

The role of migration in mutant dynamics in fragmented populations

Abstract

Mutant dynamics in fragmented populations have been studied extensively in evolutionary biology. Yet, open questions remain, both experimentally and theoretically. Some of the fundamental properties predicted by models still need to be addressed experimentally. We contribute to this by using a combination of experiments and theory to investigate the role of migration in mutant distribution. In the case of neutral mutants, while the mean frequency of mutants is not influenced by migration, the probability distribution is. To address this empirically, we performed *in vitro* experiments, where mixtures of GFP-labeled (“mutant”) and non-labeled (“wild-type”) murine cells were grown in wells (demes), and migration was mimicked via cell transfer from well to well. In the presence of migration we observed a change in the skewedness of the distribution of the mutant frequencies in the wells, consistent with previous and our own model predictions. In the presence of de novo mutant production, we used modeling to investigate the level at which disadvantageous mutants are predicted to exist, which has implications for the adaptive potential of the population in case of an environmental change. In panmictic populations, disadvantageous mutants can persist around a steady state, determined by the rate of mutant production and the selective disadvantage (selection-mutation balance). In a fragmented system that consists of demes connected by migration, a steady state persistence of disadvantageous mutants is also observed, which, however, is fundamentally different from the mutation-selection balance and characterized by higher mutant levels. The increase in mutant frequencies above the selection-mutation balance can be maintained in small ($N < N_c$) demes as long as the migration rate is sufficiently small. The migration rate above which the mutants approach the selection-mutation balance decays exponentially with N/N_c . The observed increase in the mutant numbers is not explained by the change in the effective population size. Implications for evolutionary processes in diseases are discussed, where the pre-existence of disadvantageous drug-resistant mutant cells or pathogens drives the response of the disease to treatments.

1 Introduction

Understanding the principles of mutant dynamics has been a major focus in evolutionary biology. Generation and spread of mutants is central to adaptation, where beneficial mutants tend to fix while deleterious mutants are gradually removed from the population. These processes depend on the environment and are guided by forces of selection, the rate of mutation, genetic drift, and the population structure.

Population structure is an important determinant of the evolutionary trajectories [25, 42]. Evolutionary dynamics in fragmented populations are of interest for questions connected to ecology and ecological conservation [6, 46, 41, 39, 14, 40, 22]. Human interference as well as natural factors may fragment habitats and isolate subpopulations of a species from the rest, thus influencing genetic variability and species survival. Population genetics in structured populations has been applied to studies of island biogeography, dynamics of species living in patchy environments, extinction and recolonization, see e.g. [47, 11]. Another area where population fragmentation is of great importance is host-associated microbiomes. Spatial structures are generated by host anatomy and physiology; examples include gastrointestinal crypts [8] and skin pores [5]. Understanding the microbial evolution in structured populations is critical for modeling community diversity and stability, species coexistence, and predicting the response of microbiomes to treatments. Further applications to biomedical problems are discussed below.

Mathematical models have been an important component of research into mutant dynamics. Different types of evolutionary models have been explored; most relevant for the current paper are the Moran process [34, 35] and the Wright-Fisher model [57] that assume constant population sizes.

Various evolutionary measures have been considered, including the average frequency of mutants at a given time or population size, the fixation probability of mutants, and the average time to fixation for mutants of varying relative fitness [19, 38, 27, 15, 54, 31].

Evolution in fragmented populations is described by models that are sometimes referred to as structured or subdivided population models, as well as patch or deme models [35, 26, 25, 12, 42, 11]. These models describe a group of distinct, spatially separated populations of the same type. Some amount of interaction between the separate groups occurs via migration of individuals from one group to another, and the dynamics within a single group of individuals is generally assumed to be non-spatial [35, 26, 49, 13]. Migration of individuals can either occur to the nearest neighboring regions (spatially restricted), or individuals can migrate to any region in the system. A higher rate of migration decreases population fragmentation because it results in each region’s dynamics becoming dependent on a larger portion of the overall population and thus in better mixing of individuals [57, 35, 12, 49]. The term “metapopulation models” is often reserved to describing the dynamics with local extinctions and re-colonization, see e.g. [25, 43, 47, 42, 11, 3].

Mathematical patch models with different assumptions on structure and migration between groups have been studied in the context of evolutionary dynamics. A number of important results about the effect of fragmentation and structure on mutant dynamics have been established. Commonly, it is found that the fixation probability of a mutant is largely independent of migration (depending on the explicit model assumptions) [32, 33, 49], but that other quantities such as the time to fixation and effective population size can vary based on model structure [44, 54]. The distribution of mutant numbers in individual demes has been studied in different contexts, starting with the seminal paper by Wright [57], which gave rise to the standing balance theory of evolution. Further developments include both discrete (Wright-Fisher) and overlapping (Moran) models of population dynamics and different assumptions on the migration process, see e.g. [35, 12, 49, 13]. It was found that generally, migration among demes transforms the probability distribution of mutant frequency in a deme from bimodal to unimodal. Another set of results comes from the diffusion approximation that describes selection and drift of mutants in subdivided populations. In [4, 48], a Wright-Fisher process in a subdivided population with inter-deme migrations is considered, while in [49], the local dynamics are described by a Moran process. Analytical expressions for the effective population size are derived. It is shown that although the form of the diffusion approximation is equivalent between a structured and panmictic population, fragmentation can significantly increase the effective population size and the variance of allele frequencies.

The interplay between patch dynamics and traits or alleles has also been previously studied in multiple experimental contexts [23, 18, 9, 2]. For instance, Kerr et al identified path dependent migration effects in the eco-evolutionary dynamics of E.coli-T4 phage co-cultures [18]. Excitingly, recent work on range expansions in asexually reproducing microbes has shown that an excess of spontaneous mutations (relative to Luria-Delbruck expectations) are generated during spatial range expansions by allele surfing [9].

While several aspects of mutant dynamics in fragmented populations have been mathematically elucidated, there still remain open theoretical questions and experimental gaps, some of which we address in this paper. (i) From an experimental point of view, to the best of our knowledge, direct tests of model predictions regarding mutant distributions in fragmented asexual populations in the presence and absence of migration are lacking. Here, we provide an experimental test of fundamental model predictions about mutant distributions for neutral mutants, using a system where GFP-labeled and unlabeled murine cell lines (“mutant” and “wild-type”) are co-cultured in 96-well plates, and migration is mimicked by swapping cells between demes (wells) with a pipette. This system does not contain de novo mutant production, and the experimental results are inter-

preted with a corresponding Moran process model, which reproduces previous theoretical insights and is found to be consistent with the data. (ii) The model is then used to extend this analysis assuming de novo mutant production and different mutant fitness. In particular, we study the perceived “selection-mutation balance” in populations. This refers to the persistence of disadvantageous mutants around a steady state in a population of wild-types at equilibrium, the level of which can be calculated and is determined by the mutation rate and the degree of the selective disadvantage. We show that for a fragmented habitat, disadvantageous mutants can also persist at a steady state level similar to the selection-mutation balance in a panmictic population, but that there is a fundamentally different scenario in which the expected number of mutants at the steady state can be significantly higher. Our results demonstrate that the previously obtained change in the effective population size of fragmented populations is not enough to explain the change in the level at which the disadvantageous mutants persist. This has implications for understanding the adaptive potential of a population in response to environmental change, with broad applicability ranging from ecological systems to biomedical problems, such as the emergence of drug resistant mutants.

2 Methods

2.1 Experiments with neutral mutants

To address the existing gap in the literature concerning studies of neutral mutant dynamics in fragmented populations in the presence and absence of migration, we performed experiments that represent an *in vitro* comparison to the mathematical model presented in the following section. To create a system representing neutral migration we mixed GFP labeled mammalian cells with unlabeled cells. This suspension of mammalian cells could be propagated in the wells of a 96 well plate. By systematically transferring small volumes of the cell suspension, we experimentally simulated migration between wells. Cells were continually maintained at confluent cell population densities to mimic the Moran process. The proportion of mutants in the wells both with and without migration of cells between wells was assessed at the end of the experiment. Further details on the *in vitro* experiments are presented below.

Experimental details. 96 demes (wells) were filled with cells. Cell types were wild type and mutant, which were neutral with respect to one another. To create this model we transduced murine Ba/F3 cells (DSMZ; ACC-300) with green fluorescent protein (GFP). GFP+ cells were mixed with wild-type Ba/F3 cells into populations of approximately 0.1 and 1 percent GFP+. Ba/F3 populations were maintained in RPMI 1640 Medium (Sigma Aldrich), supplemented with 10 percent FBS (Fisher), 10 ng/mL IL-3 (PeproTech), 100 U/mL penicillin and 100 μ g/mL streptomycin (Life Technologies). The control condition was no migration and the experimental condition was migration between demes, which was performed by swapping cells between demes using a pipette. Four total experiments were performed, with both conditions starting at approximately 0.1% initial mutants and also at 1% initial mutants in all demes. Wells in each plate were diluted twice daily with fresh medium on a 16/8 hour time interval to maintain competition at high confluence, with one plate from each mixed population subject to migration events at these intervals. All wells were maintained at 200 μ L — dilutions with fresh medium were performed to maintain cell confluency at stationary phase and replenish nutrients. This was done to approximate a Moran process and allow for continuous culture, and it was a significant challenge. Throughout the experiment, maximum population densities were adjusted based on viability, to maintain population density and competition. Migration events were performed immediately prior to viability dilutions on one plate of each GFP+ population through the duration of the experiment. Each well in a row was thoroughly

Notation	Description
K	number of demes
N	constant number of individuals in each deme
NK	constant total number of individuals in overall population
r	fitness of the mutant individuals
u	rate of forward mutation wild type $\rightarrow$ mutant
u_b	rate of back-mutation wild type $\leftarrow$ mutant
p_{migr}	migration probability
n_{ind}	number of individuals exchanged during a swapping event
n_{swaps}	number of swaps that occur during a migration event
$j_{sel-mut}$	selection mutation balance in each deme

Table 1: Description of model parameters.

homogenized, and 5 uL of each well’s 200 uL volume was transferred to the vertically adjacent well. Receiving wells were then homogenized and 5 uL was transferred to the next adjacent well; this process was repeated until all eight rows had received volume from its adjacent upstream row and transferred to its downstream adjacent row.

