  $X_t$  (and their frequency  $x_t$ ) fluctuates through a Moran process. After a fragmentation, we assume that each deme reaches its ecological equilibrium instantaneously, which we model by using a stochastic pure birth process or pure death process, depending whether the population size is above or under its equilibrium. At state 2, the carrying capacity of each deme is assumed to be  $N/2$ : the overall carrying capacity is constant equal to  $N$  because *e.g.* no more resources are available after a fragmentation and resources are equally divided between the two demes.

Values of  $f$ ,  $c$  and  $p$  can be used to model various dynamic processes. For example, landscape changes due to human action frequently destroy and recreate suitable habitats, which can divide a population into subpopulations of very different sizes and merge other previously isolated demes. If the total area of suitable habitats stays constant, Model 2a using a high value of  $f$  and  $c$  ( $> 0.1$ ) associated to a value of  $p$  very different from 0.5 ( $|p - 0.5| > 0.45$ ) will be appropriate to model such fast landscape changes. In contrast, repetitive fusions and fissions of islands due to sea level changes are best simulated with a small value of  $f$  and  $c$  ( $< 0.01$ ) and a value of  $p$  close to 0.5 ( $|p - 0.5| < 0.05$ ). Note that this scenario is relevant when the amount of resources of each island are approximately identical because we assume that the two demes have equal carrying capacities.

## 2.4 Model 2b: alternations of fragmentation and fusion, variable carrying capacity

The assumption of a constant overall carrying capacity is not relevant for all dynamic processes we aim to model, *e.g.* the fragmentation of a population due to the establishment of a new colony. Thus, we suggest a third landscape dynamics model, with carrying capacity variations. This Model 2b (Figure 3) is analogous to Model 2a except that (i) a fragmentation doubles the population size since each subpopulation carrying capacity is  $N$ , and that (ii) the overall population size is regulated (divided by 2) just after the two subpopulations merge, using a pure death process. For this model, the population fusion followed by the reduction of its size can be for example interpreted as the movement of one subpopulation into the territory of the other because its own habitat has become unsuitable (*e.g.* climatic events, arrival of a new predator). The population size is then assumed to be reduced because of increased competition.

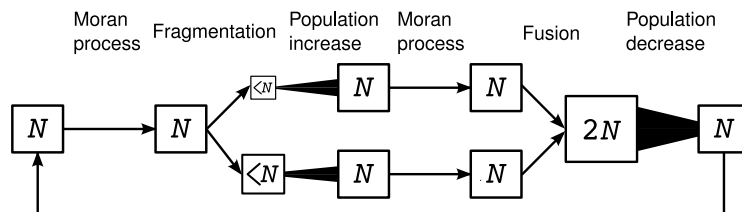


Figure 3: Model 2b, alternation of fragmentations and fusions with a variable overall carrying capacity. Model 2b is analogous to Model 2a except that the carrying capacity of each subpopulation is  $N$  and that the population size is regulated just after the two subpopulations merge, using a pure death process.

## 2.5 Reference model for a static landscape

Our goal is to understand how repetitive bottlenecks and cycles of alternating fragmentations and fusions affect the fixation probability  $U$  and the time to fixation  $T$  of a mutant allele, with or without selection. To this end, we compare results with fixation probability and time to fixation in static landscapes (*i.e.* one population of constant size). Let us recall some classical results in the latter case that we use as a reference case.

Using diffusion approximations for large populations, Kimura (1962) derived an expression for the fixation probability of a weakly advantageous allele in a Wright-Fisher undivided haploid population:  $U(x_0) \approx (1 - e^{-2Nsx_0})/(1 - e^{-2Ns})$ . As the variance in offspring number per individual is exactly twice as large in the Moran process as in the Wright-Fisher model (Ewens 2004), for a Moran process, the fixation probability of a weakly advantageous allele is

$$U(x_0) \approx \frac{1 - e^{-Nsx_0}}{1 - e^{-Ns}}. \quad (1)$$

Kimura and Ohta (1969) established approximations for the mean time to fixation, conditional on fixation, for a Wright-Fisher population. Fixation times in a Moran population can easily be deduced. In the advantageous case, the fixation time can be approximated by

$$T(x_0) \approx J_1(x_0) + \frac{1 - U(x_0)}{U(x_0)} J_2(x_0) \quad (2)$$

where  $U(x_0)$  is defined by Eq. (1),

$$J_1(x_0) = \frac{1}{s(1 - e^{-Ns})} \int_{x_0}^1 \frac{(e^{Nsx} - 1)(e^{-Nsx} - e^{-Ns})}{x(1 - x)} dx$$

and

$$J_2(x_0) = \frac{1}{s(1 - e^{-Ns})} \int_0^{x_0} \frac{(e^{Nsx} - 1)(1 - e^{-Nsx})}{x(1 - x)} dx.$$

The population effective sizes  $N_e$  given in the next sections are computed using a Moran population as the reference, that is,  $N$  terms should be replaced by  $N_e$  in Eq. (1) and (2) to obtain approximations of the fixation probability and time to fixation respectively.

## 2.6 Numerical methods

To simulate our models, we calculate the time until the next event changing allele frequencies as the expected value of this time which is the inverse of the total rate at which events occur. We then choose which event occurs according to their probability to occur at this time, which is the ratio of favorable rate over total rate. Possible events are those events modifying allele frequencies of the Moran process: the birth of a resident and the death of a mutant (occurring at rate  $X_t^i Y_t^i / (X_t^i + Y_t^i)$  where  $X_t^i$  and  $Y_t^i$  are the number of, respectively, mutants and residents in the deme  $i$  ( $i = 1, 2$ ) at time  $t$ ), or the birth of a mutant and the death of a resident (at rate  $(1 + s)X_t^i Y_t^i / (X_t^i + Y_t^i)$ ). Other possible events are bottlenecks (at rate  $g$ , only in Model 1), fragmentation of a population (at rate  $f$ , only in Model 2 when the landscape is in state 1) and fusion of two subpopulations (at rate  $c$ , only in Model 2 when the landscape is in state 2).



The stochastic pure birth process used in our three models is simulated similarly; possible events are the birth of a resident (at rate  $Y_t^i$ ) or the birth of a mutant (at rate  $(1+s)X_t^i$ ). The stochastic pure death process, used in Model 2, is analogous; residents die at rate  $(1+s)Y_t^i$  and mutants at rate  $X_t^i$ . For the reasons explained above, the duration of pure birth and death processes is not taken into account for the computation of fixation times.

We estimate the fixation probability  $U$  and the time to fixation (conditional on its fixation)  $T$  of a mutant allele initially at frequency  $x_0$  by replicating each simulation 5,000,000 times for Model 1 and 1,000,000 times for Models 2a and 2b. In Model 1,  $U$  and  $T$  are evaluated conditional on the persistence of the population until fixation of one type (mutant or resident); simulations that lead to population extinction are thus ignored. In Models 2a and 2b, estimations of  $U$  and  $T$  depend on the initial state of the landscape (state 1 or 2). To compute an average value of  $U$  and  $T$ , we started a proportion of the simulation replicates in state 2 equal to the ratio  $f/(f+c)$ , which is the long-term proportion of time spent in state 2. Note that for Model 2b, because the total population size depends on the state of the landscape and because we start simulations with the same initial mutant frequency  $x_0$  regardless of the state of the landscape, the initial number of mutants depends on the state of the landscape. The section 3.2.3 analyzes the results when there is initially one single mutant regardless of the state of the landscape.

We did simulations and numerical integrations using the GNU Scientific Library (The GSL Team 2007). We plot the results using R (R Development Core Team 2008). Error bars give 95% confidence intervals of the estimations of  $U$  and  $T$ . Table 1 gives a summary of notation, default values of parameters, and ranges of simulated values of parameters.

## 3 Results

### 3.1 Model 1: repeated bottlenecks

#### 3.1.1 Diffusion approximation and variance effective size

We use diffusion approximations (Kimura 1962; Ewens 2004) to compute the probability of fixation of an advantageous mutant allele in a large population of size  $N$  undergoing repeated bottlenecks. For large  $N$ , the mutant allele frequency  $x$  makes infinitesimal changes that can be decomposed into  $\delta x = \delta_1 x + \delta_2 x$ , where  $\delta_1 x$  is the change due to genetic drift and selection, and  $\delta_2 x$  is the change due to bottleneck events. These changes occur at a rate which depends on the current state  $x$  of the mutant frequency. The resulting diffusion approximation is then characterized by its infinitesimal mean  $a(x) = a_1(x) + a_2(x)$  and its infinitesimal variance  $\sigma(x) = \sigma_1(x) + \sigma_2(x)$ , which are the expectations of the infinitesimal changes and of their squares multiplied by their rate of occurrence. A heuristic justification of this fact can be given by the law of total probabilities applied to the different events that can occur at the *first event* changing allele frequencies (birth/death or bottleneck). We refer the reader to *e.g.* Karlin and Taylor (1981, Section 15.3) for details.

Let  $U(x_0)$  be the mutant fixation probability ( $x_0$  is the initial mutant allele frequency) and  $T_0(x_0)$  the expected time to fixation *on the event of mutant fixation*, *i.e.*  $T_0(x_0) = 0$  when all mutants are lost. The expected time to fixation *conditional on mutant fixation* is then  $T(x_0) = T_0(x_0)/U(x_0)$ . Diffusion theory then ensures that  $U$  solves the differential equation

$$\frac{1}{2}\sigma\frac{\partial^2 U}{\partial x_0^2} + a\frac{\partial U}{\partial x_0} = 0, \quad (3)$$

with boundary conditions  $U(x_0 = 0) = 0$  and  $U(x_0 = 1) = 1$ , and that  $T_0$  solves the differential equation

$$\frac{1}{2}\sigma \frac{\partial^2 T_0}{\partial x_0^2} + a \frac{\partial T_0}{\partial x_0} = -U(x_0), \quad (4)$$

with boundary conditions  $T_0(x_0 = 0) = 0$  and  $T_0(x_0 = 1) = 0$ . For a population of constant size  $N$  (static landscape), the solutions of these equations are respectively Eq. (1) and (2).

