Impact of spatial structure on early and long-term adaptation in rugged fitness landscapes

Richard Servajean[1,2], Arthur Alexandre[1,2], Anne-Florence Bitbol[1,2,*]

1 Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
2 SIB Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland
*Email: anne-florence.bitbol@epfl.ch

Abstract

We investigate the exploration of rugged fitness landscapes by spatially structured populations with demes on the nodes of a graph, connected by migrations. In the rare migration regime, we find that finite structures can adapt more efficiently than very large ones, especially in high-dimensional fitness landscapes. Furthermore, we show that, in most landscapes, migration asymmetries associated with some suppression of natural selection allow the population to reach higher fitness peaks first. In this sense, suppression of selection can make early adaptation more efficient. However, the time it takes to reach the first fitness peak is then increased. We also find that suppression of selection tends to enhance finite-size effects. We extend our study to frequent migrations, suggesting that our conclusions hold in this regime. We then investigate the impact of spatial structure with rare migrations on long-term evolution by studying the steady state of the population. For this, we define an effective population size for the steady-state distribution. We find that suppression of selection is associated to reduced steady-state effective population sizes, and reduced average steady-state fitnesses.

Introduction

Natural microbial populations often possess complex spatial structures. For instance, pathogens are transmitted between hosts during epidemics [1], and the gut microbiota [2] and the soil microbiota [3] inhabit environments with complex spatial structures. A simple and generic model of spatially structured populations considers well-mixed demes (i.e. subpopulations) connected by migrations [4–16]. Spatial structure can impact the way populations evolve, in particular the probability that a mutant fixes, i.e. takes over, in the population. If migrations are symmetric enough to maintain the overall mutant fraction, the fixation probability of a mutant is not impacted by spatial structure [17, 18], unless deme extinctions occur [19, 20], and while fixation time can be impacted [21]. However, more complex spatial structures on graphs can impact mutant fixation probability [22, 23]. A graph is said to amplify natural selection if it enhances the fixation probability of beneficial mutants and reduces that of deleterious ones, compared to a well-mixed population with the same size. A graph is said to suppress natural selection in the opposite case. In particular, the star graph can amplify or suppress natural selection [24–26]. In models with exactly one individual per node of the graph [22], this depends on the update rule that is chosen to couple birth and death on the graph [24–26]. This

extends to models with demes where migration is coupled to birth and death to exactly preserve deme size [10, 14, 15]. Models of structured populations allowing to decouple growth from migrations were recently developed [13, 16]. In these models, when migrations between demes on the nodes of a graph are rare enough not to affect mutant fixation or extinction within a deme, the asymmetry of migrations determines whether the star amplifies or suppresses natural selection [13]. Meanwhile, when migrations become more frequent, suppression of selection becomes pervasive in graphs with asymmetric migrations [16].

Given the impact of spatial structure on the fate of one mutation, it is interesting to ask more broadly how spatial structure impacts the overall evolution of a population. In a constant environment, the evolution of a population can be described as an exploration of its fitness landscape, which maps each genotype to a fitness, representing reproductive success [27, 28]. Fitness landscapes are often rugged and possess multiple peaks (i.e., local maxima, corresponding to genotypes that are locally best adapted) [29–37], due to interactions between genes known as epistasis [36, 38–40]. Such topographical properties of fitness landscapes affect the predictability of evolution [35, 36, 40]. When mutations are rare, a mutation fixes or goes extinct before a new one occurs. Hence, most of the time, only one genotype, called the wild type, exists in the population [41, 42]. When the lineage of a mutant fixes, it becomes the new wild type and the population moves to another point of the fitness landscape. Evolution on a fitness landscape under rare mutations can thus be viewed as a biased random walk, where the fixation of a mutation constitutes an elementary step [43]. In the strong selection regime [43, 44], which is appropriate for large enough populations [45], only beneficial mutations can fix and the population can only go uphill in the fitness landscape. In well-mixed populations, these trajectories, known as adaptive walks [43], have been extensively studied [42, 46–50]. However, in smaller populations, neutral or deleterious mutations may fix [51, 52], due to strong genetic drift [53]. We recently showed that finite population size can foster the early adaptation of a well-mixed population in a rugged fitness landscape [54]. What is the impact of spatial structure on the way populations explore their fitness landscape in the rare mutation regime? Historically, the impact of spatial structure on adaptation on rugged fitness landscapes was at the core of Wright's shifting balance theory [4, 27]. Here, we aim to make a quantitative study of the impact of spatial structure, without deme extinction and re-colonization ingredients, which were important in Wright's arguments. Because demes can have small sizes, finite-size effects [54] may play important roles in the exploration of fitness landscapes by spatially structured populations. Recent studies have considered the impact of spatial structure on the exploration of rugged fitness landscapes, but in the case of populations with one individual per node of the graph [55–57]. As for mutant fixation, their results are strongly impacted by the choice of the update rule.

Here, we investigate the exploration of rugged fitness landscapes by spatially structured populations with demes on the nodes of a graph, connected by rare migrations. We find that the number of demes in the structure impacts early adaptation on rugged fitness landscapes. In particular, in some landscapes, the mean fitness of the first encountered peak starting from a random genotype reaches a maximum for a finite number of demes. This means that the early adaptation of these finite structures is more efficient than that of very large ones. However, the time it takes to reach the first fitness peak is then increased. Furthermore, we show that, in most landscapes, migration asymmetries associated with suppression of selection make early adaptation more efficient, and enhance finite-size effects. We extend our study to frequent migrations, and our results suggest that these conclusions also hold in this regime where suppression of selection is pervasive. We also investigate the impact of spatial structure on long-term evolution by studying the steady state of the population. For this, we define an effective population size for the steady-state distribution. We find that suppression of selection

is associated to reduced steady-state effective population sizes, and thus to small mean fitnesses at steady state.

Model and methods

Modeling evolution in spatially structured populations

Genotypes and mutations. We consider populations of asexual haploid individuals, e.g. bacteria. We assume that each genotype possesses a fixed fitness, which defines a fitness landscape [27, 28]. Note that we do not address environment variations or frequency-dependent selection [36]. Genotypes are described by sequences of L binary variables, taking values 0 or 1, which represent genetic units, such as nucleotides, amino acids or genes. Considering genotypes as strings of binary variables is a simplification [58] that we make for simplicity, in line with many studies [32, 40]. Mutations are modelled as substitutions from 0 to 1 or vice versa at one site. In this framework, genotype space is a hypercube with 2^L vertices, each of them having L neighboring genotypes that are accessible by a single mutation. We assume that mutations appear in a deme chosen uniformly at random, which is valid both if they occur randomly upon division and if they are due to a stress that is uniform through the structure.

Evolution in the weak mutation regime. We focus on the weak mutation regime, defined by $\mathcal{N}\mu \ll 1$, where $\mathcal{N}$ denotes total population size and μ denotes mutation rate (i.e., mutation probability per generation). In this regime, mutations are sufficiently rare for the fate of a mutation (either fixation or extinction) in the whole structured population to be set before any new mutation occurs. Thus, the population almost always comprises a single genotype. Note that phenomena like clonal interference [59–62] are thus not described. When a mutant genotype takes over (i.e., fixes), it becomes the new wild type. The evolution of the population can be modelled as a biased random walk in genotype space [43], and is driven by random mutations, natural selection and genetic drift. A mutation that fixes yields one step of this random walk, whereby the population hops from one genotype to another in the hypercube representing genotype space. This description is known as the origin-fixation approach [52].

Modeling spatially structured populations. We consider spatially structured populations on graphs, using the model introduced in Ref. [13]. This model considers D well-mixed demes (i.e. subpopulations), each placed on one node of a graph, and migrations along the edges of this graph. All demes have the same carrying capacity C. Each genotype i is characterized by its fitness f_i, and individuals with genotype i have a division rate $f_i(1 - N/C)$ where N denotes deme size. All individuals have the same death rate g. Thus, we consider selection on birth. Disregarding migrations, the steady-state size of a deme is $C(1 - g/f_i)$. We focus on the regime where deme size fluctuates weakly around this steady-state value. Migration events, where an individual moves from one deme to another, occur with rate m_{ab} from deme a to deme b. As in Ref. [13], we focus on the rare migration regime, where each mutation either gets extinct or fixes within a deme before any migration event. We then briefly extend our work to more frequent migrations, using the model introduced in [16]. We consider highly symmetric graphs, illustrated in Fig. 1.

Mutant fixation within a deme. In the rare migration limit, the first step of mutant fixation is its fixation in the deme where it initially appeared. Then, the mutation may spread to other demes by migrations. Note that we assume that mutations are more rare than migrations, so that all this process happens before another mutation arises. As a deme is well-mixed, with

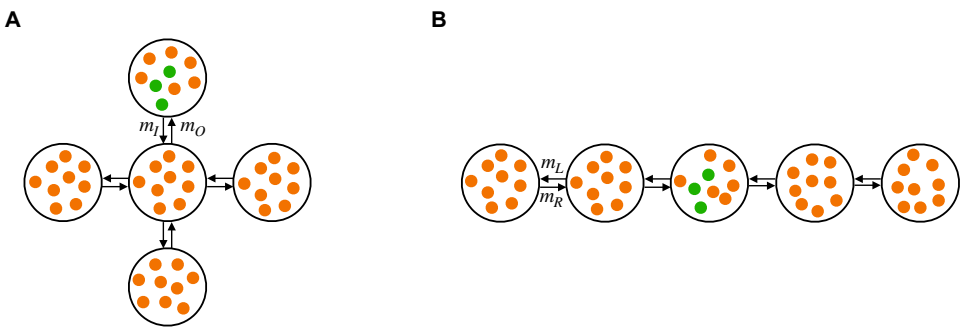

Figure 1: **Spatial structures investigated.** (A) Star; (B) Line. One well-mixed deme is placed on each node of a graph. Markers represent individuals (wild-type ones in orange, mutants in green). Demes are connected by migrations along the edges of the graphs, represented by arrows. In the star, the incoming and outgoing migration rates to and from the center are denoted by m_I and m_O, respectively. In the line, the left and right migration rates are denoted by m_L and m_R, respectively.

a size that weakly fluctuates around a steady-state value (see above), we approximate fixation within a deme by that in a population of constant size equal to its equilibrium size, see [13]. For completeness, we consider both the Moran and the Wright-Fisher models within a deme.

In the Moran model, the fixation probability ρ_{ij} of the lineage of a mutant with genotype j in a population where all other individuals are wild-types with genotype i reads [53, 63]:

$$\rho_{ij} = \frac{1-(f_i/f_j)}{1-(f_i/f_j)^{N_i}} = \frac{1-(1+s_{ij})^{-1}}{1-(1+s_{ij})^{-N_i}}\,, \tag{1}$$

where $s_{ij} = f_j/f_i - 1$ is the relative fitness advantage of the mutant and $N_i = C(1-g/f_i)$ is the steady-state size of the deme composed of N_i individuals.

In the Wright-Fisher model under the diffusion approximation, we have [64, 65]:

$$\rho_{ij} = \frac{1-e^{-2s_{ij}}}{1-e^{-2N_i s_{ij}}}\,. \tag{2}$$

Circulations. Circulations are graphs where, for each deme, the total incoming migration rate is equal to the total outgoing migration rate [13, 16, 22]. One example is the clique, or fully connected graph, or island model [4], where each deme is connected to all others and the migration rate is the same between any two demes. In a circulation in the rare migration regime, the fixation probability of the lineage of a mutant with genotype j in a population where all other individuals are wild-types with genotype i reads [13]:

$$\phi_{ij} = \rho_{ij}\frac{1-\gamma}{1-\gamma^D}, \tag{3}$$

where we introduced $\gamma = N_i\rho_{ji}/(N_j\rho_{ij})$ with $N_j = C(1-g/f_j)$, and where ρ_{ij} denotes the fixation probability of type j in a deme full of individuals of type i (see paragraph above).

Note that the fact that the fixation probability is the same for all circulation graphs extends beyond the rare migration regime [16]. This is consistent with Maruyama's theorem [17, 18] and with the circulation theorem in models with one individual per node of the graph [22]. Furthermore, the fixation probability in Eq. (3) is almost the same as that of a well-mixed population with the same total size [13]. In other words, for circulations, spatial structure has almost no effect on mutant fixation probability.

Star. In a star in the rare migration regime, the fixation probability of one mutant appearing in a deme chosen uniformly at random is [13]:

$$\phi_{ij} = \rho_{ij} \frac{(1-\gamma^2)[\gamma + \alpha D + \gamma\alpha^2(D-1)]}{D(\alpha+\gamma)[1+\alpha\gamma - \gamma^D(\alpha+\gamma)^{2-D}(1+\alpha\gamma)^{D-1}]}, \tag{4}$$

where α represents migration asymmetry, and is defined as the ratio $\alpha = m_I/m_O$ of the incoming migration rate m_I from a leaf to the center to the outgoing migration rate m_O from the center to a leaf (see Fig. 1). The star is a suppressor of selection when $\alpha < 1$, and an amplifier when $\alpha > 1$ [13]. As a reminder, in a suppressor (resp. amplifier) of selection, the fixation probability of beneficial mutations are smaller (resp. larger) than in a well-mixed population with same size, and the opposite holds for deleterious mutations. When $\alpha = 1$, the star is a circulation, with fixation probability given by Eq. (3) [13, 16].