2.2 Mathematical modeling

To study the role of population fragmentation and migration in evolution, we will consider a population of asexually reproducing (haploid) individuals of two types, which we refer to as “wild types” and “mutants”, see Figure 1. The total population of NK individuals is split into K demes of N individuals each, as in for instance [13]. We assume a finite island model, where all demes are equidistant. Competition is implemented by assuming that (neutral or deleterious) mutants may have relative fitness (denoted by $r = 1 - s \leq 1$) that is not necessarily equal to the fitness of the wild types (assumed to be 1). De novo mutations are included through forward mutation (with probability u per division of a wild type cell) and back-mutation (with probability u_b per division of a mutant cell, further details included in Section 2.2.1). Migration is modeled in the following way. A single migration event is attempted with probability $0 \leq p_{migr} \leq 1$ and performed by randomly selecting two demes, then randomly selecting n_{ind} individuals from each and swapping them with each other; n_{swaps} migration updates are completed each time.

The dynamics are set up in the following way. First, a migration update is performed, where individuals have a chance to swap demes. This is followed by a birth-death update in all K demes. Further details on birth-death updates are given in Section 2.2.1, and on migration in Section 2.2.2; Table 1 lists all the model parameters. We run simulations until long-term dynamics have been established and (quasi)-stationary states (mutant extinction/fixation or a stable average number of mutants within individual demes and the overall system) have been reached.

2.2.1 The Moran process

Within each deme, we model the stochastic birth-death dynamics by using the well-known Moran process (see e.g. [34, 35]). Therefore, the population size of each deme as well as the total population size remains constant. The process of de novo (forward and/or backward) mutation may be included in our framework. De novo mutation occurs during reproduction, so for instance if a wild type cell is chosen for reproduction, there is a $1 - u$ chance of faithful reproduction and a u chance to create a mutant. If a mutant cell divides, it creates a wild-type offspring with probability u_b (and a mutant offspring with probability $1 - u_b$).

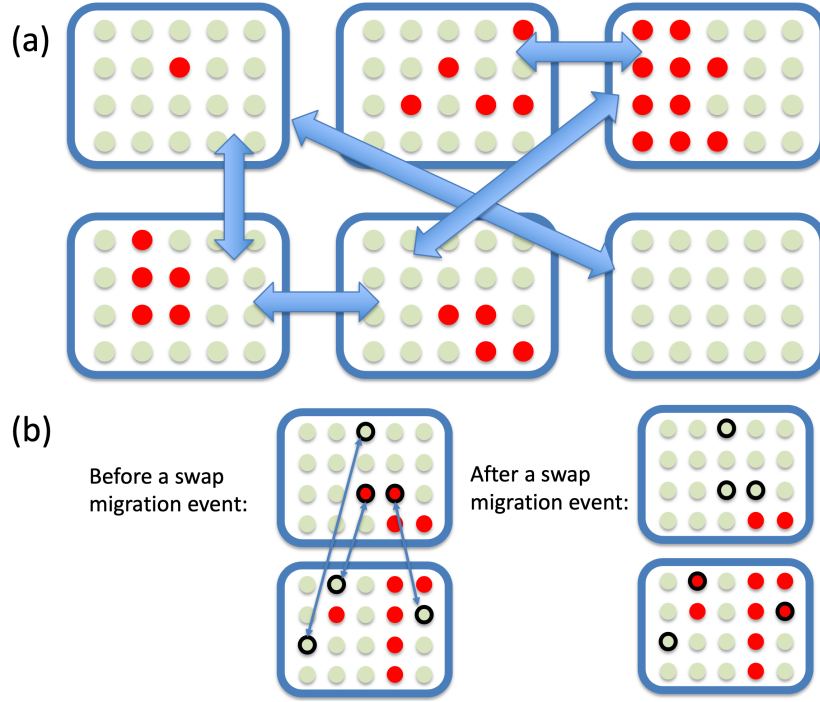


Figure 1: A schematic illustrating the mathematical model. (a) General model structure: each rectangle represents a deme, and green (red) circles represent wild-type (mutant) individuals. Two-sided arrows represent random swap-migration events within randomly chosen pairs of demes. (b) Details of a swap migration event: groups of $n_{ind} = 3$ cells are randomly selected within two demes, and exchanged. As a result of this particular event, the number of mutants in the top deme decreased, and the number of mutants in the bottom deme increased by 2.

In the absence of de novo forward mutation ($u = 0$), the model has only one absorbing state, given by mutant extinction in all demes. Similarly, in the absence of de novo back mutation ($u_b = 0$), mutant fixation in all demes is the only absorbing state. With the inclusion of both forward and back mutation, there are no absorbing states [34]. In the absence of migration, the stationary probability distribution in an individual deme can be calculated. For example, in the regime where mutant fixation happens on a much faster time-scale than mutant production, a simple expression for the stationary probability can be derived. Denoting by y_i the probability to have i mutants in the deme, we have

$$y_0 = \frac{u_b}{r^{N-1}u + u_b}, \quad y_N = 1 - y_0, \quad (1)$$

with the probability of the other states being of the order of the mutation rate, which is described in further detail in Section 2.2.3. This is similar to previous approximations of the stationary distribution for the Moran process with de novo mutation and selection and under various conditions, see for instance [35, 7, 51, 16, 49]. Another useful quantity is the selection-mutation balance, which is given (for an individual deme, in the limit of small mutation rates) by

$$j_{sel-mut} = \frac{Nu}{1-r}, \quad (2)$$

this quantity represents the number of (negatively selected, $r = 1 - s < 1$) mutants, in the Moran process with de novo mutations, which corresponds to an equal probability to increase and decrease this number in a single birth-death update. This quantity is also identical to the classical expression of the expected number of deleterious mutations at mutation-selection balance. Section 2.2.3 (see also [7, 51, 35]) provides details of the calculations for expressions (1) and (2), as well as higher

order approximations.

203

204 The process of Moran birth-death updates within an individual deme is set up as follows.
 205 Suppose the number of mutants is given by j (and thus the number of wild type individuals is
 206 given by $N - j$). For each time step a cell is chosen randomly to die and a cell is chosen based
 207 on fitness to reproduce.¹ Therefore, the chance that a mutant is chosen to reproduce is $\frac{rj}{rj+N-j}$,
 208 and the chance that a wild type is chosen to reproduce is $\frac{N-j}{rj+N-j}$. This is a Markov process, where
 209 the states of the model (the number of mutants in the deme) are integers $j \in \{0, 1, \dots, N\}$. This
 210 model has a tridiagonal transition matrix. Let us define

$$P_j^\uparrow = \frac{(N-j)}{N} \frac{rj(1-u_b) + (N-j)u}{rj+N-j}, \quad P_j^\downarrow = \frac{j}{N} \frac{rju_b + (N-j)(1-u)}{rj+N-j} \quad (3)$$

to be the probabilities to increase and decrease the number of mutants starting from j mutants, in one step. Then we have

$$P_{jk}^{Moran} = \begin{cases} P_j^\uparrow, & k = j+1, \\ P_j^\downarrow, & k = j-1, \\ 1 - (P_j^\uparrow + P_j^\downarrow) & k = j, \\ 0 & \text{otherwise.} \end{cases} \quad 0 \leq j \leq N.$$

211

Let us denote the probability to have m mutants at time t in a deme as $\pi_m(t)$, with $0 \leq m \leq N$, $\sum_{m=0}^N \pi_m(t) = 1$. The row vector $\pi(t)$ contains this information for each discrete time-step, t . The initial condition is

$$\pi_m(0) = \begin{cases} 1, & m = m_0, \\ 0, & \text{otherwise.} \end{cases}$$

In the absence of migration, we have (in matrix form)

$$\pi(t+1) = \pi(t)P^{Moran}, \quad t \geq 0.$$

212

If both populations reproduce faithfully ($u = u_b = 0$), we have two absorbing states, $j = 0$ (mutant extinction) and $j = N$ (mutant fixation) [35]. Let ϕ_j denote the probability for mutants to reach fixation given that we start with j mutants. We have

$$\phi_j = \begin{cases} \frac{1-(1/r)^j}{1-(1/r)^N}, & r \neq 1, \\ j/N, & r = 1. \end{cases}$$

213

2.2.2 Modeling migration

215 We assume a migration update is attempted each step with some migration probability $0 \leq p_{migr} \leq$
 216 1. We assume that at each migration update, n_{ind} individuals are randomly picked from one deme
 217 and replaced with n_{ind} individuals randomly selected from the second deme. In some simulations,
 218 to increase the intensity of migration, we repeated this procedure n_{swaps} times for each migration
 219 update. We assume that the probability of migration applies to all swap events jointly, and thus
 220 that n_{swaps} swap events occur with probability p_{migr} , and no swap events occur with probability

¹Note that in this formulation we do not distinguish between a birth-death or a death-birth process [17]. Assume $r = 1$. In a true death-birth process, if a mutant dies, the probability of mutant division would be given by $(j-1)/N$ (assuming that a cell that just dies cannot divide). Similarly, in a true birth-death process, the probability of mutant death following a mutant division would be $(j-1)/N$ (assuming that an individual that just divided does not immediately die). In the present model formulation we do not include these considerations.

221 $1 - p_{migr}$. n_{ind} and n_{swaps} are non-negative integers that we set before performing a simulation.