Let us first compute the infinitesimal mean and variance ( $a_1$  and  $\sigma_1$ ) due to genetic drift and selection. Recall that the birth rate of the resident is 1 and that of the mutant is  $1 + s$ . When the mutant allele frequency equals  $x$ , birth-death events occur at total rate  $(2 + s)Nx(1 - x)$ , and with probability  $(1 + s)/(2 + s)$ ,  $\delta_1 x = 1/N$ , whereas with probability  $1/(2 + s)$ ,  $\delta_1 x = -1/N$ . Then

$$E[\delta_1 x] = \frac{s}{N(2 + s)} \quad \text{and} \quad E[(\delta_1 x)^2] = \frac{1}{N^2}.$$

Multiplying by the total rate, we get

$$a_1(x) = sx(1 - x) \quad \text{and} \quad \sigma_1(x) = \frac{(2 + s)x(1 - x)}{N}.$$

The computation is more technical for bottlenecks. Recall that bottlenecks occur at rate  $g$ , and that each individual independently survives the bottleneck with probability  $1 - d$ . Given the numbers  $Nx$  of mutants and  $N(1 - x)$  of residents, the numbers  $Z_m$  of mutant survivors and  $Z_r$  of resident survivors are independent binomial variables with probability  $1 - d$  and respective parameters  $Nx$  and  $N(1 - x)$ . By the central limit theorem, one can write

$$Z_m \approx Nx(1 - d) + \sqrt{Nx}W_m \quad \text{and} \quad Z_r \approx N(1 - x)(1 - d) + \sqrt{N(1 - x)}W_r,$$

where  $W_m$  and  $W_r$  are two independent centered normal variables with variance  $d(1 - d)$ . Assuming that growth phases following a bottleneck event do not change significantly allele frequencies, the frequency change  $\delta_2 x$  due to the bottleneck event can then be written as

$$\delta_2 x = \frac{Z_m}{Z_m + Z_r} - x \approx \frac{(1 - x)\sqrt{Nx}W_m - x\sqrt{N(1 - x)}W_r}{N(1 - d) + \sqrt{Nx}W_m + \sqrt{N(1 - x)}W_r} \approx \frac{(1 - x)\sqrt{x}W_m - x\sqrt{(1 - x)}W_r}{\sqrt{N(1 - d)}}.$$

Then

$$E[\delta_2 x] \approx 0 \quad \text{and} \quad E[(\delta_2 x)^2] \approx \frac{x(1 - x)d}{N(1 - d)}.$$

Multiplying by the total rate  $g$  yields ( $a_2$  and)  $\sigma_2$ . Then recalling the values of  $a_1$  and  $\sigma_1$ , we get

$$a(x) = sx(1 - x) \quad \text{and} \quad \sigma(x) = \frac{x(1 - x)}{N} \left[ (2 + s) + \frac{gd}{1 - d} \right].$$

Then solving Eq. (3) and Eq. (4) assuming weak selection, (*i.e.*  $s$  is negligible compared to 1), we get the same expressions for  $U$  and  $T$  as those in a static landscape,

Eq. (1) and (2) respectively, but where all  $N$  terms are replaced by the *variance effective population size* defined by

$$N_e^v = \frac{N}{1 + \frac{gd}{2(1-d)}}. \quad (5)$$

Another way of presenting our results is to consider the bottleneck phase as a (rescaled) time period during which the population undergoes pure drift and no selection. Indeed, the infinitesimal variance  $\sigma(x)$  was expressed as the sum of the infinitesimal variance  $\sigma_1(x) = 2x(1-x)/N$  due to genetic drift (assuming weak selection), and of the infinitesimal variance  $\sigma_2(x) = gdx(1-x)/(1-d)N$  due to allele frequency change during bottlenecks. Then one can conclude that the effect of bottlenecks can be compared to that of genetic drift accelerated by a factor  $gd/2(1-d)$ . Put another way, the expected effects of bottlenecks on  $t = 1/g$  time units (*i.e.* on average one bottleneck) are equal to those of pure genetic drift on  $t' = d/2(1-d)$  time units during which selection would be relaxed.

### 3.1.2 Coalescent effective size

Let us now compute a *coalescent effective population size*. Relying on Sano et al. (2004), we define the coalescent effective size  $N_e^c$  as twice the expected coalescence time of a uniformly sampled pair of genes under neutrality (there is a factor 2 because we use a Moran population as the reference; see section 2.6). In a population of constant size  $N$  where allele frequencies fluctuate via a Moran process with birth rate 1, the coalescence rate of two genes is  $2/N$ . Then the expectation of the coalescence time is  $N/2$ , thereby explaining the definition of  $N_e^c$ .

In Model 1, two genes coalesce either during the Moran process, *i.e.* between bottlenecks, or during the growth phase following each population size reduction. Between bottlenecks, the population is of constant size  $N$ , so that the coalescence rate of two genes is  $2/N$ . Let us now compute the probability of coalescence during the growth phase. For a population growing from size  $N_0$  to size  $N_1$  ( $N_1 \geq N_0$ ), let  $k_i$  be the number of descendants (in the population of size  $N_1$ ) of each individual  $i$  (in the population of size  $N_0$ ). Two genes sampled at random among  $N_1$  have coalesced during this growth phase with probability

$$\sum_{i=1}^{N_0} \frac{k_i(k_i - 1)}{N_1^2}.$$

Because the variance of the number of descendants is small for a pure birth process, we can replace  $k_i$  with its expectation  $N_1/N_0$ . We can then write that two genes coalesce with probability

$$\frac{1}{N_0} - \frac{1}{N_1}.$$

In Model 1, the expected size of the population at a bottleneck event is  $N_0 = N(1-d)$ , and after the growth phase, the population size is  $N_1 = N$ . As bottlenecks occur at rate  $g$ , two genes coalesce during growth phases following bottlenecks at rate

$$g \left( \frac{1}{N(1-d)} - \frac{1}{N} \right).$$

Recalling that in Model 1, two genes coalesce either during Moran processes or during growth phases, the coalescence rate is then

$$g \left( \frac{1}{N(1-d)} - \frac{1}{N} \right) + \frac{2}{N}$$

and therefore the coalescent effective population size is

$$N_e^c = \frac{N}{1 + \frac{gd}{2(1-d)}}. \quad (6)$$

Using this coalescent effective size or the variance effective size (Eq. (5)) thus leads to the same approximation for  $U$  and  $T$ .

### 3.1.3 Comparison with harmonic mean and simulations

We can compare the previous variance/coalescent effective size to the *harmonic mean effective size*. Textbooks (see *e.g.* Wright 1938; Kimura 1970; Gillespie 2004) indeed suggest that when a population is subject to rapid size variations, the harmonic mean effective size  $N_e^h$  is applicable. Let us compute it. Between bottlenecks, population size equals  $N$  for  $1/g$  time units on average. At the time of a bottleneck, population size shrinks instantaneously from size  $N$  to expected size  $N(1-d)$ . The population then exponentially grows to size  $N$ . Assuming that each individual gives birth at rate 1, the growth phase lasts  $-\log(1-d)$  time units. Then we get

$$N_e^h \approx \frac{1/g - \log(1-d)}{\frac{1}{gN} + \int_0^{-\log(1-d)} \frac{1}{N(1-d)e^t} dt} \approx \frac{N(1 - g \log(1-d))}{1 + \frac{gd}{1-d}}. \quad (7)$$

Note that when  $g$  and/or  $d$  vanish (no bottleneck), all effective population sizes we have computed (Eq. (5), (6) and (7)) reduce to that in a static landscape, *i.e.*  $N_e^v = N_e^c = N_e^h = N$ .

Figure 4 shows that repeated bottlenecks decrease the fixation probability  $U$  of an advantageous mutant allele, compared to a static landscape:  $U$  decreases when the intensity of bottlenecks  $d$  increases and when bottlenecks are more frequent (higher bottleneck rate  $g$ ). A bottleneck is indeed very similar to a founder event: it is a sampling event of a few individuals of the population. Thus, as we consider alleles that do not confer resistance to bottlenecks, bottlenecks generate additional genetic drift counteracting the effect of selection with an efficiency which increases with the intensity of bottlenecks  $d$  and with their frequency  $g$ . Moreover, increasing  $g$  decreases the expected duration between the emergence of a mutant and the first bottleneck event. Otto and Whitlock (1997) showed that the closer in time to a bottleneck event a mutant appears, the more its fixation probability is reduced. This effect is added to the accumulation of drift due to bottlenecks to reduce  $U$ .