Line. Let us now turn to a line of demes, which corresponds to the linear stepping stone model [5]. Migration rates to the left and to the right can differ, see Fig. 1. In Section 1 of the Supplementary Material, we derive the fixation probability of one mutant appearing in a deme chosen uniformly at random within a line of demes. We obtain:

$$\begin{aligned}\phi_{ij} = \frac{\rho_{ij}}{D}\Bigg\{ & \frac{1-\gamma/\alpha}{1-(\gamma/\alpha)^D} + \frac{1-\alpha\gamma}{1-(\alpha\gamma)^D} \\ & + \frac{4}{D^2(1+\alpha)(1+\gamma)} \sum_{k=1}^{D-2}\sum_{n=1}^{D-1} \gamma^{1-\frac{n}{2}} \left[\frac{1-(\alpha\gamma)^n}{1-(\alpha\gamma)^D}\alpha^{k+1-\frac{n}{2}} + \frac{1-\left(\frac{\gamma}{\alpha}\right)^n}{1-\left(\frac{\gamma}{\alpha}\right)^D}\alpha^{\frac{n}{2}-k} \right] \left(t_{k,k+1}^{(1,n)} - t_{k,k+1}^{(n,1)} \right) \Bigg\},\end{aligned} \tag{5}$$

with

$$t_{k,k+1}^{(a,b)} = \sum_{r=1}^{D-1}\sum_{s=1}^{D-1} \frac{\sin\left(\frac{kr\pi}{D}\right)\sin\left(\frac{ar\pi}{D}\right)\sin\left(\frac{bs\pi}{D}\right)\sin\left(\frac{(k+1)s\pi}{D}\right)}{1 - \frac{2\sqrt{\alpha\gamma}}{(1+\alpha)(1+\gamma)}\left[\cos\left(\frac{r\pi}{D}\right) + \cos\left(\frac{s\pi}{D}\right)\right]}. \tag{6}$$

Notations are as above, except that the migration asymmetry is defined as $\alpha = m_R/m_L$, where m_R (resp. m_L) is the migration rate from a deme to its right (resp. left) neighbor. Like the star, when $\alpha = 1$, the line becomes a circulation, with fixation probability given by Eq. (3). The more α deviates from 1, the more the line suppresses natural selection, see Fig. S3. Note that by symmetry, the fixation probability is the same for asymmetries α and $1/\alpha$.

Beyond rare migrations. To go beyond the rare migration regime without artifacts, we use the serial dilution model introduced in [16], where demes undergo phases of local exponential growth followed by events of dilution and migration, modeling serial passage with spatial structure. Specifically, we use multinomial sampling to perform dilution and migration events, and we choose the convention where all demes contribute by C to the next bottleneck (see the Supplementary Material of [16] for details). We take a growth time $t = 1$ to have an equivalence with the structured Wright-Fisher model [16]. Then, under rare migrations, this model gives results that are fully consistent with that of [13], with the within-deme fixation probability given by Eq. (2) instead of Eq. (1), see Ref. [16]. Analytical expressions of the fixation probability beyond the rare migration regime are only available within the branching process approximation, which holds for slightly beneficial mutants [16]. Here, we aim to include neutral and deleterious ones. Hence, we estimate the fixation probability for the specific f_i and f_j in our fitness landscapes by performing simulations of the serial dilution model.

Quantifying short-term and long-term adaptation of spatially structured populations

Exploration of fitness landscapes by spatially structured populations. Here, we investigate the biased random walk performed in a fitness landscape by structured populations in the rare migration regime. For a star, the fixation probability of a mutant in the whole population is given by Eq. (4), combined with Eq. (1) if the Moran model is used to describe within-deme fixation (resp. with Eq. (2) if the Wright-Fisher model is used). Let us call these walks a Moran star walk and a Wright-Fisher star walk, respectively. Let us use analogous names for the line: Moran line walk and Wright-Fisher line walk. Recall that in Ref. [54], we studied the Moran and the Wright-Fisher walks in well-mixed populations.

As for well-mixed populations [54], the walks performed by structured populations on fitness landscapes are discrete-time Markov chains, where each time step corresponds to a mutation event. They are irreducible, aperiodic and positive recurrent (and thus ergodic). Hence, they possess a unique steady-state distribution, describing the probability that the population features a given genotype at steady state, towards which they converge for any initial condition [66, 67]. They are also reversible [68, 69]. Note however that the Wright-Fisher origin-fixation dynamics beyond the diffusion approximation (which we do not consider here) is not reversible [70].

In Ref. [54], we addressed the effect of finite population size N on the exploration of fitness landscapes by well-mixed populations. Here, we will mainly study the impact of the number D of demes, and of the migration asymmetry α, which characterize spatial structure and have a significant impact on mutant fixation probability in the star and in the line. Indeed, in the rare migration regime, the star can amplify or suppress natural selection depending on the value of α (see above and Ref. [13]), and the intensity of amplification depends on D [13, 22]. We will also investigate the effect of varying the carrying capacity C of each deme, which impacts the amplitude of fluctuations within each deme.

Characterizing early adaptation. In order to investigate the short-term evolution of a population, we will focus on the height h of a walk, which is the fitness of the first encountered peak [42], as in Ref. [54]. We consider its average $\bar{h}$ over initial genotypes chosen uniformly at random. It provides a global characterization of early adaptation by a structured population in the fitness landscape. Note that starting from a random genotype may be relevant to describe evolution after an environmental change, which makes the population ill-adapted to the current environment. To assess the variability of h, we also investigate its standard deviation σ_h over replicates, as well as the standard deviation $\overline{\sigma_{h_i}}$ over replicates of the height h_i starting from a specific initial genotype i, then averaged over i. Furthermore, for ensembles of landscapes in a given class, we study the corresponding ensemble mean heights $\langle \bar{h} \rangle$.

In addition to the height h of a walk, we examine its length $\bar{\ell}$ and time $\bar{t}$, which are defined respectively as the mean number of mutation fixation events, and of mutation events (leading to fixation or not), before the first fitness peak is reached.

We obtain $\bar{h}$, $\bar{\ell}$ and $\bar{t}$ by two different methods: by numerically solving the equations from a first step analysis (FSA), and by numerical simulations. For details on the FSA, see Ref. [54] (section 2(b) of the main text and section S2 of the Supplementary Material). Numerical simulations are performed using a Monte Carlo procedure. Note that we mainly simulate the embedded versions of the walks (i.e., we normalize the transition probabilities to all neighbors so that they sum to one, thus avoiding rejected moves) in order to save time and resources, except when computing the time t of a walk.

Characterizing steady state. We are also interested in the impact of spatial structure on the long-term evolution of structured populations. As a first step to characterize it, we investigate the steady-state distribution associated to our walks on fitness landscapes. It provides the probability that the population features a given genotype after a sufficiently long time of evolution. Specifically, we define the steady-state effective population size of the population, as the size of a well-mixed population with the same steady-state distribution.

Fitness landscapes

We investigate how spatially structured populations explore various fitness landscapes, both model ones and experimental ones.

LK fitness landscapes. In the LK model (originally called NK model [71]), the fitness of a genotype $\vec{\sigma} = (\sigma_1, \sigma_2, ..., \sigma_L) \in \{0,1\}^L$ is expressed as:

$$f(\vec{\sigma}) = \sum_{i=1}^{L} f_i \left(\{\sigma_j\}_{j \in \nu_i} \right) , \tag{7}$$

where f_i is the contribution of site i, and ν_i comprises i and its epistatic partners, which are the sites that interact with i. In this model, the first parameter L is the number of (binary) units of the genome, i.e. the dimension of the hypercube of genotypes. The second parameter K is the number of epistatic partners of each site [71]. Thus, ν_i comprises $K+1$ elements, for all i. For each site, we consider that its epistatic partners are chosen uniformly at random among other sites: each site is said to have a 'random neighbourhood' [42]. Increasing K increases epistatis, and yields more rugged landscapes. If $K = L-1$, the LK model reduces to the House of Cards model [72, 73] where the fitness values of each genotype are independent and identically distributed. Indeed, then, any mutation affects the contribution of all sites in Eq. (7). Conversely, when $K = 0$, landscapes are additive and display a single peak. The LK fitness landscape model allows for easy tuning of ruggedness. Note that we draw the f_i in a uniform distribution between 0 and 1.

Other fitness landscapes. We consider two generalizations of the LK model: the LK model with non-binary genotype states [58], and a modified LKp model [74] where the contributions f_i in Eq. (7) have a probability p to be set to a positive constant value to include neutral mutations (the constant is zero in the classic version; our modification avoids zero overall fitnesses). We also investigate the Rough Mount Fuji model [32, 75], the Eggbox model [76], the Ising model [76, 77], and tradeoff-induced landscapes [78, 79]. Details about these models are given in the Supplementary Material of Ref. [54]. Finally, we also consider experimentally measured landscapes from [80] and [81].

Code availability. Our code is freely available at
`https://github.com/Bitbol-Lab/fitness-landscapes/`.

Results

Early adaptation on small LK rugged landscapes

Small rugged landscapes. In Fig. 2, we study the height of biased random walks performed by star-structured populations on small LK landscapes with $L = 3$ and $K = 1$. We consider an ensemble of 2×10^5 such landscapes in Fig. 2(A,E). We also give particular attention to

three specific landscapes from this ensemble, denoted by A, B and C. These fitness landscapes are depicted in Fig. 2(I-K), their MAGELLAN representations [82] are shown in Fig. S4, and all their fitness values are provided in Table S1. Each of these landscapes comprises two fitness peaks. Landscapes A and B were randomly chosen among the landscapes with more than one peak obtained when generating LK landscapes. The adaptation of well-mixed populations on these two landscapes was studied in Ref. [54], which facilitates comparison. Landscape C was chosen because it comprises a very high fitness peak (genotype 111) and a much lower one (genotype 000), which is reminiscent of a (rough) Mount Fuji landscape [32, 75], see Fig. 2(K).

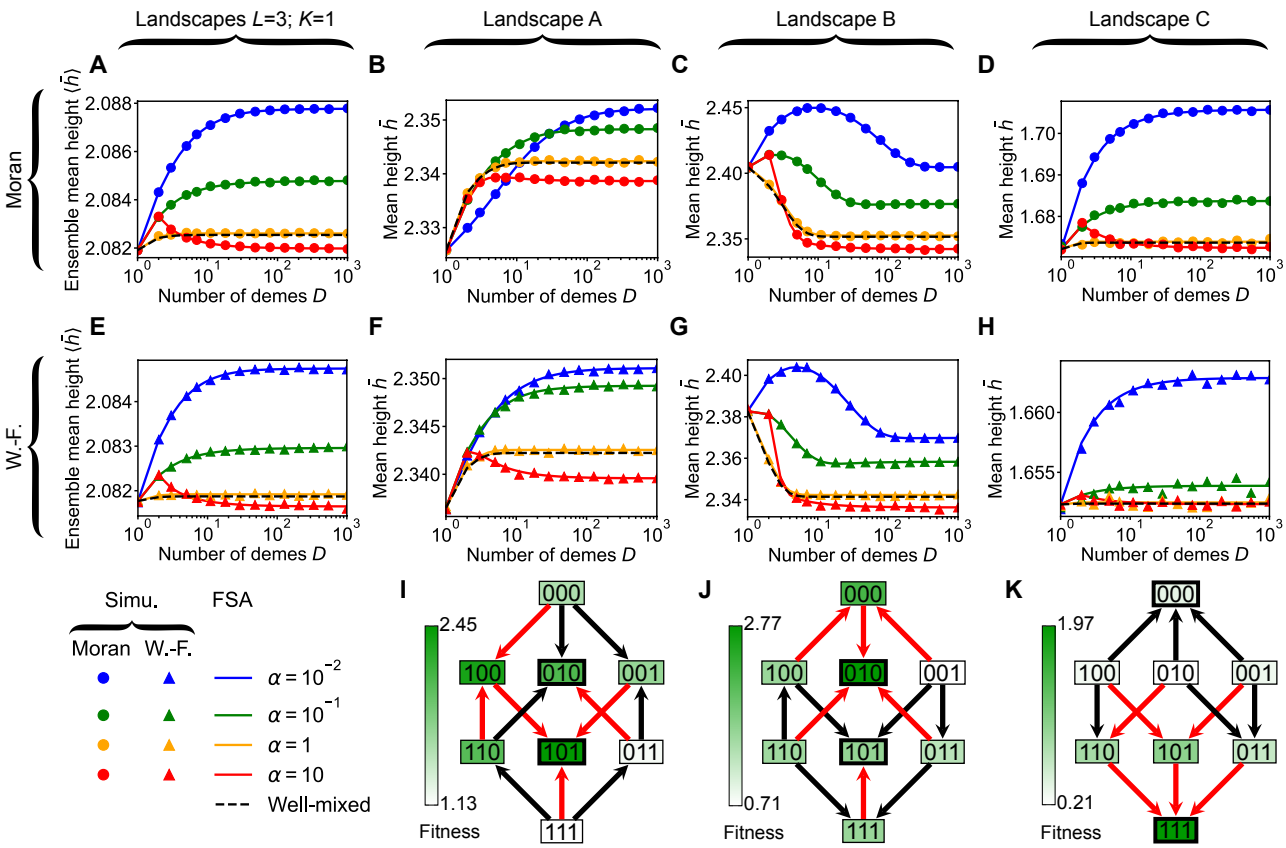

Figure 2: **Impact of the number of demes D and migration asymmetry α of a star on early adaptation in LK landscapes with $L = 3$ and $K = 1$.** (A) Ensemble mean height $\langle \bar{h} \rangle$ of the first fitness peak reached when starting from a uniformly chosen initial genotype versus the number of demes D for the Moran star walk, with various values of migration asymmetry α. Lines: numerical resolutions of the FSA equations for each landscape; markers: simulation results averaged over 10 walks per starting genotype in each landscape. In both cases, the ensemble average is performed over 2×10^5 landscapes. Dashed black line: well-mixed case with population size $DC(1 - g/f_W)$ where f_W is the fitness of the wild type, shown as reference. (B-D) Mean height $\bar{h}$ of the Moran star walk versus D in landscapes A, B and C, depicted in (I-K). (E-H) Same as (A-D) but for the Wright-Fisher star walk. Simulation results in landscapes A, B and C are averaged over at least 10^5 walks per starting genotype. (I-K) Representations of fitness landscapes A, B and C. Genotypes with thick edges are peaks; arrows point towards fitter neighbors; red arrows point toward fittest neighbors. Parameter values (all panels): $C = 20$ and $g = 0.01$.