222
223 To incorporate migration, let us assume that a migration update precedes a Moran update.
224 Then we can write,

$$\pi(t + 1/2) = \pi(t)P^{migr}, \quad \pi(t + 1) = \pi(t + 1/2)P^{Moran}, \quad (4)$$

where P^{migr} denotes the transition matrix associated with migrations. In order to formulate this transition matrix, we will need the Hypergeometric distribution, which describes the probability of picking n mutants in n_{ind} draws without replacement, if the total number of mutants is j out of N individuals:

$$\rho^j(n) = \binom{n_{ind}}{n} \binom{N - n_{ind}}{j - n} / \binom{N}{j}$$

The probability to change the number of mutants from m_1 to m_2 in a population characterized by vector $\pi(t)$ is given by

$$\mu_{m_1, m_2} = \sum_{n_1=0}^a \rho^{m_1}(n_1) \sum_{m=1}^N \pi_m(t) \rho^m(m_2 - m_1 + n_1).$$

Here, we assume that the population containing m_1 mutants loses n_1 mutants to another deme and gains n_2 mutants from the other deme; we sum over all possible values of n_1 , and note that

$$m_1 - n_1 + n_2 = m_2,$$

225 which gives us the expression for $n_2 = m_2 - m_1 + n_1$. The probability to lose n_1 individuals is
226 $\rho^{m_1}(n_1)$. The probability to gain n_2 individuals is calculated as follows: the donor deme is assumed
227 to contain m mutants (probability $\pi_m(t)$), because all individuals are assumed to obey the same
228 laws and the number of mutants in the deme at time t are drawn from the same probability distribu-
229 tion, $\pi(t)$. The probability to gain n_2 mutants from a deme containing m mutants is $\pi_m(t)\rho^m(n_2)$,
230 and to get the total probability of gaining n_2 individuals, we sum over all m .

231 Finally, the transition matrix for the migration step is given by

$$P_{ij}^{migr} = p_{migr}\mu_{ij} + (1 - p_{migr})\delta_{ij},$$

232 where δ_{ij} is the Kronecker delta.

233
234 The steady state, $\bar{\pi}$, satisfies the following equation:

$$\bar{\pi}_j = \sum_k P_{kj}^{Moran} \left(p_{migr} \sum_i \sum_m \bar{\pi}_i \bar{\pi}_m h_{imk} + (1 - p_{migr}) \bar{\pi}_k \right), \quad (5)$$

where

$$h_{imk} = \sum_n \rho^i(n) \rho^m(k - i + n).$$

235 **2.2.3 The stationary probability distribution and the selection-mutation balance in** 236 **the absence of migration**

237 In the presence of de novo forward and back mutations, let us determine the stationary probability
238 distribution for the number of mutants. Let us suppose the stationary probability distribution is
239 given by $(x_0, x_1, \dots, x_N)$ with

$$\sum_{j=0}^N x_j = 1. \quad (6)$$

240 The components x_j satisfy the following equations,

$$x_{j-1}P_{j-1}^\uparrow + x_{j+1}P_{j+1}^\downarrow - x_j(P_j^\uparrow + P_j^\downarrow) = 0, \quad 0 \leq j \leq N, \quad (7)$$

where $x_{-1} = x_{N+1} = 0$ and the transition probabilities are given by Equation (3). Let us suppose that the mutation rates are small compared with $1 - r$ and the inverse of the deme size $\frac{1}{N}$, and denote

$$u = \epsilon U, \quad u_b = \epsilon U_b,$$

$\epsilon \ll 1 - r, \epsilon \ll 1/N$. Solving Equation (7) in the zeroth order (i.e. under the assumption of no de novo mutations, setting $\epsilon = 0$) we obtain $x_i = 0$ for $1 \leq i \leq N - 1$, with probabilities x_0 and x_N undefined. To find the approximate solution under the assumption of small mutation rates, we go to the first order in ϵ . Let us set

$$x_0 = y_0, \quad x_N = y_N, \quad x_i = \epsilon y_i \quad \text{for } 1 \leq i \leq N - 1.$$

241 Taking into account only the terms of order ϵ in Equation (7), we obtain a degenerate system of
 242 $N + 1$ equations for $N + 1$ unknowns y_i . The additional condition is given by Equation (6) and in
 243 the lowest order in ϵ reduces to $y_N = 1 - y_0$. The solution is similar to what is found by [7, 51] in
 244 similar models, and is given by

$$y_0 = \frac{U_b}{r^{N-1}U + U_b} = \frac{u_b}{r^{N-1}u + u_b}, \quad (8)$$

$$y_N = 1 - y_0. \quad (9)$$

245 In particular, in the absence of de novo back mutations, the system converges to the $j = N$ state
 246 (mutant fixation), and in the absence of de novo forward mutations, we have $y_0 = 1$ (mutant
 247 extinction). If we include higher order terms, then the stationary distribution for the intermediate
 248 states is given by

$$y_i = \frac{Nr^{i-1}(N - i + ir)UU_b}{i(N - i)(r^{N-1}U + U_b)}, \quad 1 \leq i \leq N - 1. \quad (10)$$

249

Approximation (8)-(9) can also be obtained from the following simple consideration. Let us assume that the system spends most of the time in “pure” states (that is, in state $j = 0$ or in state $j = N$), which is the consequence of the time-scale separation: the waiting time to obtain a de novo mutation must be much longer than the typical time of mutant fixation. Then we can establish the balance of the following two processes. (1) If the system is in state $j = 0$ (denote this probability by y_0), then the transition to state $j = N$ happens at rate $Nu \times \phi_1$, which is the product of the mutant production rate, Nu , and the probability of a single mutant fixation given by ϕ_j with $j = 1$. (2) On the other hand, the probability of finding the system in state $i = N$ is $1 - y_0$, and the rate at which the system leaves and gets fixed at $i = 0$ is given by multiplying the rate of wild type production (Nu_b) by the probability of a wild-type fixation (ϕ_j) where we replace r with $1/r$, which is the relative fitness of the wild type compared to that of mutants. We obtain the equation

$$y_0Nu \frac{1 - 1/r}{1 - 1/r^N} = (1 - y_0)Nu_b \frac{1 - r}{1 - r^N},$$

250 whose solution y_0 is given by Equation (8).

251

252 To calculate the selection-mutation balance, we solve the equation $P_j^\uparrow = P_j^\downarrow$ for j and obtain
 253 Equation (2), where we used the largest contribution in ϵ (this is also the exact solution for the
 254 selection-mutation balance in the case of only forward mutation). We also note that back mutation
 255 does not effect the selection-mutation balance as long as $\epsilon \ll (1 - r)$. The exact solution for the
 256 selection-mutation balance is given by

$$j_{sel-mut} = N \frac{ru_b - r + u + 1 - \sqrt{((r(u_b - 1) + u + 1)^2 + 4(r - 1)u)}}{2(1 - r)}. \quad (11)$$

257 In the case that the two types are neutral with respect to each other ($r = 1$), we again solve the
 258 equation $P_j^\uparrow = P_j^\downarrow$ for j and as in [35] obtain

$$j_{sel-mut} = N \frac{u_b}{u + u_b}. \quad (12)$$

259

260 3 Results

261 3.1 Population fragmentation changes mutant distribution for neutral mutants

262 To motivate this study, we performed some simple experiments that examined the role of frag-
 263 mentation and migration in neutral mutant dynamics. 96 wells were filled with cells, such that
 264 each well contained 1% of (neutral) mutants. The process of cell migration was implemented by
 265 swapping a small percentage of cells between demes by using a pipette (which is described in more
 266 detail in Section 2.1 and in Supplementary Information Section 1). The number of mutants in the
 267 wells was assessed at the end of the experiment and compared with the control condition with no
 268 migration. The resulting experimentally obtained distribution of the mutant numbers is shown
 269 in Figures 2 and S1, where we plot the percent mutant contents in the 96 demes in the form of
 270 histograms. Figure 2 shows experimental results for the 1% mutants initial condition and Figure
 271 S1 shows experimental results for both the 0.1% and 1% mutants initial conditions.

272

273 We observed that while the mean number of mutants in the absence and in the presence of
 274 migration was the same, the distribution was significantly different; in particular, the distribution
 275 without migration had a much larger skewness, while in the presence of migration it was more
 276 symmetric (Figure 2). Specifically, the average percent of mutants without migration is 1.03%,
 277 and with migration is 1.13%, which is not significantly different using the T-test (p -value greater
 278 than 0.1). However, the Kolmogorov-Smirnov test between the two distributions gives a p -value of
 279 about 10^{-3} , which suggests that the distributions are significantly different. The skewness for the
 280 experiment without migration is 0.89, and with migration it is much smaller at 0.07.

281

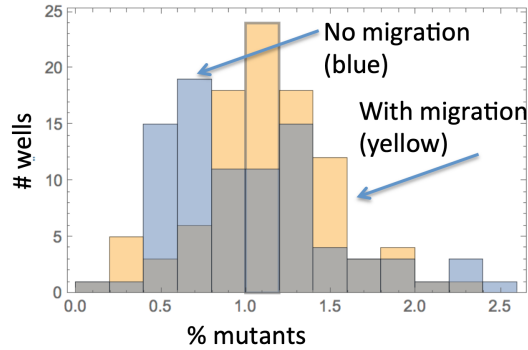


Figure 2: Effect of migration on neutral mutant distribution, experimental results. The blue bars represent the control condition without migration between the wells, the yellow bars represent the experimental condition with migration, and the gray color denotes the overlap. Initial condition of 1% mutants in each well. The average percent of mutants in each well without migration is 1.03%, and with migration is 1.13% (not significantly different using the T-test, p -value greater than 0.1). The Kolmogorov-Smirnov test between the two distributions gives a p -value of about 10^{-3} , which suggests that the distributions are significantly different. The skewness without migration is 0.89, and with migration it is much smaller at 0.07. Full experimental results are shown in Figure S1.

282 To explain these observations and extend the results to other conditions, we began by analyzing
 283 the dynamics of neutral mutants ($r = 1$).

284 Unimodal versus bimodal mutant distribution in the absence of de novo mutations.

285 To reproduce the experimental set-up, we assumed that there is no de novo mutation, and started
 286 with some small initial number of mutants in each deme (m_0). In order to analyze the effect of mi-
 287 gration/population fragmentation, we ran simulations with and without migration of cells between
 288 the distinct demes, a description of the mathematical model can be found in Section 2.