Regarding the estimation of the fixation probability, Figure 4 shows that neither the variance/coalescent effective size (Eq. (5)/Eq. (6)) nor the harmonic mean effective size (Eq. (7)) is better than the other for all bottleneck intensities and frequencies. For rare bottlenecks ( $g = 0.001$ ), all approximations give almost identical results. The variance/coalescent effective size does better for weak bottlenecks ( $d < 0.7$ ) than for strong bottlenecks. The derivation of the variance effective size  $N_e^v$  (Eq. (5)) allows to understand why: for strong bottlenecks, the assumption of small frequency variations which is necessary for a diffusion approximation to hold would not be satisfied. In particular, the expected number of survival individuals,  $N(1-d)$ , can be very small, so that mutant frequency may jump from any frequency to 0 in one generation (if all mutants die during the bottleneck). This might explain why this approximation overestimates  $U$  for strong

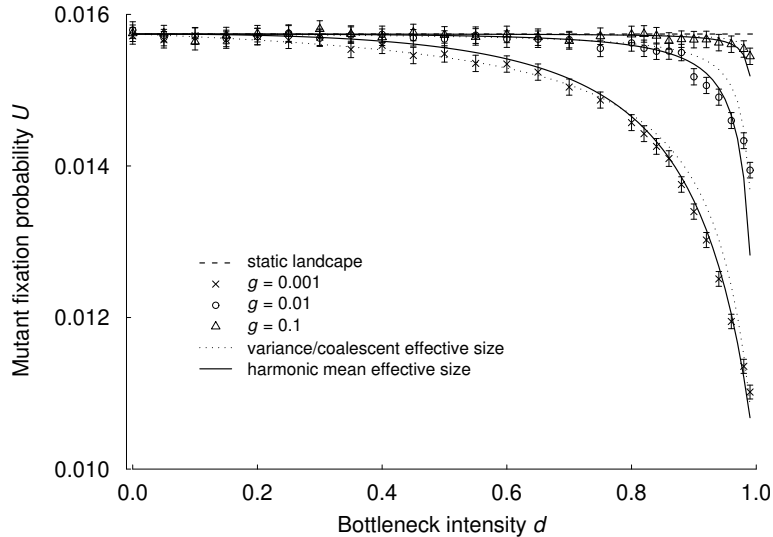


Figure 4: Model 1, fixation probability  $U$  of an advantageous mutant allele vs bottleneck intensity  $d$ . “Static landscape” (dashed line) refers to an undivided population where allele frequencies fluctuate via a Moran process (Eq. (1)). For a beneficial mutant allele, the higher the bottleneck rate  $g$  and/or bottleneck intensity  $d$ , the more additional drift due to bottlenecks, the smaller fixation probabilities. Dotted and solid lines correspond to the approximation of  $U$  obtained using respectively the variance/coalescent (Eq. (5)/Eq. (6)) and harmonic mean (Eq. (7)) effective population size. None of these approximations precisely predicts  $U$  for all values of parameters. *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.01$ .

bottlenecks. Nevertheless, the variance/coalescent effective size does better than the harmonic mean effective size for frequent and weak bottlenecks ( $d < 0.7$  and  $g = 0.1$ ). The harmonic mean effective size  $N_e^h$  seems to be the most robust approximation in the sense that its fit with simulations is overall acceptable regardless of parameter values.

Figure 5 shows that repeated bottlenecks decrease the time to fixation  $T$  of an advantageous mutant allele compared to a static landscape (Eq. (2)) and that  $T$  decreases with the intensity  $d$  and/or the rate of bottlenecks  $g$ . This is due to the fact that each time a bottleneck occurs, it generates additional drift that counteracts selection, in disfavor of the mutant (regarding its fixation probability). In contrast, conditional upon fixation, drift acts in favor of the allele to fix, so that the cumulative action of drift generated by bottlenecks can strongly reduce  $T$ : for the range of parameters values explored in Figure 5,  $T$  can be divided by 10 compared to a population of constant size (static landscape).

Figure 5 shows that both the variance/coalescent and harmonic mean effective sizes overestimate  $T$ . Sudden jumps from any frequency to 1 are possible in simulations and may explain why  $N_e^v$  overestimates  $T$ . Interestingly, none of the effective sizes can give an accurate prediction for both the fixation probability and the time to fixation.

### 3.2 Model 2: alternation of fragmentations and fusions

Defining  $U_1(x)$  as the fixation probability of mutants in frequency  $x$  appearing when the landscape is in state 1 and  $U_2(x, y)$  as the fixation probability of mutants appearing in state 2 in frequency  $x$  in one deme and  $y$  in the other deme, it is possible to apply a similar method as the one used in the case of Model 1 to derive an expression for the variance effective size. This method did not prove as fruitful for Models 2a and 2b, as it leads to a set of two integro-partial differential equations that we were unable to solve analytically. However, we have derived a coalescent effective size. We will see that it is usable only to predict times to fixation. Thus, for Models 2a and 2b, we explored with

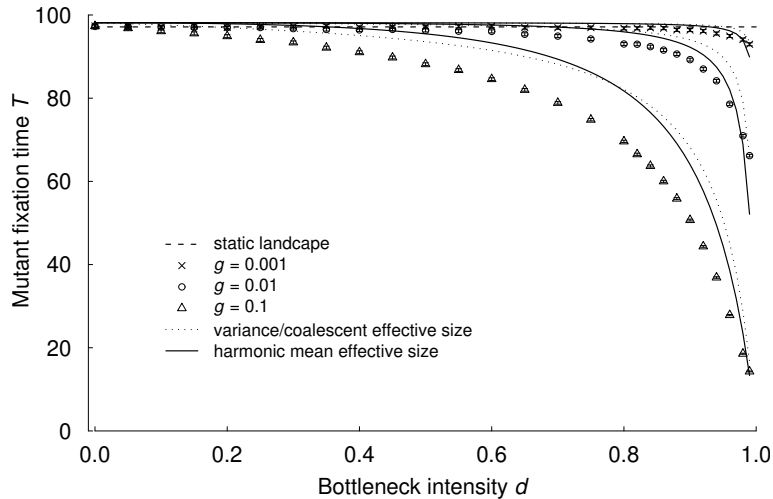


Figure 5: Model 1, fixation time  $T$  of an advantageous mutant allele vs bottleneck intensity  $d$ . “Static landscape” (dashed line) refers to an undivided population where allele frequencies fluctuate via a Moran process (Eq. (2)). The higher the bottleneck rate  $g$  and/or bottleneck intensity  $d$ , the more additional drift due to bottlenecks, the smaller fixation times  $T$ . Dotted and solid lines correspond to the approximation of  $T$  obtained using respectively the variance/coalescent (Eq. (5)/Eq. (6)) and harmonic mean (Eq. (7)) effective population size. All approximations tend to overestimate  $T$ . *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.01$ .

simulations a large range of values of all parameters (see Table 1).

### 3.2.1 Coalescent effective size

We can derive for Models 2a and 2b a coalescent effective size of the population,  $N_e^c$ , defined as the expected coalescence time of a uniformly sampled pair of genes under neutrality. Two randomly sampled genes can be either (i) in the same population when the landscape is in state 1 (“state 1 genes”), or (ii) in the same subpopulation when the landscape is in state 2 (“state 2 genes”), or (iii) in different subpopulations when the landscape is in state 2 (“state 2’ genes”).

Figure 6 indicates the backward rates of transition between each configuration of sampled genes and the coalescence rates. First, consider a pair of state 1 genes. They can coalesce during the Moran process: as they are in a population of constant size  $N$ , coalescence occurs at rate  $2/N$ . If they do not coalesce, they were previously state 2 or 2’ genes. Forward in time, fragmentation occurs at rate  $f$ ; backward in time, the transition from state 1 genes to state 2 or 2’ genes thus also occurs at rate  $f$ . As at the time of fragmentation each gene is independently segregated between the two demes, state 2 and 2’ genes are equiprobable. Therefore, both transition from state 1 genes to state 2 and 2’ genes occurs at rate  $f/2$ . Second, consider a pair of state 2 genes. They are in a population of constant size  $N/\alpha$  ( $\alpha = 2$  for Model 2a,  $\alpha = 1$  for Model 2b). They can coalesce during the Moran process, which thus occurs at rate  $2\alpha/N$ . As fusion occurs at rate  $c$ , state 2 genes can also coalesce during the birth process following fragmentation at rate  $c\gamma$ , where  $\gamma$  is the probability that two genes coalesce during the birth process following fragmentation. Finally, state 2 genes coalesce at rate  $2\alpha/N + c\gamma$ . State 2 genes can also switch to state 1 at rate  $c(1 - \gamma)$ . Last, consider a pair of state 2’ genes. They cannot coalesce because they are not in the same deme. State 2’ genes were thus previously necessarily state 1 genes. As fusion occurs at rate  $c$ , this transition occurs at rate  $c$ .

Knowing these coalescence rates and rates of transition between each state, we can write  $\tau_i$  defined as the expected coalescence time of a pair of state  $i$  genes ( $i = 1, 2, 2'$ ).

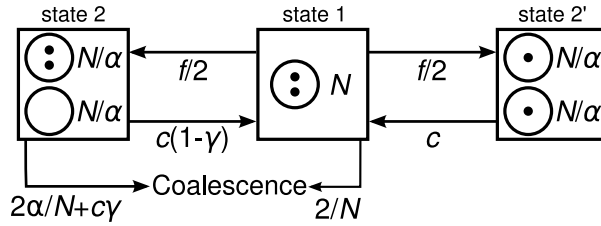


Figure 6: Models 2a and 2b, states, transitions rates, and coalescence rates of two uniformly sampled genes. We consider two uniformly sampled genes symbolized by two points. Squares indicate the possible states of these genes: (i) both in an undivided population (state 1 genes), (ii) both in the same deme (state 2 genes), (iii) each one in a different deme (state 2' genes). Circles depict (sub-)populations, their size is written on their left. In Model 2a,  $\alpha = 2$ ; in Model 2b,  $\alpha = 1$ .  $\gamma$  is the probability that two genes coalesce during the birth process following fragmentation. Arrows show the possible transitions between states. They point backward in time. Under each arrow is indicated the rate of the associated transition. See section 3.2.1 for their derivation.