Impact of the number of demes D on the height h. To assess the impact of spatial structure on early adaptation, we first vary the number of demes D in the star, for a fixed deme carrying capacity C. In Fig. 2, we observe that the way $\bar{h}$ varies with D strongly depends

on migration asymmetry α. This shows that the star structure has a notable impact on early adaptation in rugged landscapes. For $\alpha = 1$, i.e. when the star is a circulation, we obtain results that are very close to those for the well-mixed population with the same total size. Indeed, the mutant fixation probabilities in circulations are extremely similar to those in well-mixed populations [13].

While $\bar{h}$ often increases with D, there are cases where it is non-monotonic. For example, in landscape B, $\bar{h}$ versus D displays a broad maximum for $\alpha = 10^{-2}$, see Fig. 2(C,G). Then, there is a finite number of demes D that yields a more efficient overall search for high fitness peaks than the large-D limit. This is reminiscent of the maximum of $\bar{h}$ we observed for finite well-mixed populations in this landscape in Ref. [54]. However, this maximum is obtained for a larger total size here than in the well-mixed case. This suggests that population structure can increase the range of population sizes where finite size effects matter. Note that the value of D giving this maximum is slightly smaller for the Wright-Fisher star walk than for the Moran star walk. As in well-mixed populations, this comes from a factor of 2 that differs in the Moran and Wright-Fisher fixation probability formulas in the diffusion regime [54].

Fig. S5(I,M) shows the mean height $\overline{h_i}$ of the walk for each starting genotype i for $\alpha = 10^{-2}$. It reveals that the presence of the maximum of $\bar{h}$ versus D in landscape B for $\alpha = 10^{-2}$ is due to the fact that, starting from genotype 111, the transition towards the small peak becomes more favored when increasing D. Indeed, $\overline{h_{111}}$ decreases with D. Together with the constant or increasing trends of the other $\overline{h_i}$, which plateau for smaller D, this gives rise to the maximum observed when averaging over i. For larger α, $\overline{h_{111}}$ also decreases, but plateaus for smaller values of D, see Fig. S5(I-P). These observations can be explained as follows. Genotype 111 has neighbors with similar fitnesses, but only the mutation to the small peak is beneficial, see Fig. 2(J). As D increases, the walk becomes more biased toward beneficial mutations, and this transition is more favored. However, in the star, small α values lead to suppression of selection (see Fig. S6), and this transition then occurs for larger D.

Fig. S7 further shows that the standard deviation of h, starting from uniformly chosen genotypes, essentially mirrors $\bar{h}$ when plotted versus D. This holds for all values of α, and is in line with our observations for well-mixed populations [54]. Hence, the maximum of $\bar{h}$ observed for $\alpha = 10^{-2}$ in landscape B is associated with a higher predictability. Finally, Fig. S8 shows that the mean standard deviation $\overline{\sigma_{h_i}}$ (averaged over the starting points i), mainly decreases with D, which is consistent with the idea that larger values of D yield more biased walks and more reproducibility.

Beyond these rich specific features, in Fig. 2, we notice in all landscapes that large-D plateaus are reached for larger values of D when α is smaller. Again, in the star, small α leads to suppression of selection. This enhances finite-size effects, since larger values of D are then needed to obtain significant bias toward more beneficial mutations.

Impact of the migration asymmetry α on the height h. In Fig. 2(A,E), we observe that, for moderate to large D, the ensemble mean height $\langle \bar{h} \rangle$ in the star is larger when migration asymmetry α is smaller. Thus, smaller values of α make early adaptation more efficient for a star-structured population, on average. The same observation holds for the mean height $\bar{h}$ in each of our three example landscapes, see Fig. 2(B-D) and (F-H). Fig. S9 shows these heights versus α for $D = 5$, and confirms that the highest values are obtained for intermediate α values satisfying $\alpha < 1$. This is a striking observation: the star suppresses natural selection for $\alpha < 1$, thus making beneficial mutations less likely to fix, and this is precisely where we find that the highest peak is most likely to be reached first.

To understand the impact of α in more detail, let us examine how it affects the probability of fixation of available mutations. Fig. S6 shows the probability of fixation of a mutant versus its relative selective advantage s for different values of migration asymmetry α in the star, high-

lighting the values of s that exist in the three landscapes considered. Small α gives suppression of selection, and the range of small-effect mutations that are effectively neutral increases when α is decreased. Several mutations in the fitness landscapes considered fall in this range for $\alpha = 10^{-2}$ (see Fig. S6), and their fixation is thus expected to be particularly impacted.

As a concrete example, Fig. S10 focuses on the three transitions accessible from genotype 110 within landscape A, which are two beneficial mutations, one with a larger s than the other, and one deleterious mutation. Fixation of the latter can be neglected when considering large values of D (Fig. S10(A,B) and (E,F) use $D = 10^3$). Then, the key question is which of the two beneficial mutations fixes. Fig. S10(A,B) shows that the fixation probability of each of them increases when α increases, as the star shifts from suppression to amplification of selection. However, this increase requires larger α values for the least beneficial mutation of the two, which is more subject to suppression (see Fig. S10(A,B) and Fig. S6). To see how the choice between the two mutations is impacted, we normalize these two fixation probabilities by their sum for each α. Fig. S10(E,F) shows that the resulting transition probabilities vary non-monotonically with α, with the most beneficial mutation being most favored for $\alpha \approx 10^{-2}$, as the weakly beneficial one is then still strongly suppressed. Further increasing α reduces the bias toward this most beneficial mutation. In this example, the most beneficial mutation yields genotype 100, from which the only beneficial mutation leads to the highest peak (genotype 101). Thus, the more favored this transition, the larger $\bar{h}$. This is one of the reasons why the mean height $\bar{h}$ is larger as α is decreased down to 10^{-2} in landscape A for large D, see Fig. 2(B,F) and Fig. S5. Favoring the most beneficial mutations relative to the other ones often promotes reaching higher peaks. This rationalizes our striking observation that small α values, giving suppression of selection in the star, generally make early adaptation more efficient for large D, see Fig. 2. Among beneficial mutations, suppression of selection most strongly suppresses the fixation of weakly beneficial ones, giving a relative advantage to strongly beneficial ones.

Note that varying α can have various impacts. In our previous example, decreasing α below 10^{-2} reduces the bias toward the most beneficial mutation, as its fixation also becomes suppressed, see Fig. S10(A,B) and (E,F). Accordingly, Fig. S9 shows that the most efficient early adaptation is generally obtained for small but intermediate α values. Besides, for each genotype, the values of α that impact the bias between possible transitions depend on the fitness values of neighboring genotypes (and on the values of C and D). In particular, we argued that favoring the most beneficial mutations often promotes reaching higher peaks, but there are fitness landscapes in which this does not hold. For instance, Fig. S11(B,H) features a House of Cards landscape where decreasing α reduces $\bar{h}$ for large D. Detailed analysis reveals that, in this landscape, a mutation leading to the highest peak has a relatively small fitness advantage, and smaller α values yield biases against it.

Line structure. While we focused on the star so far, it is interesting to assess the impact of the number of demes D and of migration asymmetry α in other structures. Fig. 3 shows that the main conclusions obtained for the star also hold for the line, which is represented in Fig. 1(B). Namely, the height $\bar{h}$ of the first peak reached usually increases with D, with exceptions, and for large D, small α usually leads to larger $\bar{h}$, with exceptions too. Recall that, for the line, there is suppression of selection for all $\alpha \neq 1$, and a symmetry such that α and $1/\alpha$ are equivalent (see Section 1 of the Supplementary Material) – which is why the same results are obtained for $\alpha = 10$ and $\alpha = 10^{-1}$.

In landscape A, we observe that $\bar{h}$ is smaller for $\alpha = 10^{-2}$ than for $\alpha = 10^{-1}$ or $\alpha = 1$ when considering the Moran line walk (Fig. 3(A)), while the opposite holds when considering the Wright-Fisher line walk (Fig. 3(D)). Fig. S10(C,D) shows the fixation probabilities of the two beneficial mutations available from genotype 110, this time in the line. They are largest for $\alpha = 1$, as the line suppresses selection for all $\alpha \neq 1$. Furthermore, the decay of the fixation

probability when α becomes larger or smaller than 1 occurs at more extreme α values for the most beneficial mutation. Normalizing these two fixation probabilities by their sum thus yields a non-monotonic dependence on α, see Fig. S10(G,H). For the Moran line walk, there is far less bias against the least beneficial transition among the two when $\alpha = 10^{-2}$ than when $\alpha = 10^{-1}$. This least beneficial mutation leads to the small peak (010), explaining the results in Fig. 3(A).

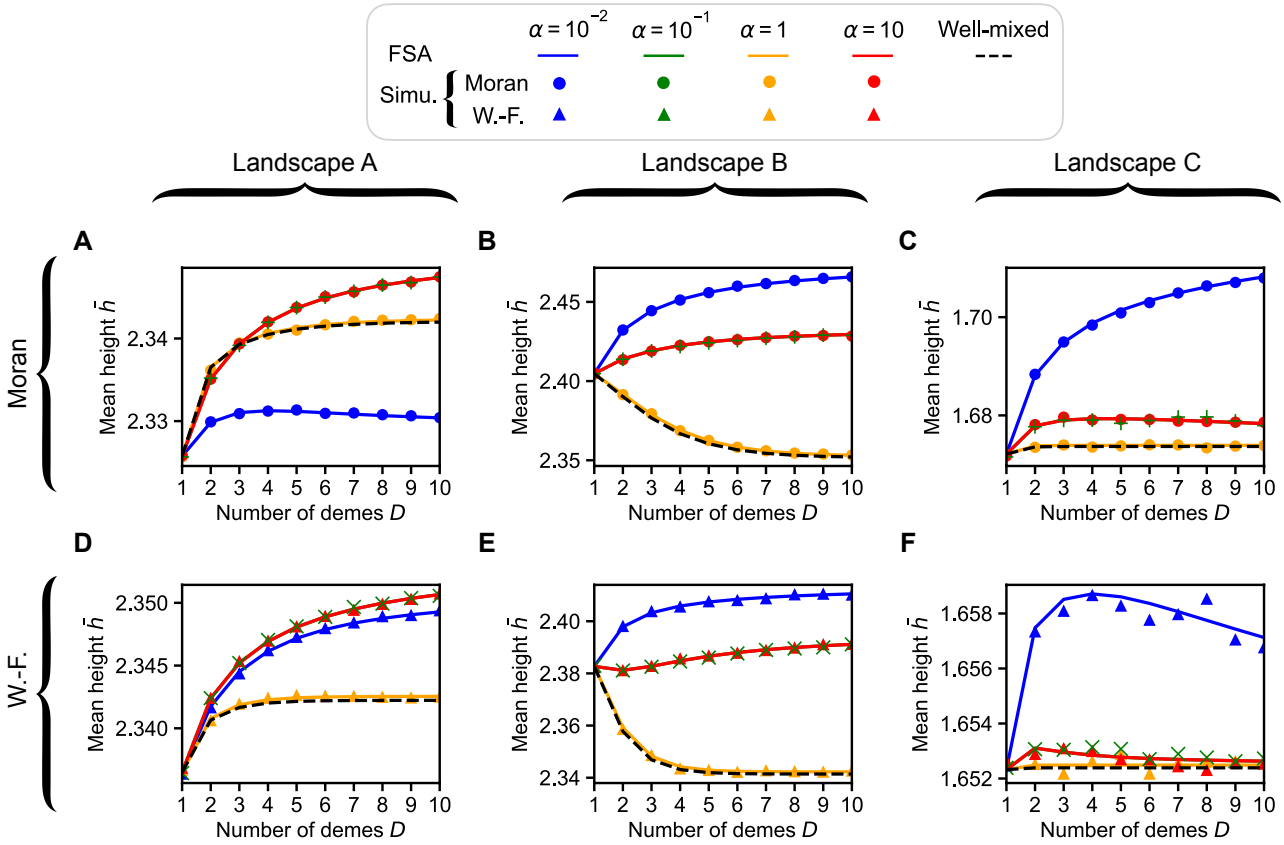

Figure 3: **Impact of the number of demes D and migration asymmetry α of a line on early adaptation in 3 LK landscapes with $L = 3$ and $K = 1$.** (A-F) Mean height $\bar{h}$ of the first fitness peak reached when starting from a uniformly chosen initial genotype versus the number of demes D for the Moran line walk (A-C) and for the Wright-Fisher line walk (D-F) for various values of migration asymmetry α, in landscapes A (A,D), B (B,E), and C (C,F) (see Fig. 2(I-K)). Dashed black line: well-mixed case with population size $DC(1-g/f_W)$, where f_W is the fitness of the wild-type genotype, shown as reference. Lines: numerical resolutions of the FSA equations for each landscape; markers: simulation results averaged over at least 10^5 walks per starting genotype in each landscape. Parameter values (all panels): $C = 20$ and $g = 0.01$.