289

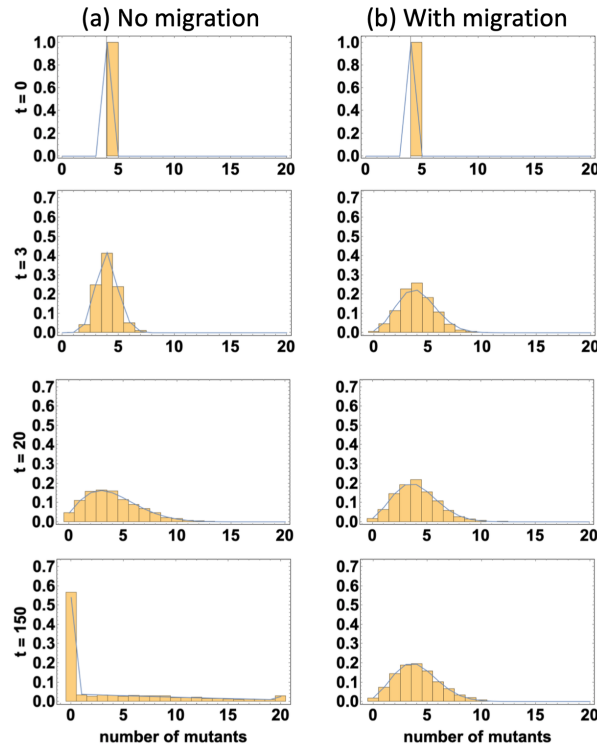


Figure 3: Stochastic simulations (histograms) and iterations of Equation (4) (blue lines). Panel (a) represents the absence ($p_{migr} = 0$) and panel (b) represents the presence ($p_{migr} = 1$) of migration. The probability distributions are presented at several moments of time (t in each plot corresponds to the number of discrete Moran steps). The rest of the parameters are $N = 20$, $m_0 = 4$, $n_{swaps} = 750$, $n_{ind} = 5$, $r = 1$, $u = u_b = 0$, and $K = 1.5 \times 10^3$.

290 As in the experimental results, in both the model simulations and analysis we found that the
 291 mean percent of neutral mutants is independent of migration, and is equal to the initial ratio of
 292 mutants in the system. This is because the transition probabilities are symmetric, see Equation
 293 (3) and Section 2 in the Supplementary Information for details. Furthermore, as the wild type and
 294 mutant are neutral, the probability for the mutant to fix within the system (assuming a non-zero
 295 migration rate) is equal to the initial frequency of mutants in the system (although the time to
 296 such fixation depends on migration [44, 54]). This is because in the context of a symmetric random
 297 walk, the fixation (i.e. absorption at the upper boundary) probability is proportional to the initial
 298 condition, see Supplementary Information Section 2 and [32, 34, 35].

299

300 On the other hand, the distribution of the frequency of mutants in the system and the dy-
 301 namics within individual demes are significantly influenced by the presence of migration (Figure
 302 3). Starting from a delta-like distribution (as initially all demes contain m_0 mutants), the distri-
 303 butions get wider with time and eventually reach a quasi-stationary distribution. In the absence

of migration (Figure 3(a)), the dynamics of each deme are independent from one another. Since the probability of fixation is simply the initial fraction of mutants (m_0/N), there is a large chance (given by $1 - m_0/N$) of mutant extinction in each deme. Therefore, the probability distribution of the number of mutants in each deme becomes flatter and develops a skew to the right, as most demes will trend toward mutant extinction, while a few will trend toward mutant fixation (i.e. a bimodal distribution, with modes at mutant extinction ($j = 0$) and fixation ($j = N$)). Eventually, as time $\rightarrow \infty$, all demes will be fixed at either mutant extinction or fixation. In the presence of migration (panel (b)), the dynamics of each deme are no longer independent from one another. As was similarly found in [49, 13, 12, 35, 57], migration makes all the demes look more similar to each other, resulting in a one-humped (unimodal) distribution. These results match well with *in vitro* experimental simulations of the computational model, which are shown in Figure 2 (see also Supplementary Information Section 1). Note that while the figure shows a long-term state, this is not an equilibrium, and the only outcomes as time $\rightarrow \infty$ is mutant fixation or mutant extinction in the whole system (individual demes, if $p_{migr} = 0$) because there is no de novo mutation [13, 35].

In addition to the extremes of no migration or a large amount of migration (where the system is well-mixed), we also investigate other regimes where there is some intermediate level of migration of individuals between the demes in the system. Figure S4 shows the time-evolution of the mutant probability distributions obtained by iterating Equation (4) (see panels (a-c) for three different values of p_{migr}), and then by plotting the resulting quasi-stationary probability distributions (panel (d)). Here we see that the effect of the Moran process is to “make” the probability distribution bimodal, and the effect of migration is to “make” it unimodal. The result is a trade-off of the two tendencies, and depending on the amount of migration, the distribution shape changes accordingly.

Quasi-stationary distributions become stationary in the presence of de novo mutation. Next, we expand the theory beyond the experimental conditions of Figure 2 to include the effect of de novo mutations. Since mutants are now generated stochastically, we alter the initial conditions to start with the wild type fixed in all demes ($m_0 = 0$). In the case of only forward mutation, the mutant will be created via mutation more often than the wild type. This mutational bias results in the neutral mutant ($r = 1$) fixing quickly in the entire population (see Figure S6(a-b)). On average, the time to fixation decreases with increasing migration, as faster migration results in more frequent introduction of the mutant into all of the demes, see [49, 44, 11, 54, 44].

In the case of both forward and back de novo mutation, the dynamics are more complex. Figure 4(a) shows simulations representing 2×10^3 demes of 20 individuals each, after 10^5 iterations with varying rates of migration. The histograms represent the number of neutral mutants per deme. In the absence of migration (left), the number of mutants will drift around, becoming extinct or fixed within a deme. In a highly fragmented population, this will happen more often, and the mutant will be at the extinction/fixation long-term state most of the time. Increasing migration (and/or decreasing population fragmentation by increasing the size of the demes, not shown) will result in fluctuation around the equilibrium value (Equation (12)) in each deme (Figure 4(a, right)). As in the simulations without de novo mutation, if the rate of forward and back mutation is equal ($u = u_b$), then migration does not change the expected mean number of neutral mutants (as the stationary distribution is symmetric around the selection-mutation balance of 50% mutants, see Figures S5 and S6(c-d)). If the rate of forward and back mutation is not equal ($u \neq u_b$), then the expected level of mutants is given by $\frac{u_b}{u+u_b}$ (see Equation (12) and [35] for details of this calculation).

Note also that the presence of mutations changes the nature of the long-term system behavior: the quasi-stationary distribution observed in the absence of de novo mutations (Figure 3(b), bottom graph) becomes a stationary distribution in the presence of forward and back mutation, as absorbing states no longer exist [13].

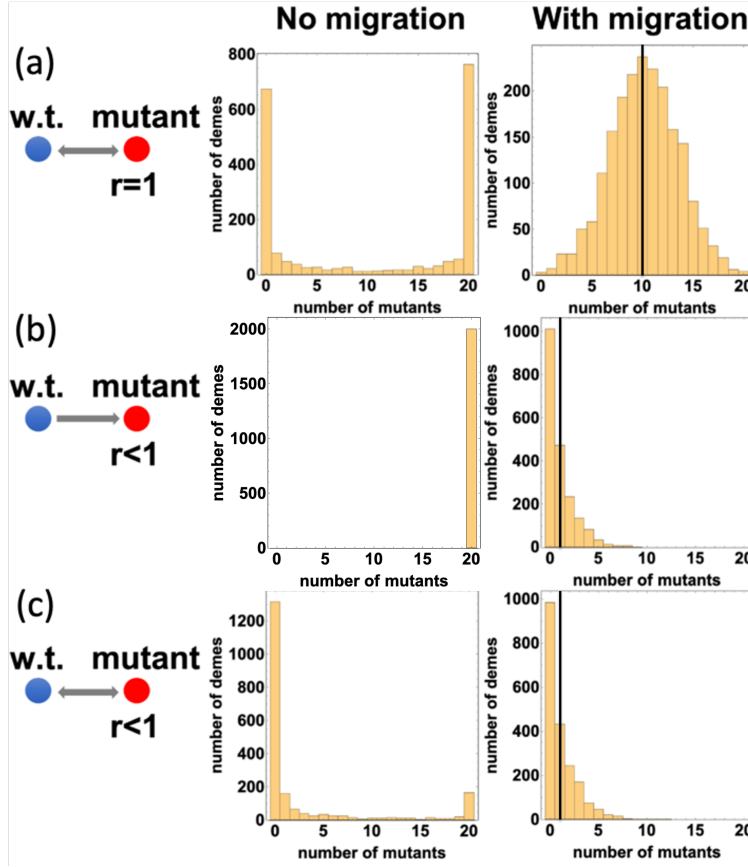


Figure 4: Histograms for the number of mutants per deme in the absence ($p_{migr} = 0$) and presence ($p_{migr} = 1$) of migration, after 10^5 Moran iterations. (a) Neutral mutant ($r = 1$) with both forward and back mutation ($u = u_b = 0.005$), see Figure S5 for intermediate migration cases; (b) disadvantageous mutant ($r = 0.9$) with forward mutation only ($u = 0.005$, $u_b = 0$), see Figure S12 for intermediate migration cases; (c) disadvantageous mutant ($r = 0.9$), with forward and backward mutation ($u = u_b = 0.005$), see Figure S13 for intermediate migration cases. The horizontal axis is the number of mutants and the vertical axis is the number of demes at that number of mutants. The vertical lines in the right panels represent the theoretical mutant equilibria, Equation (12) for (a) and the selection-mutation balance Equation (2) for panels (b-c). Other parameters are $N = 20$, $m_0 = 0$, $n_{swaps} = 100$, $n_{ind} = 10$, and $K = 2 \times 10^3$.