Each  $\tau_i$  equals the sum of the expected time to the first event, plus, because this first event is not necessarily a coalescence event, the expected coalescence time  $\tau_j$  of state  $j \neq i$  genes weighted by the probability that the first event is a transition from state  $i$  to state  $j$  genes. Then

$$\begin{cases} \tau_1 = \frac{1}{f+2/N} + \frac{f/2}{f+2/N}\tau_2 + \frac{f/2}{f+2/N}\tau_{2'} \\ \tau_2 = \frac{1}{c+2\alpha/N} + \frac{c(1-\gamma)}{c+2\alpha/N}\tau_1 \\ \tau_{2'} = \frac{1}{c} + \tau_1 \end{cases} \quad (8)$$

Let us now compute  $\gamma$ , the probability that two genes coalesce during the birth process. Its expression differs between Model 2a and 2b. We have shown in section 3.1.2 that the coalescence rate of two genes in a population growing from size  $N_0$  to size  $N_1$  is roughly  $1/N_0 - 1/N_1$ . In Model 2a, if the fragmentation is such that  $p \leq 1 - p$ , one deme grows from expected size  $N_0 = Np$  to size  $N_1 = N/2$ , whereas the size of the other deme decreases (coalescence is impossible in this deme). A pair of state 2 genes are in the growing deme with probability  $1/2$  and in the shrinking deme with probability  $1/2$ . Therefore, if  $p \leq 1 - p$ ,  $\gamma = (1/2)[1/(Np) - 1/(N/2)]$ . Similarly, if  $p \geq 1 - p$ , one deme grows from expected size  $N(1 - p)$  to size  $N/2$ , the size of the other deme decreases, and state 2 genes are in each of these two demes with the same probability. In that case,  $\gamma = (1/2)[1/(N(1 - p)) - 1/(N/2)]$ . Therefore, for Model 2a,

$$\gamma = \frac{|1 - 2p|}{2N \min(p, 1 - p)}. \quad (9)$$

In Model 2b, after a fragmentation the sizes of both demes increase: one deme grows from size  $Np$  to size  $N$ , the other from size  $N(1 - p)$  to size  $N$ . As a pair of state 2 genes is in each of these demes with equal probability, the probability that they coalesce during the birth process is

$$\gamma = \frac{1}{2} \left( \frac{1}{Np} - \frac{1}{N} \right) + \frac{1}{2} \left( \frac{1}{N(1 - p)} - \frac{1}{N} \right) = \frac{1 - 2p(1 - p)}{2Np(1 - p)}. \quad (10)$$

In a constantly undivided population of constant size  $N_e^c$ , two uniformly sampled genes would always be state 1 genes, with an expected coalescence time  $\tau_1 = N_e^c/2$ . Consequently, solving Eq. (8) gives the effective coalescent size for Models 2a and 2b

$$N_e^c = \frac{2N(f+c) + 2\alpha f/c + 4\alpha}{2c + \alpha f + fc\gamma N/2 + 4\alpha/N} \quad (11)$$

where  $\alpha = 2$  and  $\gamma$  is defined by Eq. (9) for Model 2a, and where  $\alpha = 1$  and  $\gamma$  is defined by Eq. (10) for Model 2b. Replacing  $N$  terms by this expression of  $N_e^c$  into Eq. (1) and (2) should give an approximation of, respectively, the fixation probability  $U$  and time to fixation  $T$  of an advantageous allele. Let us first analyze the results obtained using simulations, and then compare the approximations of  $U$  and  $T$  obtained using  $N_e^c$  to simulations.

### 3.2.2 Fixation probability of advantageous alleles

To provide some intuition on the allele dynamics in fragmentation-fusion models, Figure 7 shows a typical time series from simulation runs with an extreme value of the asymmetry parameter ( $p = 0.99$ ). Asymmetrical fragmentations can strongly affect allele frequencies: one newly formed deme is large (its size is close to that of the undivided population), so that just after fragmentation the mutant allele frequency in this deme is close to the overall mutant frequency just before fragmentation. The other deme is small, allowing a founder effect, as after a bottleneck in Model 1. As a result, the more asymmetrical the fragmentation, the stronger the founder effect, the more the mutant frequency in the whole population is likely to change strongly after a fragmentation, as illustrated in Figure 7.

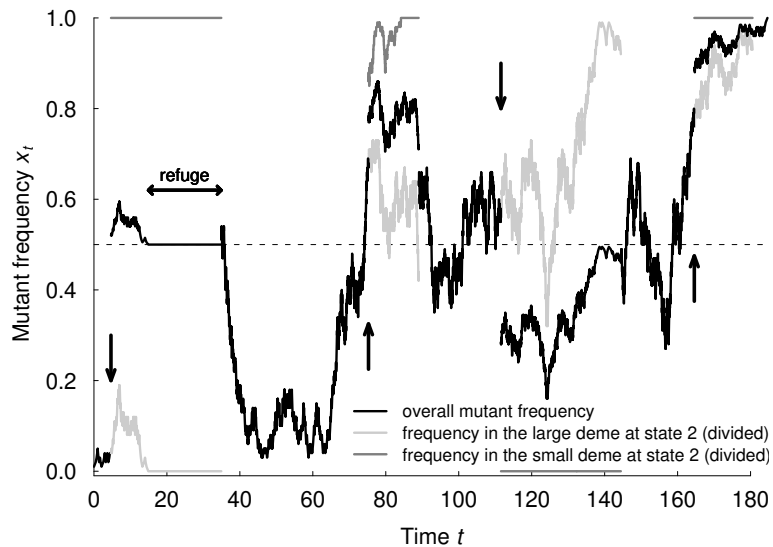


Figure 7: Model 2b, mutant frequency  $x_t$  vs time  $t$ , with asymmetrical fragmentations. This figure represents one single simulation of the Model 2b with asymmetrical fragmentations ( $p = 0.99$ ), ended by mutant fixation (similar patterns are obtained with Model 2a). The overall mutant frequency changes abruptly after each fragmentation (indicated by a vertical arrow) and mutant frequencies in each deme are very different: mutant frequency equals 0 or 1 in the “small” deme (founder events) whereas mutant frequency is close to the overall mutant frequency in the “large” deme. Note that as long as the mutant is absent from one deme, its global frequency cannot exceed 0.5 (dashed line). In this simulation, at time 15, a refuge is briefly created, mutant frequency then cannot change until a fusion occurs (here, at time 35). Except at the second fragmentation, “half” a refuge is created by fragmentation (*i.e.* fragmentation generates a monomorphic deme). *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.05$ ;  $p = 0.99$ ;  $f = c = 0.05$ .

The fragmentation rate  $f$  scales the expected time spent between fusion and fragmentation: increasing  $f$  increases the frequency of fragmentations, which increases the



cumulative number of possible founder events whose intensity is determined by  $p$ . Using our results of Model 1, we can formulate expectations based on the hypothesis that the founder effect is the principal factor in the allele dynamics. Since in Model 1 the bottleneck rate  $g$  sets the cumulative number of founder events due to bottlenecks and  $d$  their intensity, we should observe the same effects of  $f$  and  $p$  on fixation probability in Models 2a and 2b as the effects of  $g$  and  $d$  respectively in Model 1. Moreover, increasing the fusion rate  $c$  increases the frequency of fusions, and thus increases the frequency of fragmentations (if a fusion occurs earlier, the following fragmentation occurs earlier). Consequently, we expect that  $c$  alters fixation probabilities in the same way as  $f$  does.

As this reasoning predicts, in the advantageous case, compared to a static landscape the mutant fixation probability  $U$  decreases with the asymmetry of fragmentations  $p$  (Figure 8) and/or with the fragmentation rate  $f$  (Figure 9) and/or with the fusion rate  $c$  (Figure 10). The reference model is the same for Models 2a and 2b (static landscape, Eq. (1)). However, in Model 2b the overall population size fluctuates between  $N$  (state 1) and  $2N$  (state 2). The population size used in Eq. (1) for Model 2b is thus the harmonic mean of overall carrying capacities (Gillespie 2004),  $(f + c)/(c/N + f/(2N))$ , which is greater than  $N$  (*i.e.* the carrying capacity of the population in Model 2a at any time). As a result,  $U$  is higher in Model 2b than in Model 2a (Figures 8, 9 and 10). However, the effect of the landscape dynamics relative to a static landscape is identical in both models:  $U$  is reduced when the landscape dynamics are fast enough and fragmentations asymmetrical enough.

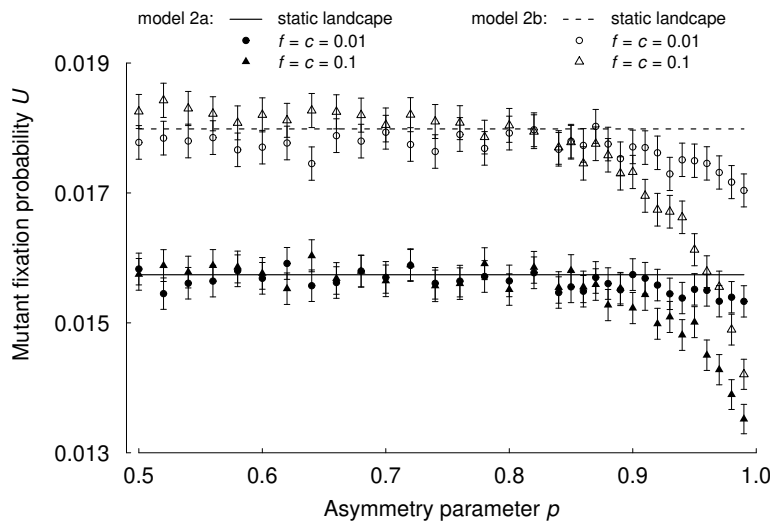


Figure 8: Model 2, fixation probability  $U$  of an advantageous mutant allele vs asymmetry parameter  $p$ . “Static landscape” (lines) refers to an undivided population where allele frequencies fluctuate via a Moran process (Eq. (1)). For this reference case, the population size (computed as the harmonic mean of population sizes) is higher in Model 2b than in Model 2a so that the static landscape fixation probability is higher in Model 2b (dashed line) than in Model 2a (solid line). For both models, in the advantageous case, when the asymmetry parameter moves away from 0.5 and/or when the landscape dynamics speed up ( $f$  and/or  $c$  increasing), fixation probabilities get smaller (stronger and/or more frequent founder effects generating additional drift). The definition of the asymmetry parameter  $p$  is such that results are symmetrical in  $p = 0.5$ . We thus plot results for  $p \in [0.5; 1]$  only. *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.01$ .