Impact of the number of demes D on the length ℓ and time t. Let us now consider the length ℓ of a walk, which represents the number of mutation fixation events that occur before a fitness peak is reached. For this, we turn back to the star structure. Fig. S12(A-B) shows that the ensemble mean length $\langle \bar{\ell} \rangle$ decreases when D increases. Indeed, larger values of D give a stronger bias towards the most beneficial mutations to the walk, so fewer steps are generally needed to reach a peak. Nevertheless, specific landscapes can lead to a non-monotonic (landscape A, see Fig. S12(C-D)) or increasing (landscape C, see Fig. S12(G-H)) behavior of $\bar{\ell}$ versus D. Recall that landscape C comprises a high peak (genotype 111) and a much lower peak (genotype 000), see Fig. 2(K). Consider a population that starts from one of the neighboring genotypes of 000. Increasing D increases the bias towards the fittest neighbors. This makes

the population more likely to mutate to a neighbor of the high peak 111, instead of hitting the low peak 000, which leads to an increase of both ℓ and h.

While ℓ describes the number of successful mutations, it is also interesting to consider the number t of all mutations that occurred before a peak is reached, whether they fixed or not. Fig. S13 shows that it increases with D. Indeed, increasing D reduces all fixation probabilities, which usually increases the number of unsuccessful mutations. Note however that, in landscape C, $\bar{t}$ decreases when D increases above 2 when $\alpha = 10$, see Fig. S13(G,H). As explained above, the specific features of this landscape entail that $\bar{\ell}$ and $\bar{t}$ have a similar dependence on D as $\bar{h}$. Another striking result is that smaller values of α lead to substantially larger values of $\bar{t}$, see Fig. S13. This is due to suppression of selection, which results in more mutations being effectively neutral and having small fixation probabilities. Thus, while suppression of selection often allows to reach higher peaks first, this tends to require more time.

Impact of the carrying capacity C. Fig. S14 shows the ensemble mean height $\langle \bar{h} \rangle$, and the mean height $\bar{h}$ in landscapes A, B and C, versus deme carrying capacity C for various values of α in the star, with $D = 5$. We observe that all curves converge to the limiting result for a large well-mixed population (dashed horizontal black lines). Indeed, when demes are large, deleterious mutations cannot fix within their deme of origin. In the rare migration regime, this entails that they cannot fix in the structured population, just as in the case of large well-mixed populations. Similarly, if a beneficial mutation fixes in the deme where it appeared, the wild type, which is less fit than the mutant, cannot re-invade this deme, and mutant fixation in the structured population is thus assured. Thus, for rare migrations and in the limit of large demes, the fate of a mutation in the whole population is entirely determined by its fate in its deme of origin. Therefore, in this limit, spatial structure does not impact mutant fixation probabilities or the mean height of a walk. Fig. S14 also shows that smaller values of α lead to a later convergence of $\langle \bar{h} \rangle$ and $\bar{h}$ versus C. Thus, once again, suppression of selection enhances finite-size effects.

Beyond the rare migration regime. So far, we considered the rare migration regime [13]. How does spatial structure impact early adaptation for more frequent migrations? To address this question, we start from the recent model developed in Ref. [16]. This model alternates phases of local within-deme growth and dilutions with migrations [16], modeling serial transfer with exchanges between demes [83, 84]. It is very close to a structured Wright-Fisher model. We checked that, for small migration rates in the star, our results from the serial dilution model are consistent with those we obtained above in the rare-migration regime, assuming the Wright-Fisher model in the diffusion approximation within demes, see Fig. S15.

In Fig. 4, we show the mean height $\bar{h}$ versus the number of demes D in landscapes A, B and C, for the star with migration asymmetry $\alpha = 4$, both with frequent migrations under the model of Ref. [16] and with our above-described Wright-Fisher star walk, which assumes rare migrations. For frequent migrations in the serial dilution model, Ref. [16] showed that all non-circulation graphs are strict suppressors of selection in the branching process approximation. This includes stars with migration asymmetry $\alpha > 1$, while they amplify selection in the rare-migration regime regime. In Fig. 4, we observe that $\bar{h}$ is higher for frequent migrations than in the rare migration regime, and that the value of D from which the curve plateaus is larger with frequent migrations. Both of these results match what we observed in the rare migration regimes for structures that suppress natural selection (namely the star with $\alpha < 1$ and the line with $\alpha \neq 1$).

In Fig. S16, we provide similar results as in Fig. 4, but for a larger carrying capacity C. Strikingly, the height $\bar{h}$ depends on D for frequent migrations in this case too, while this carrying capacity is in the regime where structure does not affect fixation probabilities or the mean height

$\bar{h}$ in the rare-migration regime (see paragraph above on the impact of C). This suggests that increasing the amount of migrations within spatially structured populations enhances finite size effects.

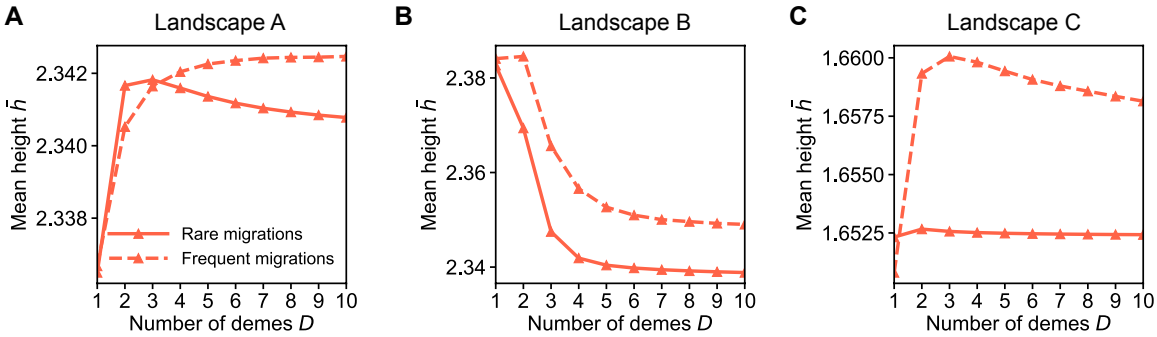

Figure 4: **Impact of the number of demes D of a star on early adaptation in three LK landscapes with $L = 3$ and $K = 1$ beyond the rare migration regime.** (A) Mean height $\bar{h}$ of the first fitness peak reached when starting from a uniformly chosen initial genotype of landscape A versus the number of demes D in a star with $\alpha = 4$, for frequent and rare migrations. (B-C) Same as (A) but in landscapes B and C. All values of $\bar{h}$ are obtained by numerically solving the FSA equations. With rare migrations, we use our analytical expressions of fixation probabilities (see "Model and methods"). With frequent migrations, we use fixation probabilities estimated from simulations of the serial dilution model (at least 10^7 replicates). In all cases, $C = 20$. For frequent migrations, $m_I = 0.2$ and $m_O = 0.05$, in the serial dilution model [16]. For rare migrations, $g = 0.01$. Markers are results and lines are guides for the eye.

Early adaptation in larger landscapes

So far, we focused on small fitness landscapes with $L = 3$ genetic units. It allowed us to link features of early adaptation to the probabilities of specific transitions in the fitness landscapes. However, real genotypes have many units (genes or nucleotides). It is thus important to extend our study to larger landscapes. For all this analysis, we focus on early adaptation in star-structured populations in the rare migration regime.

Early adaptation in star-structured populations in larger LK landscapes. First, we consider LK landscapes, and we analyze the impact of genome length L and number of epistatic partners K on early adaptation. Fig. 5(A,D) shows that the ensemble mean height $\langle \bar{h} \rangle$ reached in LK landscapes with $L = 20$ and $K = 1$ displays a maximum for a finite value of D for all values of α considered. This means that the most efficient early adaptation is obtained for stars with a finite number of demes, and not in the limit of very large stars. Recall however that demes need to be small enough for structure to matter in the rare migration regime (see above). In Fig. 5(B,E), we further observe that the position of this maximum shifts towards larger values of D when α decreases, i.e. when suppression of selection is stronger. This confirms that suppression of selection enhances finite-size effects. Finally, Fig. 5(C,F) shows that the size of the overshoot of $\langle \bar{h} \rangle$, with respect to its large-D limit, increases with L, and Fig. S17 shows that it decreases with K. These results are in line with those we obtained for well-mixed populations [54]. Thus, both in structured populations and in well-mixed ones, the adaptive advantage of some finite populations over large ones becomes stronger for larger genomes with few epistatic partners per genetic unit (gene or nucleotide). Importantly, this case is highly relevant in nature. For instance, pairwise epistasis has been shown to describe protein sequences well [85–87].

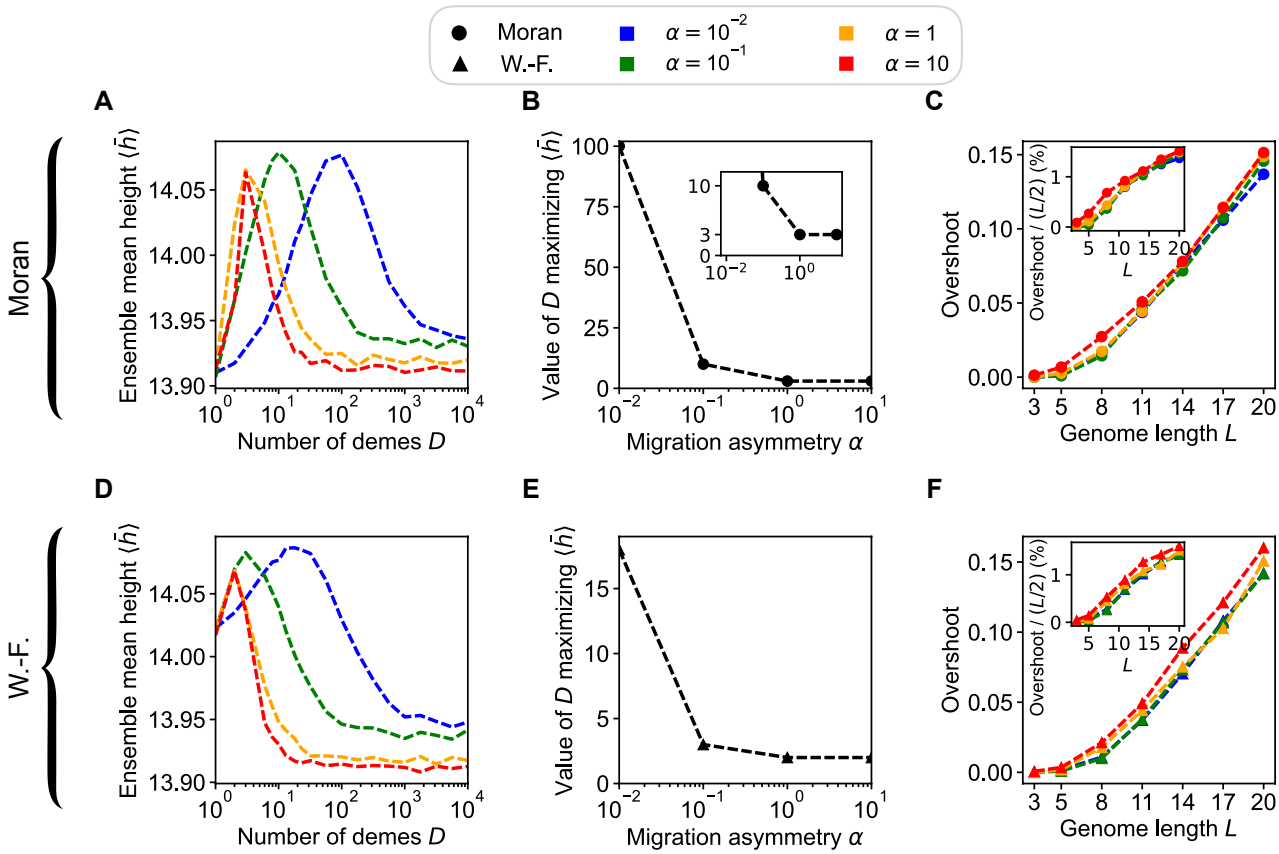

Figure 5: **Impact of genome dimension L on early adaptation in LK landscapes with $K = 1$.** (A) Ensemble mean height $\langle \bar{h} \rangle$ of the first fitness peak reached when starting from a uniformly chosen initial genotype in LK fitness landscapes with $L = 20$ and $K = 1$, versus the number D of demes for the Moran star walk with various values of migration asymmetry α. (B) Value of D maximizing the ensemble mean height $\langle \bar{h} \rangle$ of a Moran star walk versus migration asymmetry α, for LK fitness landscapes with $L = 20$ and $K = 1$. (C) Overshoot, defined as the maximum value of $\langle \bar{h} \rangle (D)$ minus the large-D limit of $\langle \bar{h} \rangle$ (evaluated at $D = 10^5$ in practice), versus L. Inset: same, but the overshoot is divided by $L/2$ and expressed in percentage. (D-F) Same as (A-C), but for the Wright-Fisher star walk. In all cases, dashed lines and markers represent simulation results (obtained using at least $\sim 10^5$ simulations per data point). Because of memory constraints associated to large genotypes space, landscapes were generated along the way and reinitialized for each walk, meaning that each walk takes place in a different landscape. Exception: the results for $L = 3$, were obtained in the 2×10^5 landscapes used in panels (A) and (E) of Fig. 2. Note that memory constraints are also the reason why FSA results are not shown here. Parameter values (all panels): $C = 20$ and $g = 0.01$.