3.2 Population fragmentation changes mutant frequencies and distribution for disadvantageous mutants

Next we turn to the dynamics of disadvantageous mutants. While this scenario is highly biologically realistic for many populations, it is often more difficult to study experimentally due in part to the lower probability of mutant growth. In the absence of de novo mutation, a small initial number of disadvantageous mutants will likely decay quickly and go extinct. Therefore, we focus on mathematical models that include de novo mutation processes. We will show that while migration changes the distribution of demes in a similar manner for both disadvantageous and neutral mutants, in the disadvantageous case migration also changes the expected number of mutants at the (quasi)-stationary state in fragmented populations. This is related to the concept of “drift load” [29, 30, 56], which describes how the accumulation of deleterious mutations can cause a gradual reduction in population size (and in small populations random genetic drift will progressively overpower selection making it easier to fix future mutations). As we assume constant population sizes,

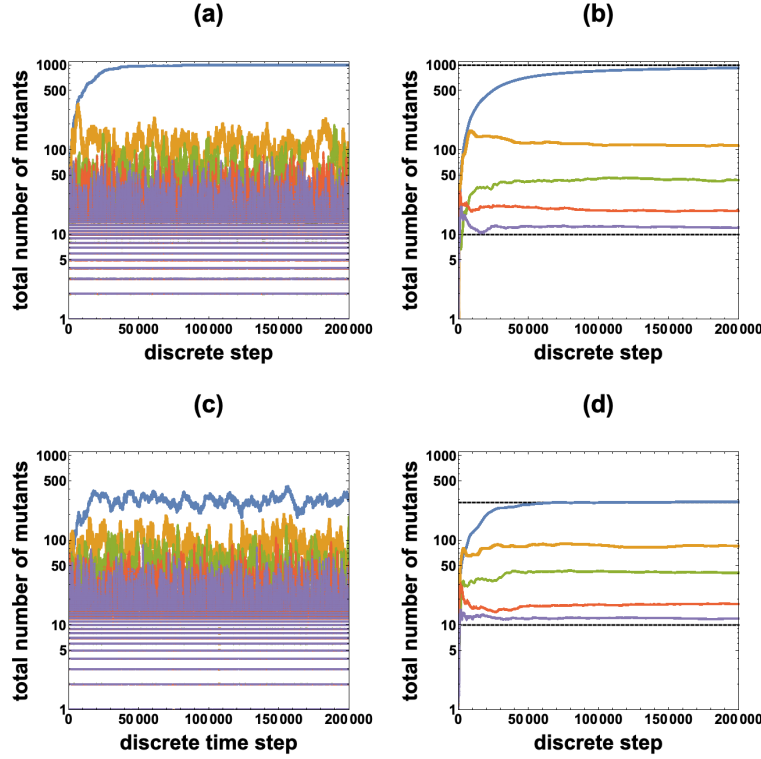


Figure 5: Number of mutants over time for varying rates of migration with a disadvantageous mutant ($r = 0.9$). Panels (a-b) include forward mutation only, and panels (c-d) include both forward and back mutation. Other parameters are $N = 10$, $m_0 = 0$, and $K = 100$. Selection-mutation balance is approximately 10 mutants in the system and mutant fixation is 10 mutants in each deme. Blue lines, no migration. The approximate expected number of mutants can be calculated using Equation (1). Yellow lines, low migration: $p_{migr} = \frac{1}{3}$, $n_{swaps} = 1$, and $n_{ind} = 1$. Green lines, medium migration: $p_{migr} = 1$, $n_{swaps} = 1$, and $n_{ind} = 1$. Red lines, high migration: $p_{migr} = 1$, $n_{swaps} = 5$, and $n_{ind} = 1$. Purple lines, very high migration: $p_{migr} = 1$, $n_{swaps} = 10$, and $n_{ind} = 5$. The approximate expected number of mutants is the selection-mutation balance. (a) Forward mutation only ($u = 10^{-3}$, $u_b = 0$), number of mutants in the system at each time step (typical runs). (b) Forward mutation only ($u = 10^{-3}$, $u_b = 0$), temporal average of the number of mutants in the system at each time step. Dashed lines represent the selection-mutation balance and mutant fixation. (c) Forward and back mutation ($u = u_b = 10^{-3}$), number of mutants in the system at each time step (typical runs). (d) Forward and back mutation ($u = u_b = 10^{-3}$), temporal average of the number of mutants in the system at each time step. Dashed lines represent the selection-mutation balance (Equation (2)) and the predicted average number of mutants under no migration (Equation (1)).

size fluctuation cannot occur; instead we observe elevated fractions of disadvantageous mutants depending on migration and population structure.

Fragmentation increases mutant numbers and decreases time to fixation. In the absence of de novo back mutations, mutant fixation in all demes is the only absorbing/stationary state, which will again eventually be reached with 100% probability. However, when there is a large amount of migration and/or a large, well-mixed population, then fixation will take a very long time and quasi-stationary states are possible [13, 44].

Figure 4(b) shows a system of small patches in histogram form in the absence and presence of migration, for disadvantageous mutants with only forward mutation. When the overall population is highly fragmented (no migration, left), fixation will occur quickly in each of the individual demes, and thus in the overall population as well. However, if the overall population is well-mixed, then fluctuation around a quasi-stationary state that is equal to the selection-mutation balance in each

deme is observed (panel (b, right)): the equilibrium value from Equation (2) is shown by the vertical line; see also time-series in Figure 5. Between these extreme scenarios, we observe that the overall system fluctuates around a quasi-stationary equilibrium that is between selection-mutation balance in each deme and complete fixation in the overall system (see Figure S12). We can see that in the case of disadvantageous mutants, population fragmentation does not only change the distribution of mutants, but also increases the expected number of disadvantageous mutants. To further illustrate this, Figure 5(a-b) shows the time course of the number of disadvantageous mutants for different rates of migration. Here we can see the (quasi)-stationary number of mutants in the system (as we run simulations over 2×10^5 discrete Moran steps), and that there are on average more mutants expected with lower migration rates (higher levels of population fragmentation), because fixation in each deme is more easily reached for fragmented (small) populations [55, 30]. In particular, in Figure 5(b) the (quasi)-stationary level of mutants goes from complete mutant fixation (blue line) under no migration to fluctuation around the selection-mutation balance (purple line, see Equation (2)) under a high migration regime.

Fragmentation increases mutant frequencies even when fixation is not an absorbing state. In the case of both de novo forward and back mutations, there are no longer any absorbing states.

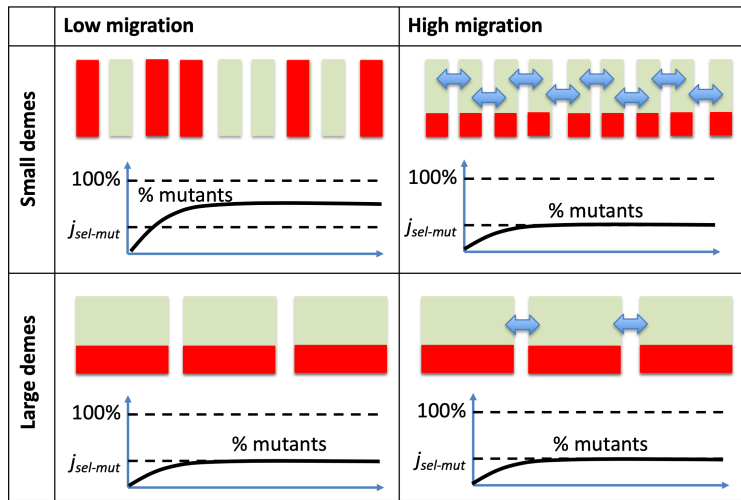


Figure 6: Summary of results: de novo forward and back mutation with varying migration (columns) and deme size (rows), assuming a constant total population. There are no absorbing states. Individual demes are represented as green rectangles, and the level of mutants in each is shown in red. Total frequency of mutants panels schematically show the percent of mutants as a function of time; the black dashed lines represent the selection-mutation balance, $j_{sel-mut}$, and 100% fixation.

Figure 4(c) shows histograms for the number of disadvantageous mutants in the absence and presence of migration, with the inclusion of back mutation. The dynamics are similar to the forward mutation only case (panel (b)), except the quasi-stationary distributions described in the preceding paragraph are now stationary distributions, as demes will not all eventually trend toward fixation (left panel). In particular, depending on the level of population fragmentation, demes will either fluctuate around a stationary value, or will individually bounce back and forth between mutant extinction and mutant fixation. In the latter case (high fragmentation), the system is characterized by a higher expected number of mutants compared to the well-mixed (or high migration rate) system. As seen in Figure 5(c-d), since mutant fixation is no longer an absorbing state, we expect a smaller number of mutants compared to when there is only forward mutation (panels (a-b)). The

expected number of mutants in the absence of migration can be computed in the case of a small mutation rate, according to Equation (1). As the amount of migration increases, the expected number of mutants converges to the selection-mutation balance given by Equation (2). In particular, in Figure 5(d) the number of mutants goes from an elevated predictable number (blue line, see Equation (1)) under no migration to fluctuation around the selection-mutation balance (purple line, see Equation (2)) under a high migration regime. The overall dynamics for different cases are summarized schematically in Figure 6.

The role of the effective population size, N_e . The effective population size (N_e) is the number of individuals in an “idealized” (panmictic) population that would be characterized by a specified quantity (measuring the strength of genetic drift such as variance, coalescence, etc) that is equal to that in the real population. Population fragmentation into subdivided demes can increase [36, 49] or decrease [53] the effective population size. As shown in [36] (Wright-Fisher type island model), and [49] (Moran type island model), in the case where the overall population structure does not change for a long evolutionary time (as is the case in our model), the effective population size of a subdivided population can be much larger than the total population size (and is larger with lower levels of migration). At the same time, the effective selection coefficient (s_e) becomes smaller. A natural question is then: is the change in N_e and s_e sufficient to explain the observed differences in mutant dynamics between non-fragmented and fragmented populations, including a higher frequency of deleterious mutations?