We could use the coalescent effective size  $N_e^c$  (Eq. (11)) to predict  $U$ . When  $f$  tends to zero, *i.e.* when the landscape tends to be most of the time in state 1,  $N_e^c$  tends to  $N$ . This is consistent with results of Figure 9: for small values of  $f$ , the fixation probability

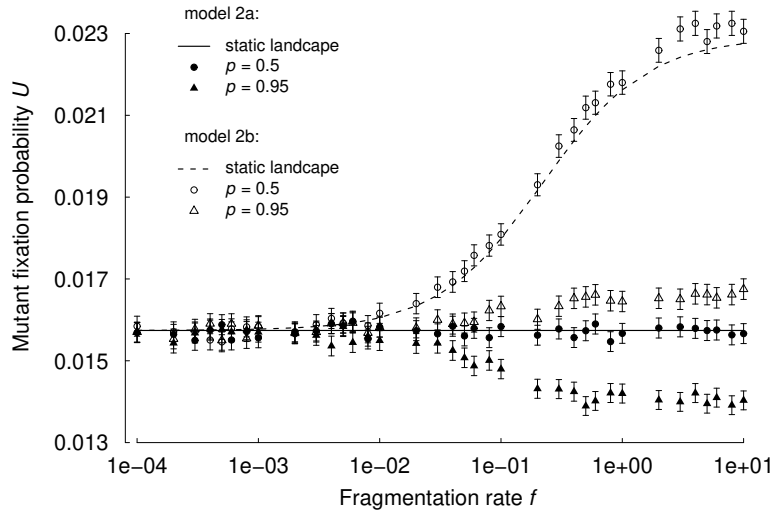


Figure 9: Model 2, fixation probability  $U$  of an advantageous mutant allele vs fragmentation rate  $f$ . For both Models 2a and 2b, in the advantageous case with symmetrical fragmentations ( $p = 0.5$ , circles), the fragmentation rate  $f$  does not alter the fixation probability compared to a static landscape (lines). For Model 2b the fixation probability increases with  $f$  (dashed line) because the mean overall population size increases with  $f$ . In contrast, with asymmetrical fragmentations ( $p = 0.95$ , triangles), the higher the fragmentation rate, the smaller the fixation probabilities relative to a static landscape (more frequent founder effects generating additional drift). *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.01$ ;  $c = 0.1$ .

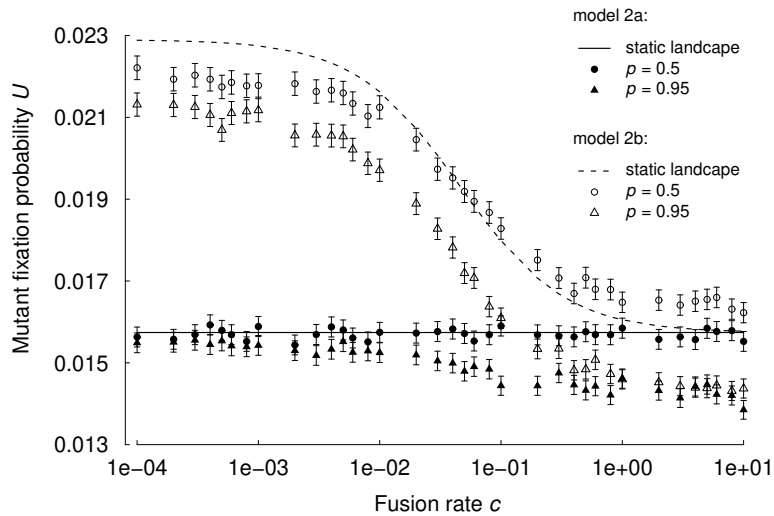


Figure 10: Model 2, fixation probability  $U$  of an advantageous mutant allele vs fusion rate  $c$ . Results of this figure are analogous to these of Figure 9. For both Models 2a and 2b, in the advantageous case with symmetrical fragmentations ( $p = 0.5$ , circles), the fusion rate  $c$  does not alter the fixation probability compared to a static landscape (lines). For Model 2b the fixation probability decreases with  $c$  (dashed line) because the mean overall population size decreases with  $c$ . With asymmetrical fragmentations ( $p = 0.95$ , triangles), the higher the fusion rate, the smaller the fixation probabilities relative to a static landscape (more frequent founder effects generating additional drift). *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.01$ ;  $f = 0.1$ .

of mutants tends to be identical to that of a static landscape. However, when  $c$  tends to zero, *i.e.* when the landscape tends to be most of the time in state 2,  $N_e^c$  tends to infinity. It means that fixation tends to become impossible when  $c$  decreases. But as long as  $c$  is not zero, fixation is nevertheless possible.  $N_e^c$  should thus tend to the effective size of a population in a static landscape, as simulations of Figure 10 show it. Therefore, for small

values of  $c$ , using the coalescent effective size  $N_e^c$  does not make sense to predict  $U$ .

### 3.2.3 State-dependent fixation probability of advantageous alleles

Figures 8, 9 and 10 give the fixation probabilities of mutants arising in state 1 or 2 averaged over the relative duration of state 1 and 2 (see section 2.6). However, the fate of one particular mutant depends on the state of the landscape when it appears.

Figure 11 shows that for Model 2b, a mutant is more likely to be fixed if it appears in state 1 (“state 1 mutant”) than in state 2 (“state 2 mutant”). This is principally because we consider here only one initial mutant; the initial overall frequency of mutants arising in state 1 is thus twice that of those arising in state 2. Figure 11 shows that for Model 2a without founder events ( $p$  close to 0.5), the mutant fixation probability is independent of the state of the landscape when the mutant arises. In Model 2a, only fragmentations may alter the fixation probability of an advantageous mutant; fusion events have no effect on  $U$ . With symmetrical fragmentations, there is no population size variation, and state 1 and 2 mutant fixation probabilities are thus identical. With founder events ( $p$  close to 1), a mutant is more likely to be fixed if it appears in state 2 than in state 1 (Figure 11). Let us recall that founder events are analogous to bottlenecks and that Otto and Whitlock (1997) showed that the closer in time to a bottleneck event a mutant appears, the more its fixation probability is reduced. A state 1 mutant appears necessarily closer in time to the first fragmentation event than a state 2 mutant, which explains why its fixation probability is smaller. Moreover, taking into account only the first fragmentation event is enough to understand this result because each fragmentation-fusion cycle reinforces the effect of the first cycle.

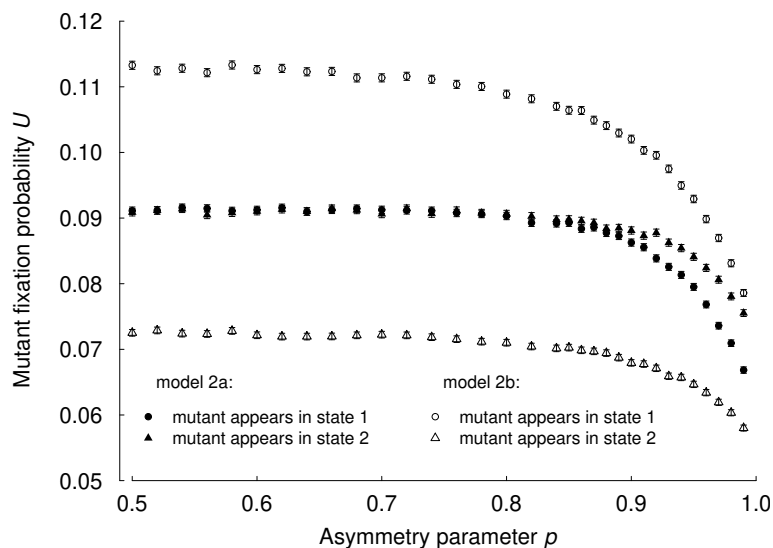


Figure 11: Model 2, state-dependent fixation probability  $U$  of an advantageous mutant allele vs asymmetry parameter  $p$ . Circles (respectively triangles) correspond to the fixation probability of one particular advantageous mutant which appears when the landscape is in state 1 (respectively state 2). Mutants are always assumed to arrive during the Moran process (*i.e.* after the possible population size variations following fragmentation and fusion). Without founder events ( $p$  close to 0.5), for Model 2a the mutant fixation probability is independent of the state of the landscape when the mutant appears, whereas for Model 2b an advantageous mutant is more likely to be fixed if it appears in state 1 than in state 2. With founder events ( $p$  close to 1), for Model 2a a mutant is more likely to be fixed if it appears in state 2 than in state 1, whereas the opposite result is still valid for Model 2b. *Parameter values:*  $N = 100$ ;  $X_{t^*} = 1$  where  $t^*$  is when the mutant appears;  $s = 0.1$ ;  $f = c = 0.1$ .

### 3.2.4 Time to fixation of advantageous alleles

We measure fixation time  $T$  as the mean time to fixation, conditional on the fixation of mutants. Fixation times can either be strongly decreased or increased compared to the case of a static landscape, depending on the relative strength of founder effects and so-called “refuge effects”. We use the term “refuge” (Figure 7) to refer to the situation in which *the mutant is fixed in one deme while the resident is fixed in the other deme* (and the landscape is necessarily in state 2). Allele frequencies then cannot change until the next fusion; neither mutants nor residents can invade or become extinct. Therefore, the effect of a refuge is to delay fixation. We quantify this “refuge effect” by the length of the time spent in refuges. The fusion rate  $c$  affects the intensity of refuge effects since it scales the waiting time until the next fusion. Besides, the asymmetry parameter  $p$  can alter refuge effects: the more asymmetrical fragmentations, the more the small deme formed by fragmentation is likely to be monomorphic, so that “half” a refuge is created by fragmentation, as it happens in the simulation showed in Figure 7.

For Model 2a, when fragmentations are symmetrical, fragmentations do not generate founder effects, so that only refuge effects affect the mutant fixation time  $T$ . Results of simulations in Figure 12 (with  $p = 0.5$ ) indeed show that in this case, the fusion rate  $c$  strongly influences  $T$ . For a small fusion rate ( $c < 0.05$ ), the time between fragmentation and fusion is long, so that if a refuge appears, the associated refuge effect can be strong.  $T$  is then strongly increased:  $T$  can be more than 10 times longer than in a static landscape. In contrast, with a high fusion rate ( $c > 0.05$ ), refuge effects are negligible. In this case, even if refuges are formed, the time until the next fusion is short and does not significantly delay fixation. Fixation times then approximately equal those in a static landscape (Eq. (2)).