Beyond the LK model. Let us now consider large fitness landscapes in other models than the LK model. Fig. S18 shows the rescaled mean height $\bar{h}$ of the first fitness peak reached versus the number D of demes in landscapes generated by various models for genomes of length $L = 6$. In most cases, smaller α yield larger values of $\bar{h}$ for large D. This is in line with our findings for small landscapes and for larger LK landscapes, and confirms their broad validity. As in the case of small landscapes, there are exceptions, for example the Rough Mount Fuji landscape of Fig. S18(E,L). Besides, as for small landscapes and for larger LK landscapes, we find that the value of D from which $\bar{h}$ plateaus becomes larger as α decreases, i.e. suppression of selection enhances finite-size effects. When $\alpha = 10^{-2}$, in two landscapes, $\bar{h}$ versus D displays a broad maximum (see Fig. S18(C,J) and (G,N), which respectively corresponds to a non-binary

LK landscape and an Ising landscape). Thus, in these landscapes, there is a finite value of D that makes the search for high fitness peaks more efficient.

Experimental fitness landscapes. Finally, we consider experimentally-measured fitness landscapes. Fig. 6 shows that the general features observed in model landscapes also hold for these experimental fitness landscapes. Namely, small α generally leads to larger values of $\bar{h}$ for large D, and enhances finite-size effects. Moreover, for experimental landscapes, we observe that the rescaled $\bar{h}$ features a large amplitude of variation when D is varied.

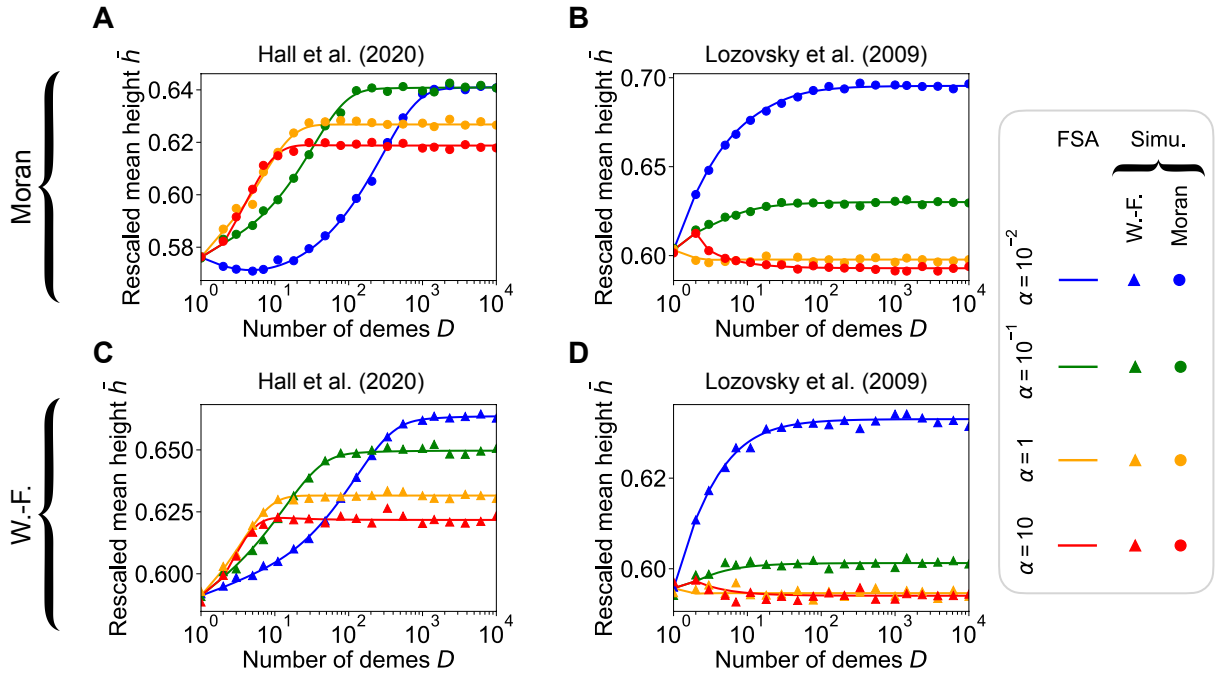

Figure 6: **Impact of the number of demes D and migration asymmetry α of a star on early adaptation in two experimental landscapes.** (A) Rescaled mean height $\bar{h}$ of the first fitness peak reached when starting from a uniformly chosen initial genotype versus the number D of demes for the Moran star, with various values of migration asymmetry α, in an experimental landscape from Ref. [80]. The rescaled mean height is defined as $(\bar{h} - h_{\min})/(h_{\max} - h_{\min})$, where $h_{\min}$ (resp. $h_{\max}$) denotes the fitness of the lowest peak (resp. highest peak). Note that $h_{\min} < \bar{h} < h_{\max}$. The fitness landscape corresponds to Fig. 2E of Ref. [80] ($L = 7$; organism: *Escherichia coli*; fitness proxy: half-maximal effective concentration EC_{50} using chloramphenicol). (B) Same as (A) but in the experimental landscape given in Table S2 from Ref. [81] ($L = 4$, organism: *E. coli* carrying dihydrofolate reductase from the malaria parasite *Plasmodium falciparum*; proxy for fitness: relative growth rate in the absence of pyrimethamine). (C-D) Same as (A-C) but for the Wright-Fisher star walk. In all panels, lines are numerical resolution of the FSA equations, while markers are simulation results averaged over at least 10^3 walks per starting genotype. Parameter values (all panels): $C = 20$ and $g = 0.01$.

Steady state and associated effective population sizes

So far, we investigated early adaptation by focusing on the first fitness peak reached. However, the long-term properties of adaptation are also of interest, for instance if the environment remains constant for a long time. Let us investigate the steady-state probability that a population possesses each given genotype in the long-term. These steady-state distributions were studied in well-mixed populations in Ref. [68]. How does spatial structure impact them?

We derive analytically the steady-state distribution of probabilities on a fitness landscape reached by a circulation, and by a star and a line in the limits $\alpha \to 0$ and $\alpha \to \infty$ in the rare migration regime. For this, we approximate the fixation probability within a deme by the Moran one for fixed deme size C, see Supplementary Material Section 2. For all these spatial structures, we obtain a steady-state probability distribution that has the same form as that associated to a well-mixed population in the Moran model [54, 68]. This motivates us to define a steady-state effective population size N_e as the size of the well-mixed population that possesses the same steady-state distribution as the structured population considered. Our analytical results for these steady-state effective population sizes are given in Table 1. For circulations, the steady-state effective population size is close to the actual population size. For a star with a very small α, as well as for a line with either a very small or very large α, featuring strong suppression of selection, the effective population size is the size of a deme. Conversely, when α is very large in the star, yielding amplification of selection, the effective population size is roughly twice the actual total population size.

Structure	Population size	Effective population size
Circulation	CD	$(C-1)D+1$
Star ($\alpha \to 0$)	CD	C
Star ($\alpha \to \infty$)	CD	$(C-1)(2D-3)+1$
Line ($\alpha \to 0$)	CD	C
Line ($\alpha \to \infty$)	CD	C

Table 1: **Steady-state effective population size of different structured populations.**

While our analytical approach only provides results for extreme values of α in the star and in the line, we also numerically estimate the steady-state effective population size N_e of these spatially structured populations for various values of α. Fig. S19 shows N_e versus α. The asymptotic values are in very good agreement with our analytical predictions for the star and for the line, and the result for $\alpha = 1$ is in good agreement with our analytical prediction for circulations (see Table 1). For the star, we observe that $\alpha > 1$ and amplification of selection are associated to steady-state effective population sizes larger than population size. The opposite holds when $\alpha < 1$ both in the star and in the line, which is associated to suppression of selection. We further observe in Fig. S19 that the range of values of α leading to intermediate effective population sizes depends on the landscape. In particular, both in the star and in the line, the transition between the small-α and the large-α regime is broader in landscape C than in landscapes A and B.

Our results show that suppression of selection is associated to reduced steady-state effective population sizes compared to actual sizes, while amplification is associated to larger ones. Smaller effective sizes yield smaller mean fitnesses at steady state, see [54]. This entails that suppression of selection is not beneficial to fitness on average on the very long term.

Discussion

In this paper, we investigated the impact of spatial structure on early adaptation in rugged fitness landscapes. We observed a rich dependence of early adaptation on the number of demes in the population, depending on the specific landscape considered, and on migration asymmetry. In particular, for some fitness landscapes, early adaptation is most efficient for a finite number of demes. Importantly, this tends to be the case for large genomes with few epistatic partners per site, which are relevant in practice. Furthermore, we showed that, in most landscapes, migration asymmetries associated with suppression of selection allow the population to reach

higher fitness peaks first when the number of demes is large. In this sense, and perhaps surprisingly, suppression of selection tends to make early adaptation more efficient. We showed that this can arise because suppression of selection increases the bias toward the most beneficial mutations. However, more time is then needed to reach this first fitness peak. We also found that migration asymmetries associated with suppression of selection enhance finite-size effects. Importantly, these observations apply to fitness landscapes generated by multiple models, as well as to experimentally measured landscapes. Our extension to frequent migrations suggests that these conclusions still hold in this regime where suppression of selection is pervasive. Our results show that spatial structure has a significant impact on early adaptation in rugged fitness landscapes. Besides, they suggest that more frequent migrations may enhance the impact of spatial structure on early adaptation. We further investigated the impact of spatial structure on long-term adaptation by considering the steady-state distribution, and introducing an effective population size that describes it. We found that suppression of selection is associated to effective population sizes smaller than the actual one, and amplification of selection to larger ones. Small steady-state effective sizes are associated to small steady-state mean fitnesses. Hence, suppression of selection is not beneficial on the very long term. The possible early advantages we identified here are thus transient.

It would be very interesting to study the impact of frequent migrations in more detail. A challenge is that fixation probabilities with frequent migrations are not known analytically beyond the branching process regime [16]. Here, we bypassed this challenge by using simulations to estimate them. However, developing more general descriptions of complex structured populations beyond rare migrations would allow more progress. Other possible directions include taking into account frequency-dependent selection [88, 89], which impacts fixation probabilities (see e.g. [90]), as well as genotype-dependent migration rates [91]. Moreover, while we focused on early and long-term adaptation, it would be interesting to consider intermediate timescales. In particular, studying adaptation after the first fitness peak is reached would involve fitness valley crossing, a process which is itself impacted by spatial structure [57, 92]. Finally, it would be of interest to consider more frequent mutations [93]. Note that, if considering fitness valley crossing, this would lead to tunneling [94]. In the framework of biased walks on fitness landscapes, this can be seen as a population hopping from a genotype to another one that is not among its nearest neighbors. Frequent mutations would also lead to multiple types being present in the population, and to clonal interference [59–62], which could have an interplay with population spatial structure.

Acknowledgments

The authors thank Alia Abbara for helpful discussions about the serial dilution model [16]. This research was partly funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 851173, to A.-F. B.), by the Swiss National Science Foundation (SNSF) (grant No. 315230_208196, to A.-F. B.), and by the Chan-Zuckerberg Initiative (CZI).

Supplementary material

Contents

1 Fixation probability of a mutant in a line with rare migrations **18**

2 Effective population sizes for steady-state distributions **23**

3 Supplementary figures **26**

4 Supplementary table **41**

1 Fixation probability of a mutant in a line with rare migrations

Let us consider a line of D demes, which may comprise individuals of two types, denoted by i and j. Each individual has a migration rate m_R to the right and m_L to the left. We focus on the rare migration regime, such that migrations are slower than extinction or fixation events within demes [13]. Thus, each deme only comprises individuals of one type, at the timescale relevant for migrations. In this section, we derive the fixation probability of the lineage of one mutant of type j within a line initially composed of individuals of (wild) type j in the rare migration regime.

For mutant fixation to occur in the line in the rare migration regime, local fixation first needs to occur in the deme where the mutant appeared, and then mutants have to spread to other demes by successive migration and fixation steps [13]. Therefore, at all times, there is one contiguous cluster of demes of mutant type j. Let us denote by $\varphi_{k,l}$ the probability that mutants fix in the line starting from a cluster of $l-k$ adjacent fully mutant demes. Here, $k \in [0, D]$ indicates the position of the first mutant deme along the line, while $l \in [0, D]$ is the position of the first wild-type deme along the line, with the constraint that $k \leq l$, see Fig. S1. Note that demes are numbered from 0 to $D-1$, that $k = l$ corresponds to the case where there is no mutant deme (including $k = l = D$) and that $l = D$ (resp. $k = 0$) means that there is no wild type deme on the right-hand side (resp. left-hand side) of the mutant cluster.