In Supplementary Information Section 3.1 we discuss the diffusion approximation for our fragmented model with migration, and obtain the following approximations for the effective population size and effective selection coefficient:

$$N_e = NK \left(1 + \frac{K}{2Np_{migr}n_{swaps}n_{ind}} \right), \quad (13)$$

$$s_e = s \left(1 + \frac{K}{2Np_{migr}n_{swaps}n_{ind}} \right)^{-1}. \quad (14)$$

Figure S7 illustrates the validity of the diffusion approximation and motivates the definition of the “variance” effective population size, Equation (13), see also Figure S8 that compares this approximation with numerically obtained values for N_e .

In order to obtain the predicted average number of mutants in a fragmented system, based on the effective population size, we can apply Equation (11), where the system size is given by N_e and mutant fitness $r = 1 - s_e$. The results are presented in Figure S9. While the general trend is qualitatively captured (higher degrees of fragmentation result in a larger number of disadvantageous mutants), the size of the effect predicted by this substitution is not quantified correctly.

This suggests that, as noted in [49], it is not only a change in effective population size that distinguishes the subdivided population with many demes and migration from the singular panmictic one. In other words, a rescaling of population size (and adjusting the selection coefficient) does not make the two populations equivalent. If a fragmented population with effective population size N_e is replaced with a well-mixed population of size N_e , the number of mutants will increase compared to a well-mixed population of size NK . In addition, the frequency of mutants in each individual deme will be very different: a single large deme at selection-mutation balance (well-mixed population) versus many demes that are either completely wild type or completely mutant (fragmented population). As a consequence, the mutant dynamics in a fragmented system will proceed in a

qualitatively different way, and the expected number of mutants will not be governed by a simple balance between production and selection, as it is in a panmictic system.

When can we expect to see more mutants than predicted by selection-mutation balance? The number of disadvantageous mutants is amplified relative to a well-mixed population when the population is highly fragmented (that is, the individual patches are sufficiently small) and the migration rate is not too high, see Figure 6.

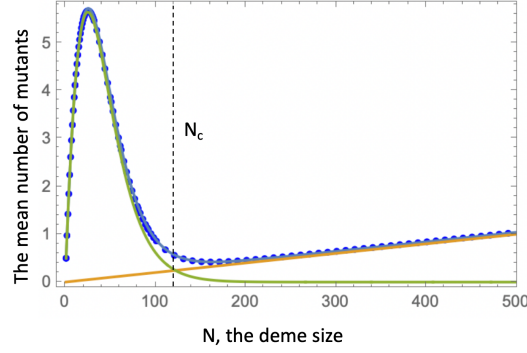


Figure 7: Estimating N_c . The expected number of mutants in a single deme in the absence of migration is shown as a function of N ; it is computed numerically (blue circles) by determining the principal eigenvector of the transition matrix, see Equation (3), and also by using approximations (8) and (10) (blue line). The green line represents the fast fixation regime (Ny_N , Equation (1)); the yellow line is the selection-mutation balance, $j_{sel-mut}$ (Equation (2)). The parameters are $u = u_b = 10^{-4}$ and $r = 0.95$. The threshold value N_c is shown by the dashed vertical line.

Even in the absence of migration, if each deme size is too large, then fixation will almost never be reached and fluctuation around the selection-mutation balance in each deme will be observed instead. On the other hand, if the deme size is very small ($N = 2$), then the expected number of mutants is approximately 50% of the system, as each deme will spend about 50% of the time at mutant extinction and 50% of the time at mutant fixation because of the small mutation rate (not shown). As the number of individuals per deme increases, this effect of fixation continues to elevate the number of mutants, but contributes less and less as the fixation probability decreases. Therefore, as the deme size grows, the expected number of mutants (the blue line in Figure 7) will extrapolate between two regimes: (i) the fast fixation regime, where the mean number of mutants in a deme is given by Ny_N (Equation (1), green line in Figure 7) and drift dominates, and (ii) the selection-mutation balance (Equation (2), yellow line in Figure 7) where selection dominates. To estimate the threshold deme size, N_c , above which the expected frequency of mutants becomes close to selection-mutation balance, we find the intersection of the fast-fixation (green) and selection-mutation balance (yellow) lines by solving the equation $Ny_N = j_{sel-mut}$ for N :

$$N_c = \frac{\log\left(\frac{1-r-u}{ru_b}\right)}{-\log r}. \quad (15)$$

As N_c represents the threshold value for which drift can overpower selection if $N < N_c$ and selection (of the disadvantageous mutant) overpowers drift for $N > N_c$, we have that N_c can be thought of as an approximation of the “selection effective population size” for our model in the absence of migration [28, 50, 10]. For more details on these calculations, see Section 3 of the Supplementary Information. If the deme size is smaller than N_c , a significantly larger number of mutants compared to the selection-mutation balance is expected. However, if demes are connected

to each other and migration is present, this may weaken the effect. Under intense migration, the expected number of mutants tends to that predicted by selection-mutation balance. Therefore, an important question is: what is the level of migration that is sufficient to lower the mutant levels back to that of selection-mutation balance?

In this model, the overall intensity of migration (monotonically) depends on several parameters (see Table 1): the probability of a migration event per update (p_{migr}), the number of swaps during a migration event (n_{swaps}), and the number of individuals exchanged during a swapping event (n_{ind}). To simplify the discussion, we will fix two of these to $n_{swaps} = K/5$ and $n_{ind} = N/5$, focusing on the parameter p_{migr} as the one parameter determining the rate of migration.

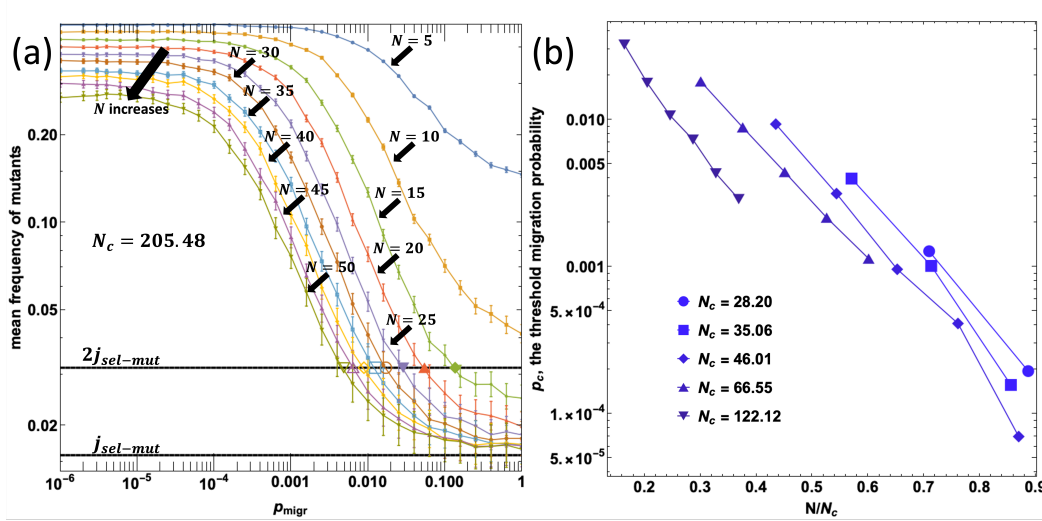


Figure 8: The role of migration in the level of mutants. (a) The mean frequency of mutants as a function of p_{migr} , calculated as a temporal average over $10^{8.5}$ time-steps; the bars represent the standard error. Different curves correspond to different values of N . The horizontal lines are $j_{sel-mut}$ and $2j_{sel-mut}$. The parameters are $u = u_b = 10^{-3.5}$, $r = 0.98$, and $N_c = 205.48$. (b) The threshold values, p_c , are plotted against the corresponding N/N_c , for several values of N_c . The exponent B (Equation (16)) is 10.7 ± 0.9 . The rest of the parameters are $K = 20$, $n_{swaps} = K/5$, and $n_{ind} = N/5$.

Figure 8(a) demonstrates how a threshold value of the migration probability can be calculated. Fixing the values of u , u_b , and r , simulations were run for different choices of the deme size, $N < N_c$, and the mean frequency of mutants (that is, the mutant number divided by the total population size, NK) was determined for each p_{migr} . As anticipated, the expected mutant frequencies are higher than the level predicted by the selection-mutation balance; also, they decrease with the deme size, N , and migration probability, p_{migr} . To quantify the migration probability that, for each N , corresponds to a significant decay in the mutant population, we defined p_c as the value of p_{migr} that leads the frequency of mutants to fall to twice the selection-mutation balance. In Figure 8(a), intersections of the mutant frequencies with $2j_{sel-mut}$ are marked with colored symbols and their horizontal coordinate gives p_c . This quantity decreases with N .

Figure 8(b) shows the threshold migration rate as a function of N/N_c for several different values of N_c . We observe that the dependence is exponential, and propose the empirical law:

$$p_c = Ae^{-BN/N_c}, \quad (16)$$

where the constants A and B do not depend on N . The value of the exponent, B , can be found by fitting (see Figure 8); it is difficult to derive analytically because it falls in the intermediate

migration rate regime (where our approximations of no migration or strong migration do not apply). Therefore, we use numerical approximations to calculate it, as shown in Figure 8. Overall, Figure 8(b) shows that the threshold migration rate (p_c) decays exponentially with N/N_c , which is the deme size divided by the threshold deme size (N_c). This implies that even for small increases in deme size $N < N_c$, drastic decreases in the migration rate $p_{migr} < p_c$ are needed to maintain the inflated number of mutants in the fragmented population compared to the expected number of mutants at selection-mutation balance. Section 3.2 of the Supplementary Information presents similar results obtained in the case of the Wright-Fisher model.

4 Discussion and Conclusion

Divided and fragmented habitats are common in nature. Some examples are naturally occurring habitats such as islands (in the context of biogeography), aquatic habitats separated by land (such as ponds or lakes), different parts of a plant that can be inhabited by lower organisms, or hosts that are inhabited by ecto- and endoparasites. Human activities may lead to further fragmentation of natural environments, which has implications for conservation biology. In general, most natural habitats are spatially structured, and are likely to be characterized by demes or microdemes, with population movement between them, such as patches of high moisture or nutrient availability across a larger habitat. Given the ubiquity of fragmented and deme-structured habitats in nature, it is important to obtain a better understanding of how such structures impact evolutionary dynamics. As mentioned in the introduction, several aspects of evolution in fragmented populations have been explored in the literature. Importantly, it has been shown that mutant fixation times can be significantly increased in deme-structured habitats, even though the probability of mutant fixation remains unaltered. Other aspects of evolution in deme-structured and fragmented habitats, however, remain to be explored in more detail from an evolutionary theory point of view, but also from an experimental point of view to verify model predictions. A more complete understanding of these dynamics is crucial for better understanding evolutionary processes in natural populations.