In the case of asymmetrical fragmentations, fragmentations generate founder effects which accelerate fixation because of additional drift they induce, as do bottlenecks in Model 1. However, they can also increase the strength of refuge effects because founder events facilitate the formation of refuges. Figure 12 suggests that there is a phase transition in parameter  $c$ . For small fusion rates ( $c < 0.05$ ), refuge effects are so strong that the acceleration of fixation due to founder effects is not significant. Fixation times are then very long, identical to those in the case of symmetrical fragmentations. For high fusion rates ( $c > 0.05$ ), refuge effects are negligible and founder effects speed up fixation which then becomes faster than in a static landscape. Even if asymmetrical fragmentations facilitate the formation of refuges, the time spent in refuges is short and does not significantly delay fixation.

For Model 2b (Figure 13), fixation times  $T$  are similar to those of Model 2a except that the mean overall population size is greater than  $N$ , so that  $T$  is higher in Model 2b than in Model 2a. However, relative to a static landscape,  $T$  is almost identical for both models (in Model 2b, for a high fusion rate  $c$ , even with symmetrical fragmentations,  $T$  is slightly smaller than in a static landscape due to the pure death process following fusions).

Figures 12 and 13 show that the coalescent effective population size  $N_e^c$  we have derived (Eq. 11) is accurate to predict times to fixation under weak founder effects or weak refuge effects. When founder effects are strong,  $N_e^c$  tends to underestimate them; when refuge effects are strong,  $N_e^c$  predicts their existence, but significantly underestimates their intensity. Simplifying the expression of  $N_e^c$  highlights both founder effects and refuge effects, and under which landscape dynamics parameters they take place. Let us write the most noteworthy. First, consider that the landscape dynamics is fast enough so that refuge effects are small ( $c \gg 1/N$ , and  $f/c \sim 1$  or  $f/c \ll 1$ ) and that fragmentations are symmetrical ( $p \approx 1/2$ ). Then, for Model 2a,  $N_e^c$  approximately equals  $N$ . For Model 2b,

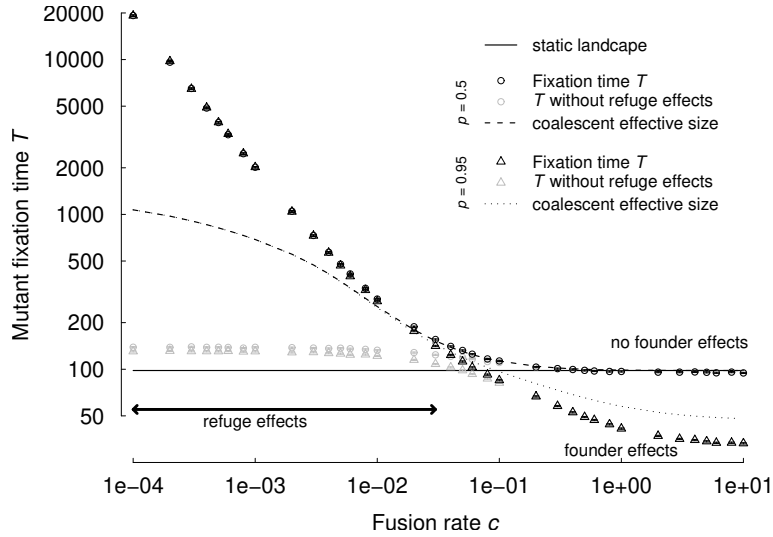


Figure 12: Model 2a, fixation time  $T$  of an advantageous mutant allele vs fusion rate  $c$ . “Static landscape” (solid line) refers to an undivided population where allele frequencies fluctuate via a Moran process (Eq. (2)). The caption “ $T$  without refuge effects” (gray symbols) refers to the corresponding simulations “Fixation time  $T$ ” (black symbols) to which the time spent in refuges has been subtracted. The difference between “Fixation time  $T$ ” and “ $T$  without refuge effects” thus quantifies refuge effects. With symmetrical fragmentations ( $p = 0.5$ , circles), only refuge effects affect  $T$ . The higher the fusion rate  $c$ , the weaker the refuge effects. For a high fusion rate ( $c > 0.05$ ), refuge effects are negligible:  $T$  equals the fixation time without landscape dynamics. For a small fusion rate ( $c < 0.05$ ), refuge effects are very strong,  $T$  can be more than 10 times longer than without landscape dynamics. With asymmetrical fragmentations ( $p = 0.95$ , triangles) both refuge and founder effects affect  $T$ . For a high fusion rate ( $c > 0.05$ ) refuge effects are negligible, founder effects determine  $T$ ;  $T$  decreases with  $c$  (more frequent founder effects generating additional drift). For a small fusion rate ( $c < 0.005$ ) founder effects are negligible, refuge effects determine  $T$  which equals  $T$  in the case of symmetrical fragmentations. Dashed and dotted lines corresponds to the approximation of  $T$  we obtained using the coalescent effective size we derived (Eq. (11)). This approximation is accurate under weak founder effects or weak refuge effects. For high  $c$  and high  $p$ , this approximation tends to underestimate founder effects; for small  $c$ , it predicts refuge effects, but underestimate them. *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.01$ ;  $f = 0.5$ .

$N_e^c$  is slightly less than the harmonic mean size of the population (*i.e.* the size of the reference population in a static landscape):

$$N_e^c \approx \frac{f + c}{c/N + f/(2N) + fc/(4N)}.$$

The term  $fc/(4N)$  corresponds to the effect of population size variations that occurs in Model 2b even when fragmentations are symmetrical (Figure 3). It explains the discrepancy between the expected time to fixation in a static landscape and these for Model 2b with symmetrical fragmentations (Figure 13).

Still assuming weak refuge effects, when fragmentations are not symmetrical, compared to a static landscape, we obtain an effective size decreased by a term  $fc\gamma/4$  for Models 2a and 2b:

$$N_e^c \approx \frac{f + c}{c/N + \alpha f/(2N) + fc\gamma/4}.$$

In contrast, when refuges can take place ( $c \ll 1/N$ ) and last a long time compared to duration of the state 1 of the landscape ( $f/c \gg 1$ ),  $N_e^c$  is strongly increased compared to

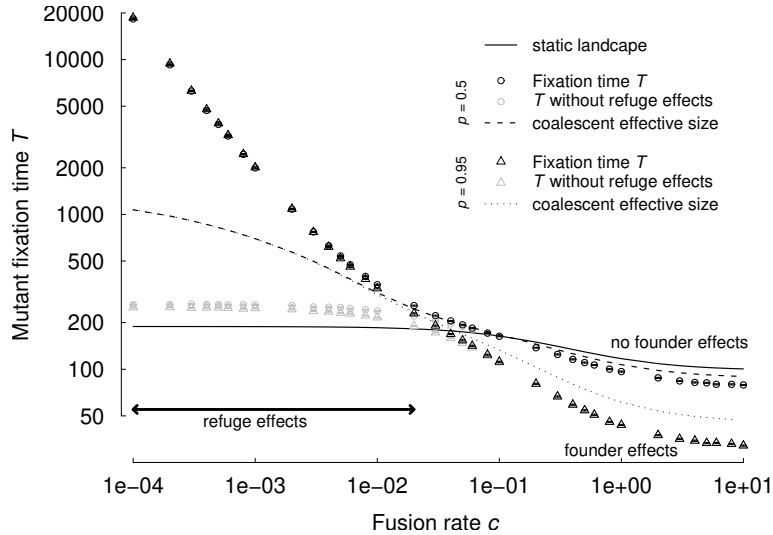


Figure 13: Model 2b, fixation time  $T$  of an advantageous mutant allele vs fusion rate  $c$ . Results of this figure (Model 2b) are analogous to these of Figure 12 (Model 2a). The only difference is that the fixation time  $T$  decreases with the fusion rate  $c$  also due to the fact that the mean overall population size decreases with  $c$ . *Parameter values:*  $N = 100$ ;  $x_0 = 0.01$ ;  $s = 0.01$ ;  $f = 0.5$ .

a static landscape and is independent of the asymmetry parameter  $p$ :

$$N_e^c \approx \frac{2}{\alpha} N \left[ \frac{1 + \alpha/(Nc)}{1 + 4/(fN)} \right].$$

$\frac{2}{\alpha} N$  is the size of the population when the landscape is in state 2 (which is the predominating state when refuges last a long time) and the term multiplying it is higher than 1 under the assumption we have made here. This term corresponds to the intensity of refuge effects.

It is interesting to note that the coalescent effective size cannot satisfactorily describe the fixation probability of an advantageous mutant in Models 2a and 2b, but that it gives a faithful prediction of (at least) the variations of the time to fixation with the parameters describing the spatial structure of the population.

### 3.3 Fixation probability and time to fixation of neutral and deleterious alleles

In the neutral case, as expected, mutant fixation probability is not altered by landscape dynamics of our three models (data not shown). After some time, all individuals of the population have indeed the same ancestor from the initial population. Given the spatial structure we assume in our models, all individuals of the initial population - neutral mutants and residents - are interchangeable. In particular, in Model 2 when the population is subdivided into two demes, the two demes are interchangeable. As a result, all individuals of the initial population have the same probability to be the future common ancestor of the population. Thus, the fixation probability of neutral mutants equals their initial frequency. Moreover, additional genetic drift due to founder effects speeds up the fixation of a mutant allele, even neutral, and refuges defined in Model 2 can be created in the neutral case also, delaying fixation. Consequently, the time to fixation of a neutral allele behaves qualitatively like an advantageous allele (data not shown).