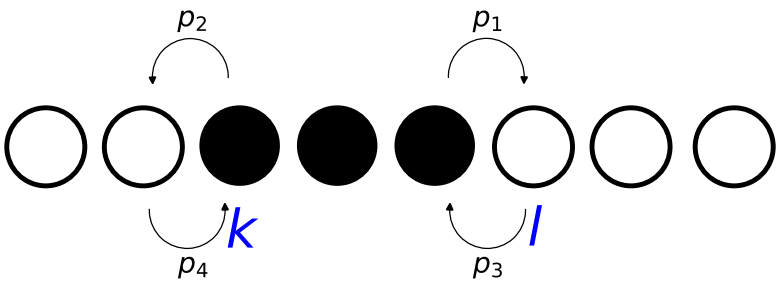

Figure S1: **State of a line graph in the rare migration regime.** We show a schematic of a line graph, where white demes only comprise wild-type individuals of type i, while black demes only comprise mutant individuals of type j. Arrows denote the transition probabilities given in Eq. (S5). Indices k and l correspond to the position of the first mutant (type j) deme and the position of the first wild type (type i) deme, respectively.

The only migration steps that change the state of the population are those where either a mutant migrates to a wild-type deme and fixes there, or the opposite. Because migrations

are to nearest neighbouring demes, these events happen at the edges of the cluster of mutant demes. Discriminating on the first migration step gives the following equation on $\varphi_{k,l}$:

$$\begin{aligned}\varphi_{k,l} &= T_{(k,l)\to(k,l+1)}\varphi_{k,l+1} + T_{(k,l)\to(k-1,l)}\varphi_{k-1,l} + T_{(k,l)\to(k,l-1)}\varphi_{k,l-1} + T_{(k,l)\to(k+1,l)}\varphi_{k+1,l} \\ &+ \left(1 - T_{(k,l)\to(k,l+1)} - T_{(k,l)\to(k-1,l)} - T_{(k,l)\to(k,l-1)} - T_{(k,l)\to(k+1,l)}\right)\varphi_{k,l}\,, \end{aligned} \tag{S1}$$

with the transition probabilities

$$T_{(k,l)\to(k,l+1)} = \frac{m_R N_j \rho_{ij}}{Z_{k,l}}, \quad T_{(k,l)\to(k-1,l)} = \frac{m_L N_j \rho_{ij}}{Z_{k,l}}, \tag{S2}$$

$$T_{(k,l)\to(k,l-1)} = \frac{m_L N_i \rho_{ji}}{Z_{k,l}}, \quad T_{(k,l)\to(k+1,l)} = \frac{m_R N_i \rho_{ji}}{Z_{k,l}}, \tag{S3}$$

where $Z_{k,l}$ is a normalisation constant. Eq. (S1) can be rewritten as follows:

$$\varphi_{k,l} = p_1\varphi_{k,l+1} + p_2\varphi_{k-1,l} + p_3\varphi_{k,l-1} + p_4\varphi_{k+1,l}, \tag{S4}$$

with

$$\begin{aligned} p_1 &= \frac{\alpha}{(1+\alpha)(1+\gamma)}, \quad p_2 = \frac{1}{(1+\alpha)(1+\gamma)}, \\ p_3 &= \frac{\gamma}{(1+\alpha)(1+\gamma)}, \quad p_4 = \frac{\alpha\gamma}{(1+\alpha)(1+\gamma)}, \end{aligned} \tag{S5}$$

where we introduced the dimensionless coefficients $\alpha = m_R/m_L$ and $\gamma = N_i\rho_{ji}/N_j\rho_{ij}$, where ρ_{ij} is the fixation probability of the lineage of one mutant with genotype j within a deme of individuals with genotype i.

In addition to Eq. (S4), the following boundary conditions hold:

$$\varphi_{k,k} = 0, \quad \varphi_{0,D} = 1, \tag{S6}$$

$$\varphi_{k,D} = U_k, \quad \varphi_{0,k} = V_k, \tag{S7}$$

where U_k and V_k will be computed below.

In Fig. S2, we illustrate the Markov chain described here, with the different states of the line and the possible transitions between them.

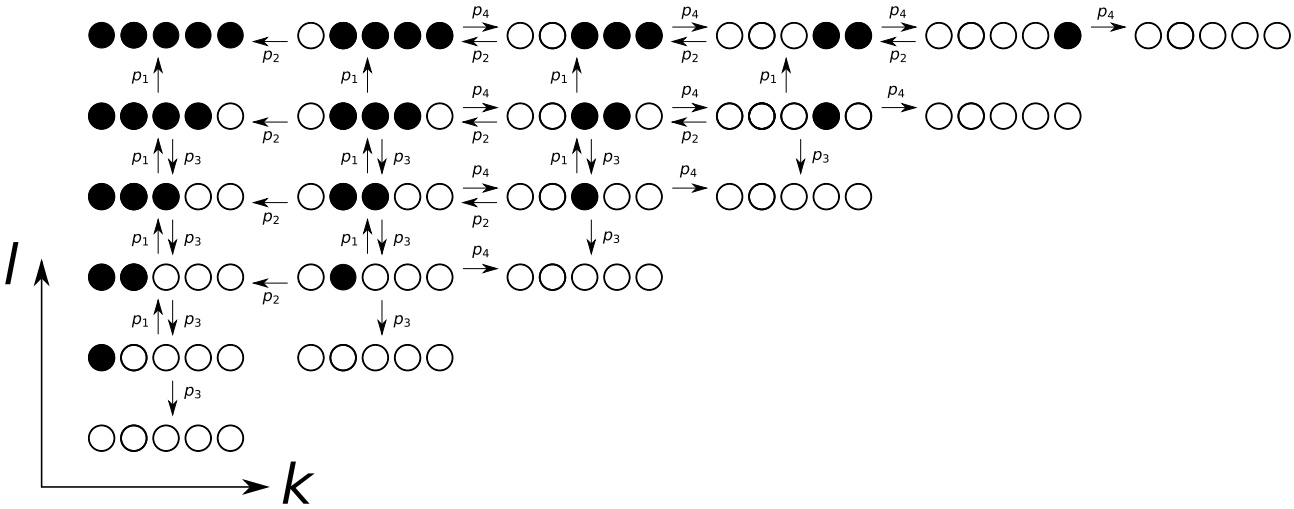

Figure S2: **Markov chain for a line graph in the rare migration regime.** We show a schematic of all different possible states for a line of $D = 5$ demes. The arrows correspond to the possible transitions, and are annotated with their probabilities. This figure is adapted from Fig. 5 of Ref. [95].

To find the expression of $\varphi_{k,l}$, we need the boundary conditions U_k and V_k. Let us first determine $U_k = \varphi_{k,D}$. It satisfies

$$U_k = W_{k\to k-1}U_{k-1} + W_{k\to k+1}U_{k+1} + (1 - W_{k\to k-1} - W_{k\to k+1})U_k\,, \tag{S8}$$

with the following transition probabilities:

$$W_{k\to k-1} = \frac{m_L N_j \rho_{ij}}{L_k}, \quad W_{k\to k+1} = \frac{m_R N_i \rho_{ji}}{L_k}, \tag{S9}$$

where L_k is a normalisation constant. We can be rewrite Eq. (S8) as:

$$0 = \alpha\gamma U_{k+1} - (1 + \alpha\gamma)U_k + U_{k-1}\,. \tag{S10}$$

The solution of this equation has the form

$$U_k = \frac{A}{(\alpha\gamma)^k} + B\,, \tag{S11}$$

where the constants A and B are determined by the boundary conditions:

$$U_0 = 1\,, \tag{S12}$$

$$U_D = 0\,. \tag{S13}$$

This gives

$$U_k = \frac{\dfrac{1}{(\alpha\gamma)^k} - \dfrac{1}{(\alpha\gamma)^D}}{1 - \dfrac{1}{(\alpha\gamma)^D}}. \tag{S14}$$

Let us now perform a similar calculation for $V_k = \varphi_{0,k}$, which satisfies

$$V_k = R_{k\to k-1}V_{k-1} + R_{k\to k+1}V_{k+1} + (1 - R_{k\to k-1} - R_{k\to k+1})V_k, \tag{S15}$$

with the following transition probabilities:

$$R_{k\to k-1} = \frac{m_L N_i \rho_{ji}}{M_k}, \quad R_{k\to k+1} = \frac{m_R N_j \rho_{ij}}{M_k}\,, \tag{S16}$$

where M_k is another normalisation constant. This leads to:

$$0 = \alpha V_{k+1} - (\alpha+\gamma)V_k + \gamma V_{k-1}\,. \tag{S17}$$

We obtain

$$V_k = C_1 \left(\frac{\gamma}{\alpha}\right)^k + C_2\,, \tag{S18}$$

with constants C_1 and C_2 that are determined by the following boundary conditions:

$$V_0 = 0\,, \tag{S19}$$

$$V_D = 1\,. \tag{S20}$$

This gives

$$V_k = \frac{1-\left(\frac{\gamma}{\alpha}\right)^k}{1-\left(\frac{\gamma}{\alpha}\right)^D}\,. \tag{S21}$$

Now that Eq. (S4) together with Eqs. (S6) and (S7) constitute a fully posed problem, we can determine $\varphi_{k,l}$. In Ref. [96], a general solution of Eq. (S4) was obtained for a two-dimensional square lattice. It reads:

$$\varphi_{k,l} = \sum_{n=1}^{D-1}\left[p_1 P_{n,D} T_{k,l}^{(n,D-1)} + p_2 P_{0,n} T_{k,l}^{(1,n)} + p_3 P_{n,0} T_{k,l}^{(n,1)} + p_4 P_{D,n} T_{k,l}^{(D-1,n)}\right], \tag{S22}$$

where the $P_{i,j}$ terms correspond to the four boundary conditions at each edge of the square and the transition probability $T_{k,l}^{(a,b)}$ reads

$$\begin{aligned} T_{k,l}^{(a,b)} &= \frac{4}{D^2}\left(\frac{p_2}{p_4}\right)^{(k-a)/2}\left(\frac{p_3}{p_1}\right)^{(l-b)/2}\sum_{r=1}^{D-1}\sum_{s=1}^{D-1}\frac{\sin\left(\frac{kr\pi}{D}\right)\sin\left(\frac{ar\pi}{D}\right)\sin\left(\frac{bs\pi}{D}\right)\sin\left(\frac{ls\pi}{D}\right)}{1-2\sqrt{p_2p_4}\cos\left(\frac{r\pi}{D}\right)-2\sqrt{p_1p_3}\cos\left(\frac{s\pi}{D}\right)} \\ &= \frac{4}{D^2}\left(\frac{1}{\alpha\gamma}\right)^{(k-a)/2}\left(\frac{\gamma}{\alpha}\right)^{(l-b)/2}\sum_{r=1}^{D-1}\sum_{s=1}^{D-1}\frac{\sin\left(\frac{kr\pi}{D}\right)\sin\left(\frac{ar\pi}{D}\right)\sin\left(\frac{bs\pi}{D}\right)\sin\left(\frac{ls\pi}{D}\right)}{1-\frac{2\sqrt{\alpha\gamma}}{(1+\alpha)(1+\gamma)}\left[\cos\left(\frac{r\pi}{D}\right)+\cos\left(\frac{s\pi}{D}\right)\right]}. \end{aligned} \tag{S23}$$

However, this solution does not directly apply to our problem since the boundary conditions are different. Indeed, we have $P_{n,D} = U_n$ and $P_{0,n} = V_n$ but we do not have $P_{D,n}$ or $P_{n,0}$. Instead, we have the following condition on the diagonal $k = l$: $\varphi_{k,k} = 0$. To address this difference, we generalize the method developed in Ref. [95] in the context of models with one individual per node of the line to the present case, and in particular to asymmetric transition rates. The method of Ref. [95] consists in finding the missing boundary conditions $P_{n,0}$ and $P_{D,n}$ by using the constraint $\varphi_{k,k} = 0$. To do so, we first note that

$$T_{k,k}^{(a,b)} = \frac{\left(\frac{p_2}{p_4}\right)^{b/2}\left(\frac{p_3}{p_1}\right)^{a/2}}{\left(\frac{p_2}{p_4}\right)^{a/2}\left(\frac{p_3}{p_1}\right)^{b/2}} T_{k,k}^{(b,a)} = \gamma^{a-b} T_{k,k}^{(b,a)}\,. \tag{S24}$$

Using this property, we obtain

$$\varphi_{k,k} = 0 = \sum_{n=1}^{D-1} T_{k,k}^{(n,D-1)}\left[p_1 U_n + p_4 P_{D,n}\gamma^{D-1-n}\right] + \sum_{n=1}^{D-1} T_{k,k}^{(1,n)}\left[p_2 V_n + p_3 P_{n,0}\gamma^{n-1}\right]. \tag{S25}$$

This equality is satisfied if

$$p_1 U_n + p_4 P_{D,n} \gamma^{D-1-n} = 0\,, \tag{S26}$$
$$p_2 V_n + p_3 P_{n,0} \gamma^{n-1} = 0\,. \tag{S27}$$

This yields

$$P_{D,n} = -\gamma^{n-D} U_n\,, \tag{S28}$$
$$P_{n,0} = -\gamma^{-n} V_n\,. \tag{S29}$$