Experimentally testing the validity of evolutionary mathematical models of fragmented and deme-structured populations can be challenging, yet is an important component of this work. We set up experiments in which murine cell colonies were grown in 96 wells, with migration implemented as swapping small numbers of cells between randomly chosen wells with a pipette. We used this system to quantify the distribution of neutral mutants across the demes / wells. The experimental findings confirmed that while the mean number of mutants is not influenced by migration, the probability distribution is, consistent with theoretical predictions. We used the same model to also investigate the distribution of disadvantageous mutants across the demes, which was not feasible to follow experimentally. Furthermore, we extended the model to include de novo mutations and examined the average level at which the mutants were expected to persist, and compare this to the level that is expected due to the balance between mutation and negative selection. We showed that the mutant numbers experience an increase in frequency compared to that of the selection-mutation balance of a non-fragmented system. We investigated this phenomenon; using the diffusion approximation, we found that this increase cannot be simply explained by an elevation in the effective population size and a decrease in the selection coefficient, which are consequences of fragmentation. In fact, an increase in the effective population size does not capture a profound change in mutant dynamics brought about by fragmentation and expressed in a shift from selection-mutation balance to the dynamics of intermittent mutant fixation and extinction events. In a single deme, we found that the increase (compared to the selection-mutation level) is observed when the deme size is lower than the critical size, N_c , given by Equation (15). In a fragmented system that consists of connected demes with a probability of migration, the increase in mutant numbers above the selection-mutation balance can be observed in small ($N < N_c$) demes as long as the migration rate is sufficiently small. The migration rate above which the mutants approach the selection-mutation balance decays

exponentially with N/N_c , see Equation (16).

Implications for evolutionary biology. This work has important implications for issues surrounding standing genetic variation in populations. Alleles that are mildly deleterious in current environmental conditions can become advantageous later on, when the environment changes to their favor. Such changes in the environment can arise naturally, or can be induced by xenobiotics, such as drugs, pesticides, or pollutants. Resistance evolution (such as pesticide or antibiotic resistance) is of particular applied relevance in this respect, because resistant mutants often suffer a fitness cost in the absence of the treatment and only become advantageous once treatment is initiated. The pre-existence of resistant mutants at the selection-mutation balance is an important determinant of outcome in such cases. Our mathematical modeling results indicate that the pool of pre-existing resistant mutants, and more generally the pool of deleterious mutants that persist in a population, can be significantly higher in a deme-structured or fragmented population compared to the predictions made by models without such population structure. This means that in a lot of natural settings, slightly deleterious mutants might be present at significantly greater abundances than previously thought, which has relevance for understanding the adaptability of populations, as well as for preventing xenobiotic adaptation. Besides a change in environment, the higher level at which disadvantageous mutants might persist can also speed up the crossing of fitness valleys, where a first mutation can lead to a selective disadvantage, but an additional mutation results in an overall advantage. An important factor that determines whether these effects occur is the population size within individual demes relative to the migration rate, as mentioned above. Higher levels of deleterious mutant persistence will most likely occur if the within deme population size is relatively small, as defined mathematically in our modeling framework.

Somatic evolutionary processes. In addition, and on a more speculative level, there are also biomedical implications of our findings, if cell populations *in vivo* are viewed as a kind of ecosystem in which cells evolve over time. Clonal evolution takes place within tissues as individuals age. This has been clearly documented in the hematopoietic system [24], where a variety of mutant clones with different characteristics emerge over time. These mutant clones can potentially lead to a functional deterioration of the healthy tissue, and also in the longer term to the development of malignancies. These mutants can be disadvantageous, neutral, or advantageous, and can form the basis for further mutation accumulation. These evolutionary processes take place in the bone marrow, where stem cells exist in niches, with traffic between different parts of the bone marrow via the blood [52]. There is currently no detailed data that quantifies the rate at which cell populations across the individual niches/demes communicate and the rate at which they migrate. Given the intricate anatomical structures in the hematopoietic system, and given the tendency for cells to home to their specific microenvironments [45], it is likely that cell dynamics are largely governed by local processes, with a relatively low rate at which cells move from one location to another. If this is true, then according to our model, the deme structure can influence the exact spatial genetic composition of the cell population, with mutants being dominant in some parts of the bone marrow but not others, especially for disadvantageous and neutral mutants. This in turn can have implications for bone marrow biopsies when trying to assess or monitor the clonal composition of the hematopoietic system.

Tumor evolution. Similar considerations can apply to tumors in the hematopoietic system once they have started to expand. Although tumors grow over time and we have considered the evolutionary dynamics in constant populations, tumors can be characterized by periods of slow growth or temporary stasis until further mutants are generated that allow the cells to overcome specific selective barriers. In the hematopoietic system, an example could be slowly growing/indolent cases of chronic lymphocytic leukemia (CLL), where cells grow in spatially separated lymph nodes, the

spleen, and the bone marrow. The evolution of drug-resistant mutants is a major problem that results in the eventual failure of therapies, and the level at which resistant mutants exist before the start of treatment tends to be an important determinant of the time to disease relapse [20, 1]. Mathematical models have been used to calculate the number of drug-resistant mutants, for example in chronic lymphocytic leukemia [20] or chronic myeloid leukemia [21], and these models assumed a spatially homogeneous growing tumor cell population. If drug-resistant mutants carry a fitness cost and are therefore disadvantageous before the start of therapy, the models analyzed here indicate that the organization of cells into demes can have a significant influence on the abundance of pre-existing mutants. Depending on the deme size and the migration rate, the number of resistant cells can be significantly larger than predicted by the selection-mutation balance, which could dramatically speed up the rate at which the tumor relapses during therapy. If the drug-resistant mutants are advantageous, which happens in the presence of treatment, then the spatial structure could also result in a significant effect on quantities such as the timing of mutant expansion (and therefore treatment failure). Advantageous mutant dynamics, however, are beyond the scope of the present study. There are also implications for sampling strategies when attempting to assess the burden of drug-resistant mutants before therapy, such that the true genetic diversity across the different locations is determined, rather than a skewed picture arising from the analysis of one or just a few locations. This concept of optimal sampling of tumors has also been explored with spatially explicit computational models in a different context [58, 37]. Our analysis particularly highlights the need to experimentally measure the rate at which lymphocytes redistribute from one lymph node compartment to another, which would allow a more accurate prediction about how the deme structures in the hematopoietic system influence the evolution of CLL cells.

Spatial migration. Finally, we note that in our model analysis, migration is assumed to be among randomly (uniformly) chosen demes. For migration that is spatially restricted, disadvantageous mutant levels will be elevated compared to non-spatially restricted migration because once a region of demes becomes fixed with the mutant, it is less likely that the wild type will be reintroduced (for geometric reasons). Therefore, spatially restricted migration increases population fragmentation compared to mixed migration, which increases the number of mutants. Including different spatially restricted patterns of migration could be an interesting extension of our current work.

References

- [1] Jan A Burger, Dan A Landau, Amaro Taylor-Weiner, Ivana Bozic, Huidan Zhang, Kristopher Sarosiek, Lili Wang, Chip Stewart, Jean Fan, Julia Hoellenriegel, et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. Nature communications, 7(1):1–13, 2016.
- [2] Partha Pratim Chakraborty, Louis R. Nemzer, and Rees Kassen. Experimental evidence that metapopulation structure can accelerate adaptive evolution. bioRxiv preprint, July 2021.
- [3] Joshua L Cherry. Selection in a subdivided population with local extinction and recolonization. Genetics, 164(2):789–795, 2003.
- [4] Joshua L. Cherry and John Wakeley. A diffusion approximation for selection and drift in a subdivided population. Genetics, 163(1):421–428, January 2003.
- [5] Arolyn Conwill, Anne C. Kuan, Ravalika Damerla, Alexandra J. Poret, Jacob S. Baker, A. Delphine Tripp, Eric J. Alm, and Tami D. Lieberman. Anatomy promotes neutral coexistence of strains in the human skin microbiome. Cell Host & Microbe, January 2022.