We have built our three models such that advantageous and deleterious cases should

be symmetrical because only the relative difference  $s$  in birth rates between mutants and residents matters. A Moran process with an advantageous mutant birth rate  $1 + s$  and a resident birth rate 1 is identical to a Moran process with a deleterious mutant birth rate 1 and a resident birth rate  $1 + s$ : reversing mutant and resident roles allows to look at the deleterious case instead of the advantageous case. However an asymmetry due to the mutant initial frequency remains: an advantageous mutant with initial frequency  $x_0$  is symmetrical to a deleterious mutant with initial frequency  $1 - x_0$ . However, we are of course interested in results with an initial deleterious mutant frequency  $x_0$ . We thus checked with simulations that this asymmetry does not significantly alter the expected results (data not shown): for our three models in the parameters values where the advantageous mutant fixation probability is lower than in a static landscape, the deleterious mutant fixation probability is higher. Moreover (and as expected), times to fixation of deleterious alleles are qualitatively similar to those of advantageous and neutral cases (data not shown), for all landscape dynamics we examined.

### 3.4 Comparison between Moran and Wright-Fisher processes

We obtained all results presented above using a Moran process for population dynamics. We did the same analyses for the more classical Wright-Fisher model (Wright 1931). For the landscape dynamics we studied, results are qualitatively identical, for both fixation probability and time to fixation (data not shown). Fixation probabilities of advantageous (respectively deleterious) alleles are a higher (respectively smaller) since for a same census population size, the effective size of a Wright-Fisher population is twice higher than that of a Moran population. For the same reason, fixation times are longer.

## 4 Discussion

We have proposed three models for understanding how dynamic landscapes influence the fixation probability and the time to fixation of a mutant allele. We have shown that compared to a static landscape (undivided population of constant size), in the case of repeated bottlenecks (Model 1), a succession of founder events decreases the fixation probability of an advantageous mutation, but accelerates its fixation (conditional on fixation). These effects are stronger when the landscape dynamics are faster. Also in the case of the repeated alternation of fragmentation and fusion of demes (Models 2a and 2b) do founder events decrease the fixation probability of an advantageous mutation and accelerate fixation. However, the coexistence of two temporarily disconnected demes (state 2) generates a “refuge effect” which can strongly delay fixation. If population fusions are rare, refuge effects are the principal factor determining fixation times which are longer than in a static landscape. In contrast, if fusions are frequent, founder effects are the principal factor and fixation times are then shorter than in a static landscape. Note that founder effects are only observed in the case of asymmetrical fragmentations.

We have derived for Model 1 a variance (Eq. (5)), a coalescent (Eq. (6)) and a harmonic mean (Eq. (7)) effective size. The variance and coalescent effective sizes lead to the same approximation. None of the effective sizes can give an accurate prediction for both the fixation probability and the time to fixation of an advantageous mutant. We have also shown that none of them does better than the others for all bottleneck intensities and frequencies (Figures 4 and 5). The harmonic mean effective size nevertheless appears to be the most robust. For Models 2a and 2b, we have derived a coalescent effective size (Eq. (11)) which fairly predicts time to fixation under weak founder effects or weak refuge effects (Figures 12 and 13). Depending on the characteristics of the landscape dynamics,

this effective size is decreased or increased compared to that of a static landscape. Because refuge effects strongly increase time to fixation but do not affect fixation probability, there cannot exist one single effective population size which can describe both fixation probability and time to fixation for spatially structured populations subject to landscape dynamics generating refuges.

Otto and Whitlock (1997) studied the effect of a single bottleneck on the fixation probability of a beneficial mutation. Later, Wahl and Gerrish (2001) and Heffernan and Wahl (2002) analyzed bottlenecks occurring cyclically with an extreme regularity. In spite of the differences with our Model 1, our results are qualitatively similar to theirs. The occurrence of bottlenecks at stochastic times appears to weakly alter fixation compared to regular bottlenecks. However, an additional feature of our work is that the variance effective size we derived (Eq. (5)) takes into account the variance of allele frequencies between bottlenecks (*i.e.* during Moran processes). Our approximation may thus be used also for weak reduction of population size, when an advantageous mutant is more likely to be lost between bottlenecks than because of a bottleneck.

Extinction-recolonization models (Slatkin 1977; Barton 1993; Whitlock 2003) can be compared to Models 2a and 2b of repetitive fragmentations and fusions. They are island models with a large number of demes. In each generation, extinction of one colony occurs with probability  $e$ , and a number  $k$  of individuals recolonize instantaneously the extinct colony (allowing a founder event if  $k$  is small). Both extinction-recolonization models and our Models 2a and 2b allow to model fragmenting populations where the number of occupied patches varies and the overall population size is nevertheless limited. However, they assume different mechanisms that result in the fluctuation of the number of patches: extinctions for the first case, secondary contacts (with, in Model 2b, competition following fusion, modelled by the pure death process) for the second case. Therefore, geographical and biological processes that can be modelled with these models are quite different. In extinction-recolonization models, excluding the case of extinction of small demes due to stochastic demographic fluctuations, extinctions implicitly assume dramatic and sudden events (such as violent climatic events), with immediate recolonization. All alleles in a deme which goes extinct are necessarily lost. In contrast, in our models, population size reductions are less dramatic events: we assume a complete fusion of demes, so that all alleles compete with each other and can survive after the population size reduction. It thus corresponds rather to the movement or expansion of a subpopulation to the place where another is already established, thereby increasing competition. In spite of these differences, our conclusions on fixation probability are qualitatively similar to the results of extinction-recolonization models (reduction of the fixation probability compared to an undivided population). In contrast, fixation times can be drastically higher in Models 2a and 2b (Figures 12 and 13) than in extinction-recolonization models for which the diffusion approximation in a static landscape (Eq. (2)) gives good predictions (Whitlock 2003). The reason is that the dynamics we assume allow the creation of refuges. Consequently, alleles can be retained for a long time without risking to disappear, and then be reintroduced in the total population. Hence, geographical and biological processes that can be modelled with our models should probably allow to keep a higher global diversity than those that are modelled with extinction-recolonization models.

Our three models of landscape dynamics assume (sub-)population size variations. Our results of Model 2a and 2b about the mean mutant fixation probability  $U$  (Figures 8, 9 and 10) show that the effect of these variations can be considered as analogue to repetitive founder events. However, when considering the state of the landscape when a mutant appears, a more accurate analysis is possible. Otto and Whitlock (1997) analyzed fixation probability in populations of changing size. They showed that the fate of a



mutation strongly depends on the direction of variation of population size (increase or decrease) when the mutation occurs. Our results on state-dependent fixation probability for Model 2a and 2b with successive population increases and decreases confirm that the closer in time to a population size variation a mutant arise, the more its fixation probability is altered.

Our question was to understand how dynamic landscapes alter mutant fixation. One main modification is that founder effects and refuge effects are repetitive and cumulative. These effects can exist in static landscapes, but can occur only once. Landscape dynamics thus strengthen alterations of fixation probabilities and times to fixation. This is particularly true for times to fixation which can be increased by a factor of more than 10 (Figures 12 and 13). This stresses the importance to consider the isolation of some population and their fusion after a possibly long time. Depending on the characteristics of the dynamics, fixation can be disfavored or unaffected, delayed or accelerated. Therefore, one needs to describe the whole dynamics of the landscape, and to specify characteristic time scales. Fast dynamics are appropriate to model ecological processes such as dispersal and recolonization events (establishment of new colonies and their later fusion (DeHeer and Kamble 2008; Vasquez and Silverman 2008) because of their expansion or because one habitat becomes unsuitable), or to model geographical processes such as changes in the fragmentation of habitat due to human action (Davies et al. 2006). In that case, beneficial mutations have “one small chance of doing very well”: fixation is unlikely (frequent founder effects) but very fast if it occurs (limited refuge effects). In contrast, slow dynamics are appropriate to geographical events such as the separation of populations due to glacial events followed by postglacial contacts (Young et al. 2002) or such as repetitive fragmentation and fusion of islands (or lakes) due to water level variations caused by climatic events (Owen et al. 1990; Delvaux 1999; Galis and Metz 1998; Stiassny and Meyer 1999; Cook 2008). In that case, fixation of a beneficial allele would be more likely (rare founder effects) but very slow when it occurs (strong refuge effects). On top of specifying characteristic times scales, knowing how the overall population size varies is essential. In Model 2a the overall carrying capacity is constant whereas it fluctuates in Model 2b. Results of these two models are very close when compared to equivalent static landscapes (*i.e.* the effect of dynamic landscapes is the same for both models), but they can be significantly different when compared to each other.

Our models are based on some strong assumptions. First, we assume that the pure birth (or death) process used after a bottleneck (Model 1) or a fragmentation or a fusion (Models 2a and 2b) is instantaneous. This assumption is justified because this exponential growth phase is very short compared to the timescale of allele frequency change. Nevertheless, we checked that relaxing this assumption does not significantly change our results. We thus have explicitly computed the durations of the exponential growth and decline phases, which appeared indeed to be much shorter than other processes; fixation times were almost unchanged (data not shown).

Second, Models 2a and 2b assume that in state 2, the two demes do not exchange migrants. This may be a strong assumption, we thus relaxed it (data not shown). Migration has only a small effect on fixation probability: with a migration rate of 0.1 per individual, fixation probabilities are only 10% (at the most) higher than without migration (for the range of values we explored, Table 1). In contrast, high migration rates have a significant effect on time to fixation which approaches the time to fixation in a static landscape (Eq. (2)) when migration increases. When the landscape dynamics generate refuges, each migration event indeed breaks refuges, which shortens the time to fixation. Our results are nevertheless still valid in the case of weak migration: for a per capita migration rate less than 0.01, we still observe refuge effects, *i.e.* the time to fixation is

significantly higher than that in a static landscape. Also note that the effect of migration weakly depends on the fragmentation rate  $f$  and on the fusion rate  $c$ .

Last, Models 2a and 2b assumes that the population is divided in at most two demes. Fragmented populations generally consist in many subpopulations (Hanski and Gaggiotti 2004). In such landscapes, the number of founder events and of refuges would thus be higher; advantageous mutations would thus be less likely to be fixed and fixation, when it occurs, would be more delayed. For the extreme scenario of continually increasing fragmentation of the population, fixation could even never occur because refuges would always exist and each one may fix a different allele. However, highly fragmented populations often consist of small and weakly interconnected demes whose existence is usually ephemeral: subpopulations go extinct stochastically and habitats are recolonized. Extinctions may then allow fixation since it destroys refuges.