Finally, the solution of Eq. (S4) with the boundary conditions in Eqs. (S6) and (S7) reads

$$\varphi_{k,l} = \sum_{n=1}^{D-1} p_1 U_n \left[T_{k,l}^{(n,D-1)} - \gamma^{n-D+1}\, T_{k,l}^{(D-1,n)} \right] + p_2 V_n \left[T_{k,l}^{(1,n)} - \gamma^{1-n}\, T_{k,l}^{(n,1)} \right], \tag{S30}$$

which finally leads to:

$$\varphi_{k,l} = \frac{1}{(1+\alpha)(1+\gamma)} \sum_{n=1}^{D-1} \left\{ \alpha\, \frac{(\alpha\gamma)^{-n} - (\alpha\gamma)^{-D}}{1-(\alpha\gamma)^{-D}} \left[T_{k,l}^{(n,D-1)} - \gamma^{n-D+1}\, T_{k,l}^{(D-1,n)} \right] \right.$$
$$\left. + \frac{1-(\gamma/\alpha)^n}{1-(\gamma/\alpha)^D} \left[T_{k,l}^{(1,n)} - \gamma^{1-n}\, T_{k,l}^{(n,1)} \right] \right\}. \tag{S31}$$

The fixation probability in the line starting from one mutant in deme k can be expressed as $\rho_{ij}\varphi_{k,k+1}$. By averaging out over all demes, assuming that the mutant appears in a deme chosen uniformly at random, we can write the overall mutant fixation probability as

$$\phi_{ij} = \frac{\rho_{ij}}{D} \sum_{k=0}^{D-1} \varphi_{k,k+1}, \tag{S32}$$

where $\varphi_{0,1} = V_1$, $\varphi_{D-1,D} = U_{D-1}$ and, $\forall k \in [1, D-2]$, we have

$$\varphi_{k,k+1} = \frac{4}{D^2(1+\alpha)(1+\gamma)} \sum_{n=1}^{D-1} \left\{ \frac{(\alpha\gamma)^{-n} - (\alpha\gamma)^{-D}}{1-(\alpha\gamma)^{-D}} \frac{\gamma^{(n-D+2)/2}}{\alpha^{(2k-D-n)/2}} \left[t_{k,k+1}^{(n,D-1)} - t_{k,k+1}^{(D-1,n)} \right] \right.$$
$$\left. + \frac{1-(\gamma/\alpha)^n}{1-(\gamma/\alpha)^D} \frac{\gamma^{(2-n)/2}}{\alpha^{(2k-n)/2}} \left[t_{k,k+1}^{(1,n)} - t_{k,k+1}^{(n,1)} \right] \right\}, \tag{S33}$$

with

$$t_{k,k+1}^{(a,b)} = \sum_{r=1}^{D-1} \sum_{s=1}^{D-1} \frac{\sin\left(\frac{kr\pi}{D}\right) \sin\left(\frac{ar\pi}{D}\right) \sin\left(\frac{bs\pi}{D}\right) \sin\left(\frac{(k+1)s\pi}{D}\right)}{1 - \frac{2\sqrt{\alpha\gamma}}{(1+\alpha)(1+\gamma)} \left[\cos\left(\frac{r\pi}{D}\right) + \cos\left(\frac{s\pi}{D}\right) \right]}\,. \tag{S34}$$

Using the symmetry properties of the line, the overall mutant fixation probability can be rewritten as

$$\phi_{ij} = \frac{\rho_{ij}}{D} \left\{ \frac{1-\gamma/\alpha}{1-(\gamma/\alpha)^D} + \frac{1-\alpha\gamma}{1-(\alpha\gamma)^D} \right. \tag{S35}$$
$$\left. + \frac{4}{D^2(1+\alpha)(1+\gamma)} \sum_{k=1}^{D-2} \sum_{n=1}^{D-1} \gamma^{1-\frac{n}{2}} \left[\frac{1-(\alpha\gamma)^n}{1-(\alpha\gamma)^D} \alpha^{k+1-\frac{n}{2}} + \frac{1-\left(\frac{\gamma}{\alpha}\right)^n}{1-\left(\frac{\gamma}{\alpha}\right)^D} \alpha^{\frac{n}{2}-k} \right] \left(t_{k,k+1}^{(1,n)} - t_{k,k+1}^{(n,1)} \right) \right\},$$

The ratio of the fixation probability given in Eq. (S35) to that in one deme (ρ_{ij}) is plotted in Fig. S3 for different values of migration asymmetry α. This ratio represents the fixation probability in the line starting from one fully mutant deme chosen uniformly at random, in the rare migration regime. For $\alpha = 1$, we recover the fixation probability found for a circulation graph (see Eq. (3) in the main text). For other values of α, we observe in Fig. S3 that the line graph with asymmetric migrations is a strong suppressor of selection compared to circulation graphs. In addition, we have the symmetry property $\phi_{ij}(\alpha, \gamma) = \phi_{ij}(1/\alpha, \gamma)$ and in the limit of strong migration rate asymmetries, we find

$$\lim_{\alpha \to 0} \frac{\phi_{ij}}{\rho_{ij}} = \lim_{\alpha \to +\infty} \frac{\phi_{ij}}{\rho_{ij}} = \frac{1}{D} \tag{S36}$$

which corresponds to the strongest suppression of selection, where the fixation probability reduces to that of a neutral mutant whatever the fitness of the mutant.

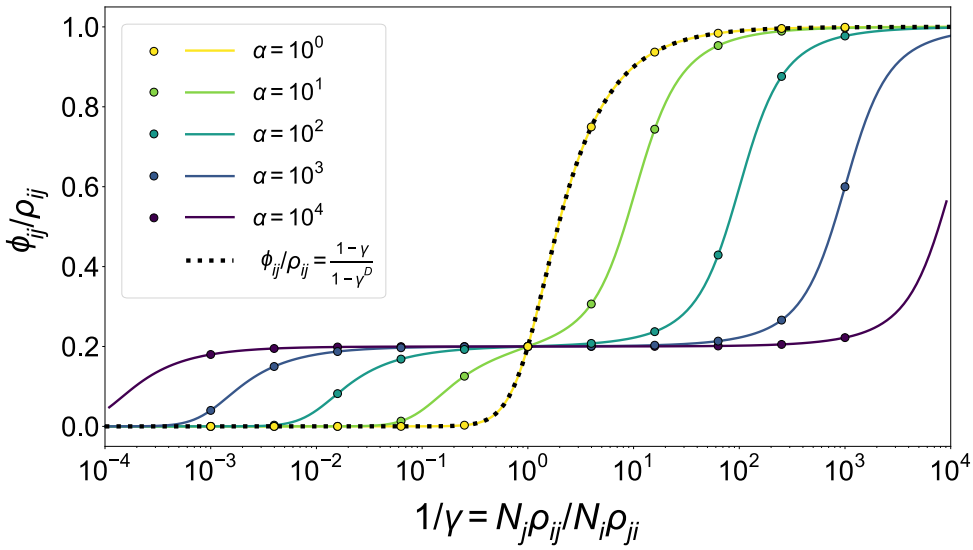

Figure S3: **Mutant fixation probability in the line with rare migrations, starting from a fully mutant deme.** The mutant fixation probability starting from a fully mutant deme, obtained as ϕ_{ij}/ρ_{ij}, is plotted versus $1/\gamma$ for different values of α, with $D = 5$ demes. Coloured lines correspond to our analytical prediction in Eq. (S35). The black dotted line shows the fixation probability for circulation graphs (e.g. clique or cycle graphs, see [13, 16]). Markers correspond to the numerical simulation of the Markov chain given in Eq. (S4), and are obtained from 10^7 simulation replicates.

2 Effective population sizes for steady-state distributions

In this section, we focus on the steady state of the walks of structured populations on fitness landscapes. As mentioned in Section Model and methods in the main text, the Markov chains corresponding to our walks in fitness landscape possess a (unique) steady state towards which they converge from any starting genotype, since they are irreducible, aperiodic and positive recurrent [66, 67]. They are also reversible and comply with the detailed balance equation. Thus, for any pair of neighboring states β and θ (i.e., states that are neighbors in the hypercube of genotype space), we have [68, 69]:

$$\frac{\pi_\theta}{\pi_\beta} = \frac{P_{\beta\theta}}{P_{\theta\beta}}, \tag{S37}$$

where $P_{\theta\beta}$, which denotes an element of the transition matrix P, corresponds to the transition probability from θ to β upon one mutation event, while π_θ is the stationary probability that the population is in state θ.

Consider a path composed of n successive steps between neighboring genotypes α_i, with $i \in [1, n+1]$. This path connects the end genotypes α_1 and α_{n+1}, which have a Hamming distance of n. Using Eq. (S37) for each step of the path gives

$$\frac{\pi_{\alpha_1}}{\pi_{\alpha_{n+1}}} = \prod_{i=1}^{n} \frac{P_{\alpha_{i+1}\alpha_i}}{P_{\alpha_i\alpha_{i+1}}} \,. \tag{S38}$$

For a well-mixed population of size N in the Moran model, Ref. [68] showed that the stationary distribution can be expressed explicitly as:

$$\pi_\theta^{\text{Moran}} = \frac{f_\theta^{N-1}}{\sum_{\delta \in H} f_\delta^{N-1}} \,, \tag{S39}$$

where H is the set of all genotypes (i.e. all vertices of the hypercube). Let us consider our spatially structured populations on graphs in the rare migration regime. We will show below that, in some limits, we can follow the line of reasoning of Ref. [68], and obtain explicit formulas for stationary distributions. Each spatial structure is associated with a specific matrix P. We focus on Moran walks in these structures (i.e., the fixation probability within a deme is the Moran fixation probability; see Section Model and methods in the main text).

Circulations. Let us consider a population with current genotype θ, structured as a circulation with D demes that each have carrying capacity C, and let us denote genome dimension by L. In the rare migration regime, the transition probability from genotype θ to a neighboring genotype β upon one mutation event can be written as:

$$P_{\theta\beta} = \frac{1}{L} \times \phi_{\theta\beta} \,. \tag{S40}$$

The first term of this product is the probability that the mutation that occurs is toward genotype β. Indeed, there are L possible mutations starting from a given genotype, and they are all equally likely. The second term, $\phi_{\theta\beta}$, is the probability that a mutant of type β fixes in the circulation. As mentioned in Section Model and methods in the main text, for rare migrations, it can be expressed as [13]:

$$\phi_{\theta\beta} = \rho_{\theta\beta} \times \frac{1-\gamma}{1-\gamma^D} \,, \quad \text{with } \gamma = \frac{\rho_{\beta\theta}}{\rho_{\theta\beta}} \,, \tag{S41}$$

where $\rho_{\theta\beta}$ denotes the probability that a mutant of type β fixes in a deme where all other individuals are of type θ. Assuming for simplicity that $g \ll f_\alpha$ for any genotype α, where g is the death rate, we can approximate the stationary population sizes of demes by their carrying capacity C, and use the Moran process to describe fixation in the demes (see Section Model and methods in the main text). This yields:

$$\rho_{\theta\beta} = \frac{1-r}{1-r^C} \,, \quad \text{with } r = \frac{f_\theta}{f_\beta} \,, \tag{S42}$$

and $\gamma = r^{C-1}$. Thus, we can write:

$$\frac{P_{\theta\beta}}{P_{\beta\theta}} = \gamma^{-1} \frac{1-\gamma}{1-\gamma^{-1}} \frac{1-\gamma^{-D}}{1-\gamma^D} = \gamma^{-D} = r^{-D(C-1)} = \left(\frac{f_\beta}{f_\theta}\right)^{D(C-1)} \,. \tag{S43}$$

Using Eq. (S43) and Eq. (S38), we find:

$$\frac{\pi_{\alpha_1}}{\pi_{\alpha_{n+1}}} = \left[\prod_{i=1}^{n} \frac{f_{\alpha_i}}{f_{\alpha_{i+1}}}\right]^{D(C-1)} = \left(\frac{f_{\alpha_1}}{f_{\alpha_{n+1}}}\right)^{D(C-1)} . \tag{S44}$$

This only depends on the fitness values of the starting and the final genotypes of the path considered. Therefore, such an equation can be written for any pair of genotypes. Considering two genotypes θ and δ, we can write

$$\pi_\delta = \pi_\theta \left(\frac{f_\delta}{f_\theta}\right)^{D(C-1)} , \tag{S45}$$

Combining this with the normalization condition of stationary probabilities, namely $\sum_{\delta\in H} \pi_\delta = 1$, leads to the following explicit formula for the stationary probability in a circulation:

$$\pi_\theta^{\text{circ.}} = \frac{f_\theta^{D(C-1)}}{\sum_{\delta\in H} f_\delta^{D(C-1)}} . \tag{S46}$$

Star. Let us now consider the star in the rare migration limit. First, in the limit $\alpha \to 0$, where α denotes migration asymmetry (see Section Model and methods in the main text), the transition probability between two neighboring genotypes θ and β simplifies to:

$$P_{\theta\beta} = \frac{1}{L} \times \rho_{\theta\beta} \times \frac{1}{D} . \tag{S47}$$

Indeed, the fixation probability starting from one fully mutant deme then tends to $1/D$ [13]. Thus, the ratios of stationary probabilities are the same as in a deme. Following the same line of reasoning as in a well-mixed population [68] and as for circulations above, we obtain:

$$\pi_\theta^{\text{star, } \alpha\to 0} = \frac{f_\theta^{C-1}}{\sum_{\delta\in H} f_\delta^{C-1}} . \tag{S48}$$

This coincides with the stationary distribution for a well-mixed population with size C in the Moran model, see Eq. (S39).