- [6] David J. D. Earn, Simon A. Levin, and Pejman Rohani. Coherence and Conservation. Science, 290(5495):1360–1364, 2000. Publisher: American Association for the Advancement of Science
_eprint: <https://science.sciencemag.org/content/290/5495/1360.full.pdf>.
- [7] Drew Fudenberg, Martin A. Nowak, Christine Taylor, and Lorens A. Imhof. Evolutionary game dynamics in finite populations with strong selection and weak mutation. Theoretical Population Biology, 70(3):352–363, November 2006.
- [8] Connie Fung, Shumin Tan, Mifuyu Nakajima, Emma C Skoog, Luis Fernando Camarillo-Guerrero, Jessica A Klein, Trevor D Lawley, Jay V Solnick, Tadashi Fukami, and Manuel R Amieva. High-resolution mapping reveals that microniches in the gastric glands control *helicobacter pylori* colonization of the stomach. PLoS biology, 17(5):e3000231, 2019.
- [9] Diana Fusco, Matti Gralka, Jona Kayser, Alex Anderson, and Oskar Hallatschek. Excess of mutational jackpot events in expanding populations revealed by spatial Luria–Delbrück experiments. Nature Communications, 7:12760, 2016.
- [10] Simon Gravel. When Is Selection Effective? Genetics, 203(1):451–462, May 2016.
- [11] Ilkka Hanski. Metapopulation dynamics. Nature, 396(6706):41–49, November 1998.
- [12] Ilkka Hanski and Michael Gilpin. Metapopulation dynamics: brief history and conceptual domain. Biological Journal of the Linnean Society, 42(1-2):3–16, January 1991.
- [13] Christoph Hauert and Lorens A. Imhof. Evolutionary games in deme structured, finite populations. Journal of theoretical biology, 299:106–112, April 2012. Place: England.
- [14] Konrad Kalarus and Piotr Nowicki. How do landscape structure, management and habitat quality drive the colonization of habitat patches by the dryad butterfly (Lepidoptera: Satyri-
nae) in fragmented grassland? PloS one, 10(9):e0138557, 2015.
- [15] Peter Kareiva, A. Mullen, and R. Southwood. Population dynamics in spatially complex environments: Theory and data. Philosophical Transactions: Biological Sciences, 330(1257):175–190, 1990. Publisher: The Royal Society.
- [16] Samuel Karlin and James McGregor. On a genetics model of Moran. Mathematical Proceedings of the Cambridge Philosophical Society, 58(2):299–311, 1962.
- [17] Kamran Kaveh, Natalia L. Komarova, and Mohammad Kohandel. The duality of spatial death-birth and birth-death processes and limitations of the isothermal theorem. Royal Society open science, 2(4):140465, April 2015.
- [18] Benjamin Kerr, Claudia Neuhauser, Brendan J. M. Bohannan, and Antony M. Dean. Local migration promotes competitive restraint in a host–pathogen ‘tragedy of the commons’. Nature, 442:75–78, 2006.
- [19] M Kimura. On the probability of fixation of mutant genes in a population. Genetics, 47(6):713–719, June 1962.
- [20] Natalia L Komarova, Jan A Burger, and Dominik Wodarz. Evolution of ibrutinib resistance in chronic lymphocytic leukemia (CLL). Proceedings of the National Academy of Sciences, 111(38):13906–13911, 2014.
- [21] Natalia L Komarova and Dominik Wodarz. Drug resistance in cancer: principles of emergence and prevention. Proceedings of the National Academy of Sciences, 102(27):9714–9719, 2005.

- [22] Tonya A. Lander, Stephen A. Harris, Patricia J. Cremona, and David H. Boshier. Impact of habitat loss and fragmentation on reproduction, dispersal and species persistence for an endangered Chilean tree. Conservation Genetics, 20(5):973–985, October 2019.
- [23] Claire Lavigne, Xavier Reboud, Madeleine Lefranc, Emmanuelle Porcher, Fabrice Roux, Isabelle Olivieri, and Bernard Godelle. Evolution of genetic diversity in metapopulations: *Ara-bidopsis thaliana* as an experimental model. Genetics Selection Evolution, 33, 2001.
- [24] Henry Lee-Six, Nina Friesgaard Øbro, Mairi S Shepherd, Sebastian Grossmann, Kevin Dawson, Miriam Belmonte, Robert J Osborne, Brian JP Huntly, Inigo Martincorena, Elizabeth Anderson, et al. Population dynamics of normal human blood inferred from somatic mutations. Nature, 561(7724):473–478, 2018.
- [25] Simon A. Levin. Dispersion and Population Interactions. The American Naturalist, 108(960):207–228, March 1974. Publisher: The University of Chicago Press.
- [26] Richard Levins. Some Demographic and Genetic Consequences of Environmental Heterogeneity for Biological Control1. Bulletin of the Entomological Society of America, 15(3):237–240, 1969. eprint: <https://academic.oup.com/ae/article-pdf/15/3/237/18805526/besa15-0237.pdf>.
- [27] Laurence Loewe and William G Hill. The population genetics of mutations: good, bad and indifferent. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 365(1544):1153–1167, April 2010. Publisher: The Royal Society.
- [28] Michael Lynch. The Origins of Genome Architecture. Oxford University Press, Incorporated, June 2007.
- [29] Michael Lynch, John Conery, and Reinhard Burger. Mutation Accumulation and the Extinction of Small Populations. The American Naturalist, 146(4):489–518, 1995. Publisher: [University of Chicago Press, American Society of Naturalists].
- [30] Michael Lynch and Wilfried Gabriel. Mutation load and the survival of small populations. Evolution, 44(7):1725–1737, November 1990. Publisher: John Wiley & Sons, Ltd.
- [31] Eugenio López-Cortegano, Ramón Pouso, Adriana Labrador, Andrés Pérez-Figueroa, Jesús Fernández, and Armando Caballero. Optimal Management of Genetic Diversity in Subdivided Populations. Frontiers in Genetics, 10:843, 2019.
- [32] T. Maruyama. On the fixation probability of mutant genes in a subdivided population. Genetical research, 15(2):221–225, April 1970. Place: England.
- [33] T. Maruyama. A simple proof that certain quantities are independent of the geographical structure of population. Theoretical population biology, 5(2):148–154, April 1974. Place: United States.
- [34] P. A. P. Moran. Random processes in genetics. Mathematical Proceedings of the Cambridge Philosophical Society, 54(1):60–71, 1958. Publisher: Cambridge University Press.
- [35] P. A. P. Moran. The statistical processes of evolutionary theory. Clarendon Press, Oxford, 1962.
- [36] Masatoshi Nei and Naoyuki Takahata. Effective population size, genetic diversity, and coalescence time in subdivided populations. Journal of Molecular Evolution, 37(3), September 1993.
- [37] Luka Opasic, Da Zhou, Benjamin Werner, David Dingli, and Arne Traulsen. How many samples are needed to infer truly clonal mutations from heterogenous tumours? BMC cancer, 19(1):1–11, 2019.

- [38] Z Patwa and L M Wahl. The fixation probability of beneficial mutations. Journal of the Royal Society, Interface, 5(28):1279–1289, November 2008. Publisher: The Royal Society.
- [39] Laura R. Prugh, Karen E. Hodges, Anthony R. E. Sinclair, and Justin S. Brashares. Effect of habitat area and isolation on fragmented animal populations. Proceedings of the National Academy of Sciences of the United States of America, 105(52):20770–20775, December 2008.
- [40] E. Rochat, S. Manel, M. Deschamps-Cottin, I. Widmer, and S. Joost. Persistence of butterfly populations in fragmented habitats along urban density gradients: motility helps. Heredity, 119(5):328–338, November 2017.
- [41] Daniel Rubinoff and Jerry A. Powell. Conservation of fragmented small populations: endemic species persistence on California’s smallest channel island. Biodiversity & Conservation, 13(13):2537–2550, December 2004.
- [42] Simon A. Levin and Thomas M. Powell. Patch Dynamics, volume 96 of Lecture Notes in Biomathematics. Springer-Verlag Berlin Heidelberg, 1993.
- [43] M. Slatkin. Gene flow and genetic drift in a species subject to frequent local extinctions. Theoretical population biology, 12(3):253–262, December 1977. Place: United States.
- [44] Montgomery Slatkin. Fixation probabilities and fixation times in a subdivided population. Evolution; international journal of organic evolution, 35(3):477–488, May 1981. Place: United States.
- [45] Beatriz Suárez-Álvarez, Antonio López-Vázquez, and Carlos López-Larrea. Mobilization and homing of hematopoietic stem cells. Stem cell transplantation, pages 152–170, 2012.
- [46] J A Thomas, N A Bourn, R T Clarke, K E Stewart, D J Simcox, G S Pearman, R Curtis, and B Goodger. The quality and isolation of habitat patches both determine where butterflies persist in fragmented landscapes. Proceedings. Biological sciences, 268(1478):1791–1796, September 2001.
- [47] Michael J. Wade and David E. McCauley. Extinction and recolonization: their effects on the genetic differentiation of local populations. Evolution; international journal of organic evolution, 42(5):995–1005, September 1988. Place: United States.
- [48] John Wakeley. Polymorphism and divergence for island-model species. Genetics, 163(1):411–420, 2003.
- [49] John Wakeley and Tsuyoshi Takahashi. The many-demes limit for selection and drift in a subdivided population. Theoretical population biology, 66(2):83–91, September 2004. Place: United States.
- [50] J. Wang, E. Santiago, and A. Caballero. Prediction and estimation of effective population size. Heredity, 117(4):193–206, October 2016. Number: 4 Publisher: Nature Publishing Group.
- [51] Warren J. Ewens. Mathematical Population Genetics, volume 27 of Interdisciplinary Applied Mathematics. Springer-Verlag New York, 2007.
- [52] Samiksha Wasnik, Wanqiu Chen, Abu S.I. Ahmed, Xiao-Bing Zhang, Xiaolei Tang, and David J. Baylink. Chapter 5 - hsc niche: Regulation of mobilization and homing. In Ajaykumar Vishwakarma and Jeffrey M. Karp, editors, Biology and Engineering of Stem Cell Niches, pages 63–73. Academic Press, Boston, 2017.
- [53] M C Whitlock and N H Barton. The effective size of a subdivided population. Genetics, 146(1):427–441, May 1997.

- 775 [54] Michael C Whitlock. Fixation probability and time in subdivided populations. Genetics,
776 164(2):767–779, June 2003.
- 777 [55] Michael C. Whitlock and Reinhard Bürger. Fixation of New Mutations in Small Populations. In
778 Denis Couvet, Régis Ferrière, and Ulf Dieckmann, editors, Evolutionary Conservation Biology,
779 Cambridge Studies in Adaptive Dynamics, pages 155–170. Cambridge University Press, Cam-
780 bridge, 2004.
- 781 [56] Y. Willi, P. Griffin, and J. Van Buskirk. Drift load in populations of small size and low density.
782 Heredity, 110(3):296–302, March 2013.
- 783 [57] Sewall Wright. Evolution in Mendelian populations. Genetics, 16(2):97, March 1931.
- 784 [58] Boyang Zhao, Justin R. Pritchard, Douglas A. Lauffenburger, and Michael T. Hemann.
785 Addressing Genetic Tumor Heterogeneity through Computationally Predictive Combination
786 Therapy. Cancer Discovery, 4(2):166–174, 2014. Publisher: American Association for Cancer
787 Research _eprint: <https://cancerdiscovery.aacrjournals.org/content/4/2/166.full.pdf>.