Our models need to be improved to take into account more complex dynamics involving for example more demes, extinction and recolonization of demes, or different local selective pressures which would fix different alleles in different demes. They also need to consider that new mutations may arise before fixation of older mutations. In particular new mutants can appear in refuges and invade them. Taking into account these additional features will allow understanding not only allele fixation probability and time to fixation but also fixation flux, genetic diversity or adaptation rate in populations involved in the geographical and biological processes that dynamic landscapes allow to model.

## Acknowledgments

The authors are grateful to four anonymous referees, whose comments significantly improved the scope and quality of our work. We thank Sébastien Ballesteros for helpful discussions. R. A. is supported by a PhD fellowship from the French Ministère de la Recherche et de la Technologie. Traveling was funded by the project Modèles Aléatoires de l'Evolution du Vivant (MAEV) of the French Agence Nationale pour la Recherche (ANR).

## References

- Barton, N. H., 1993. The probability of fixation of a favoured allele in a subdivided population. *Genet. Res.* 62, 149–158.
- Barton, N. H., Whitlock, M. C., 1997. The evolution of metapopulations. Academic Press, San Diego, Ch. Metapopulation biology: ecology, genetics and evolution, pp. 183–210.
- Batzli, O., 1992. Dynamics of small mammal populations: a review. Elsevier Applied Science, London.
- Bohme, M. U., Schneeweiss, N., Fritz, U., Schlegel, M., Berendonk, T. U., 2007. Small edge populations at risk: genetic diversity of the green lizard (*Lacerta viridis viridis*) in Germany and implications for conservation management. *Conserv. Genet.* 8, 555–563.
- Charlesworth, B., Charlesworth, D., Barton, N. H., 2003. The effects of genetic and geographic structure on neutral variation. *Annual Review of Ecology, Evolution, and Systematics* 34, 99–125.
- Colas, B., Thomas, C. D., Hanski, I., 2002. Evolutionary conservation biology. Cambridge University Press, Ch. Evolutionary responses to landscape fragmentation, pp. 284–300.

- Cook, L. M., 2008. Species richness in Madeiran land snails, and its causes. *Journal of Biogeography* 35, 647–653.
- Davies, R. G., Orme, C. D. L., Olson, V., Thomas, G. H., Ross, S. G., Ding, T.-S., Rasmussen, P. C., Stattersfield, A. J., Bennett, P. M., Blackburn, T. M., Owens, I. P. F., Gaston, K. J., 2006. Human impacts and the global distribution of extinction risk. *Proc. R. Soc. B.* 273, 2127–2133.
- DeHeer, C. J., Kamble, S. T., 2008. Colony genetic organization, fusion and inbreeding in *Reticulitermes flavipes* from the midwestern US. *Sociobiology* 51, 307–325.
- Delvaux, D., 1999. Age of Lake Malawi (Nyasa) and water level fluctuations. Tech. rep., Musée royal de l’Afrique Centrale (Tervuren), Département de Géologie et Minéralogie, rapport annuel, 1993-1999, pp.99-108.
- Ewens, W. J., 1967. The probability of survival of a new mutant in a fluctuating environment. *Heredity* 22, 438–443.
- Ewens, W. J., 2004. *Mathematical Population Genetics*. Springer, New York.
- Fisher, R. A., 1922. On the dominance ratio. *Proc. R. Soc. Edinb.* 42, 321–341.
- Galis, F., Metz, J. A. J., 1998. Why are so many cichlid species. *Trends in Ecology and Evolution* 13, 1–2.
- Gao, L. Z., Zhang, C. H., 2005. Comparisons of microsatellite variability and population genetic structure of two endangered wild rice species, *Oryza rufipogon* and *O. officinalis*, and their conservation implications. *Biodiv. Conserv.* 14, 1663–1679.
- Gillespie, J., 2004. *Population genetics - a concise guide*. Johns Hopkins University Press.
- Haldane, J. B. S., 1927. A mathematical theory of natural and artificial selection. V. Selection and mutation. *Proc. Camb. Phil. Soc.* 23, 838–844.
- Hanski, I., Gaggiotti, O. E., 2004. *Ecology, Genetics, and Evolution of Metapopulations*. Elsevier Academic Press.
- Heffernan, J., Wahl, L., 2002. The effects of genetic drift in experimental evolution. *Theoretical Population Biology* 62, 349–356.
- Heinemann, J. A., 1999. How antibiotics cause antibiotic resistance. *Drug Discov. Today* 4, 72–79.
- Iizuka, M., 2001. The effective size of fluctuating populations. *Theoretical Population Biology* 59, 281–286.
- Iizuka, M., Tachida, H., Matsuda, H., 2002. A neutral model with fluctuating population size and its effective size. *Genetics* 161, 381–388.
- Jesus, F., Wilkins, J., V.Solferini, Wakeley, J., 2006. Expected coalescence times and segregating sites in a model of glacial cycles. *Genet. Mol. Res.* 5, 466–474.
- Karlin, S., Taylor, H., 1981. *A second course in stochastic processes*. Academic Press.
- Keymer, J., Marquet, P., Velasco-Hernandez, J., Levin, S., 2000. Extinction thresholds and metapopulation persistence in dynamic landscapes. *The American Naturalist* 156, 478–494.

- Kimura, M., 1962. On the probability of fixation of mutant genes in a population. *Genetics* 47, 713–719.
- Kimura, M., 1970. *Mathematical topics in population genetics*. New York, NY: Springer, Ch. Stochastic processes in population genetics, with special reference to distribution of gene frequencies and probability of gene fixation, pp. 178–209.
- Kimura, M., Ohta, T., 1969. The average number of generations until fixation of a mutant gene in a finite population. *Genetics* 61, 763–771.
- Kimura, M., Ohta, T., 1974. Probability of gene fixation in an expanding finite population. *Proc. Natl. Acad. Sci. USA* 71, 3377–3379.
- Lambert, A., 2006. Probability of fixation under weak selection: a branching process unifying approach. *Theoretical Population Biology* 69, 419–441.
- Malhi, Y., Roberts, J. T., Betts, R. A., Killeen, T. J., Li, W., Nobre, C. A., 2008. Climate change, deforestation, and the fate of the amazon. *Science* 319, 169–172.
- McLean, A. R., 1995. Vaccination, evolution and changes in the efficacy of vaccines: a theoretical framework. *Proc. R. Soc. B.* 261, 389–393.
- Moran, P. A. P., 1962. *The Statistical Processes of Evolutionary Theory*. Clarendon, Oxford, UK.
- Otto, S. P., Whitlock, M. C., 1997. The probability of fixation in populations of changing size. *Genetics* 164, 723–733.
- Owen, R., Crossley, R., Johnson, T. C., Tweddle, D., Kornfield, I., Davison, S., Eccles, D. H., Engstrom, D. E., 1990. Major low levels of Lake Malawi and their implication for speciation rates in cichlid fishes. *Proc. R. Soc. Lond. B Biol. Sci.* 240, 519–553.
- Patwa, Z., Wahl, L., 2008. The fixation probability of beneficial mutations. *J R Soc Interface* 5, 1279–1289.
- R Development Core Team, 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. URL <http://www.R-project.org>
- Roze, D., Rousset, F., 2003. Selection and drift in subdivided populations: a straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. *Genetics* 165, 2153–2166.
- Sano, A., Shimizu, A., Iizuka, M., 2004. Coalescent process with fluctuating population size and its effective size. *Theoretical Population Biology* 65, 39–48.
- Slatkin, M., 1977. Gene flow and genetic drift in species subject to frequent local extinctions. *Theoretical Population Biology* 12, 253–262.
- Stiassny, M. L. J., Meyer, A., 1999. Cichlids of the rift lakes. *Sci Am* 280, 44–49.
- The GSL Team, 2007. *GNU Scientific Library - Reference Manual*. URL <http://www.gnu.org/software/gsl/>

- Theodorou, K., Couvet, D., 2006. Genetic load in subdivided populations: interactions between the migration rate, the size and the number of subpopulations. *Heredity* 96, 69–78.
- Turchin, P., 2003. *Complex population dynamics - A theoretical/empirical synthesis*. Princeton University Press.
- Vasquez, G. M., Silverman, J., 2008. Intraspecific aggression and colony fusion in the Argentine ant. *Animal Behaviour* 75, 583–593.
- Vuilleumier, S., Yearsley, J. M., Perrin, N., 2008. The fixation of locally beneficial alleles in a metapopulation. *Genetics* 178, 467–475.
- Wahl, L. M., Gerrish, P. J., 2001. The probability that beneficial mutations are lost in populations with periodic bottlenecks. *Evolution* 55, 2606–2610.
- Whitlock, M. C., 2002. Selection, load and inbreeding depression in a large metapopulation. *Genetics* 160, 1191–1202.
- Whitlock, M. C., 2003. Fixation probability and time in subdivided populations. *Genetics* 164, 767–779.
- Whitlock, M. C., 2004. *Ecology, genetics and evolution of metapopulations*. Elsevier Academic Press, Ch. Selection and drift in metapopulations, pp. 153–174.
- Whitlock, M. C., Barton, N. H., 1997. The effective size of a subdivided population. *Genetics* 146, 427–441.
- Wilcox, B. A., Murphy, D. D., 1985. Conservation strategy: the effects of fragmentation on extinction. *The American Naturalist* 125, 879–887.
- Wright, S., 1931. Evolution in Mendelian populations. *Genetics* 16, 97–159.
- Wright, S., 1938. Size of population and breeding structure in relation to evolution. *Science* 87, 430–431.
- Young, A., Torres, C., Mack, J., Cunningham, C., 2002. Morphological and genetic evidence for vicariance and refugium in Atlantic and Gulf of Mexico populations of the hermit crab *Pagurus longicarpus*. *Marine Biology* 140, 1059–1066.