Second, if $\alpha \to \infty$, using the corresponding limit of the fixation probability starting from one fully mutant deme [13] leads to the transition probability

$$P_{\theta\beta} = \frac{1}{L} \times \rho_{\theta\beta} \times \frac{D-1}{D} \frac{1-\gamma^2}{1-\gamma^{2(D-1)}} , \tag{S49}$$

which leads to

$$\pi_\theta^{\text{star, } \alpha\to\infty} = \frac{f_\theta^{(C-1)(2D-3)}}{\sum_{\delta\in H} f_\delta^{(C-1)(2D-3)}} . \tag{S50}$$

Line. We showed above that, in the line with rare migrations in the limit of strong migration asymmetries, the fixation probability starting from one fully mutant deme tends to $1/D$, see Eq. (S36). Thus, the stationary distribution is given by Eq. (S47), as for the star in the limit $\alpha \to 0$.

Effective population sizes. Comparing Eqs. (S46), (S48) and (S50) to the well-mixed case in Eq. (S39) reveals a strong formal similarity. Specifically, only the exponent of the fitnesses differs. This leads us to define an effective population size for the steady state, as the size of a well-mixed population that would display the same stationary distribution. Table 1 in the main text lists the effective population sizes corresponding to the different structures we studied above.

While we focused on Moran walks here, similar derivations could be performed for Wright-Fisher walks by taking the approximation from Ref. [68] for the fixation probability within a deme, instead of the classical expression used in the main text (see Ref. [68]).

3 Supplementary figures

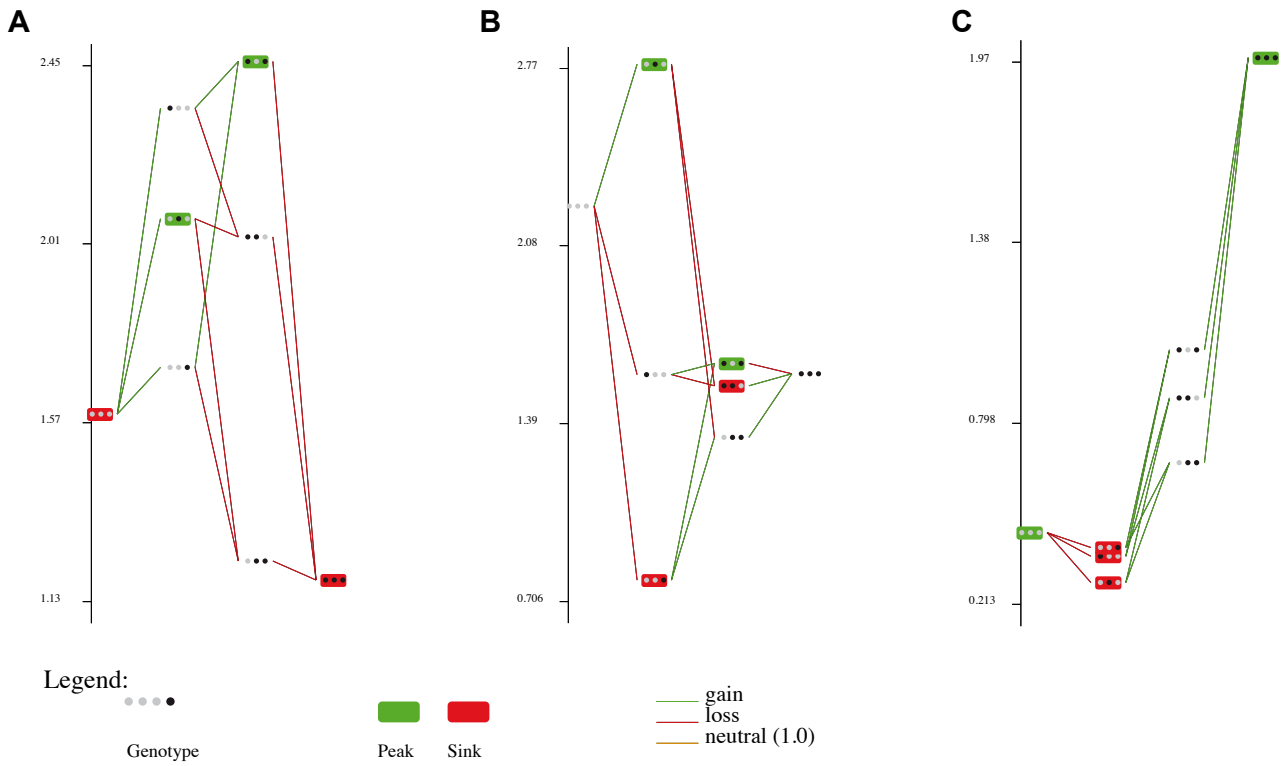

Figure S4: **MAGELLAN representation [82] of the fitness landscapes A, B and C from Fig. 2(I-K).** Fitness values are given in Table S1. The vertical axis denotes fitness, while the horizontal axis corresponds to the Hamming distance to the wild type. At each site, a gray marker denotes a wild-type genetic unit (nucleotide, amino acid or gene), while a black one denotes a mutated one.

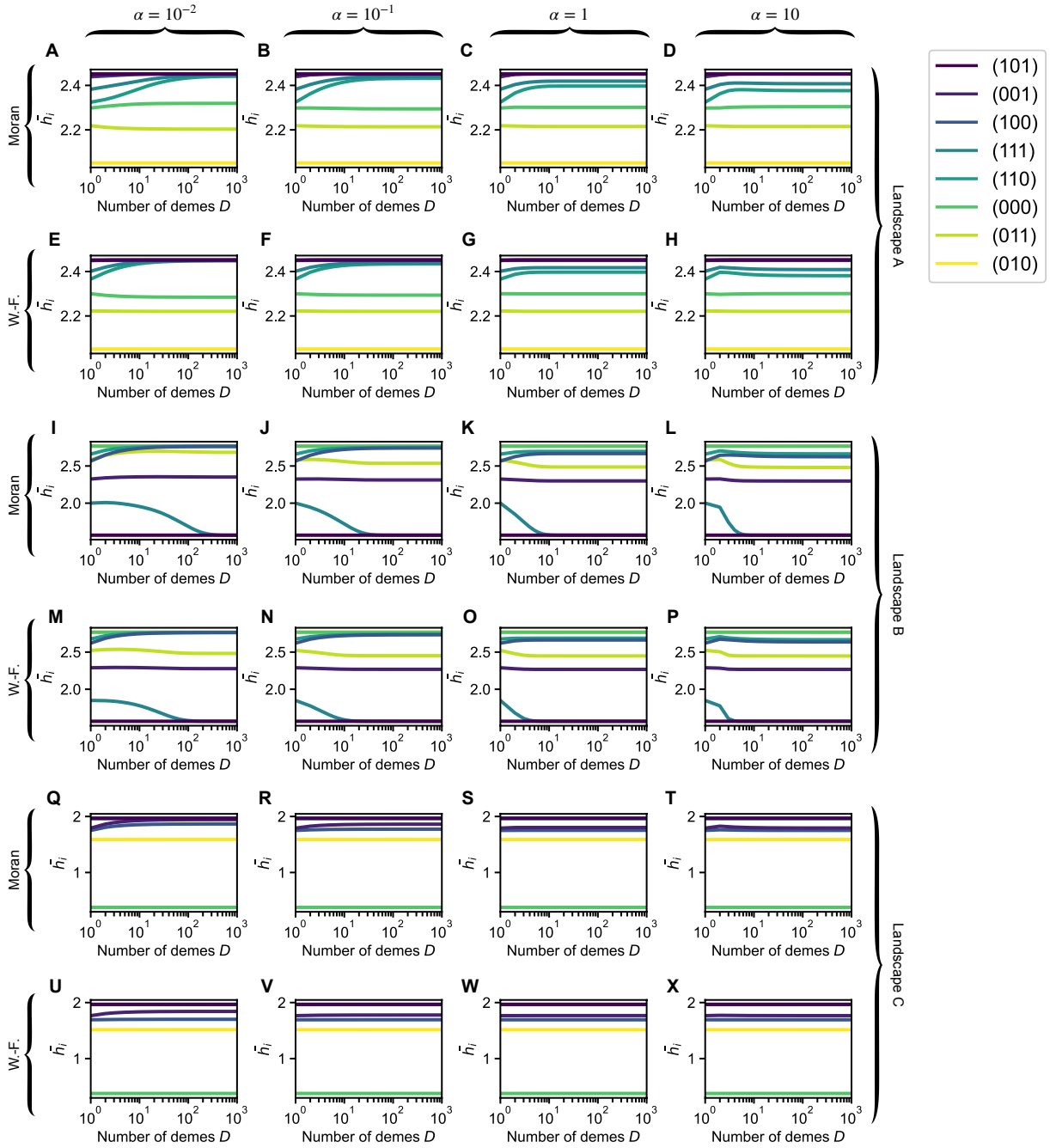

Figure S5: **Impact of the number of demes D and migration asymmetry α of a star on early adaptation starting from a specific genotype i, in LK landscapes with $L = 3$ and $K = 1$.** Mean height $\bar{h}_i$, starting from a specific genotype i, versus D in landscape A (A-H), landscape B (I-P), and landscape C (Q-X), for the Moran star walk (first row for each landscape) and the Wright-Fisher star walk (second row for each landscape), for different migration asymmetry values (different columns). Landscapes A, B and C are shown in Fig. 2(I-K) and Fig. S4, and the corresponding fitness values are provided in Table S1. Lines are numerical resolutions of the FSA equations for each landscape. Colors denote the starting genotype. Note that in landscape A, $h_{(101)}$, $h_{(001)}$ and $h_{(100)}$ are indistinguishable. In landscape B, $h_{(010)}$ and $h_{(000)}$ are indistinguishable. In landscape C, $h_{(111)}$, $h_{(011)}$, $h_{(101)}$ and $h_{(110)}$ are indistinguishable. In all cases, parameter values are $C = 20$ and $g = 0.01$, as in Fig. 2.

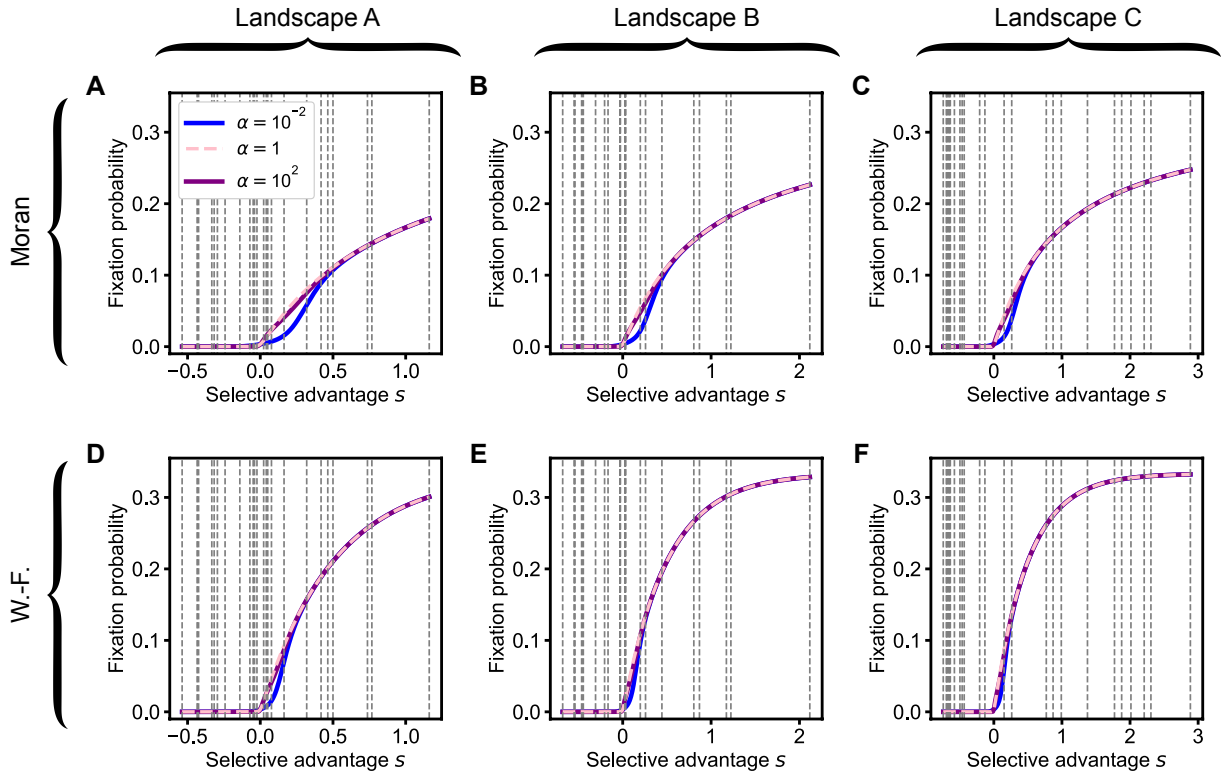

Figure S6: **Fixation probability of mutations within a star-structured population with rare migrations.** (A) Fixation probability of a mutation in a (Moran) star versus its relative selective advantage s for different values of migration asymmetry α in a star with rare migrations [13]. The values of s in landscape A (see Table S1) are indicated by vertical gray dashed lines. (B-C) Same as (A) but for landscapes B and C, respectively (same curves, but different range and different vertical gray dashed lines). (D-F) Same as (A-C) but for the Wright-Fisher star. In all cases, parameter values are $D = 5$, $C = 20$, and $g = 0.01$.

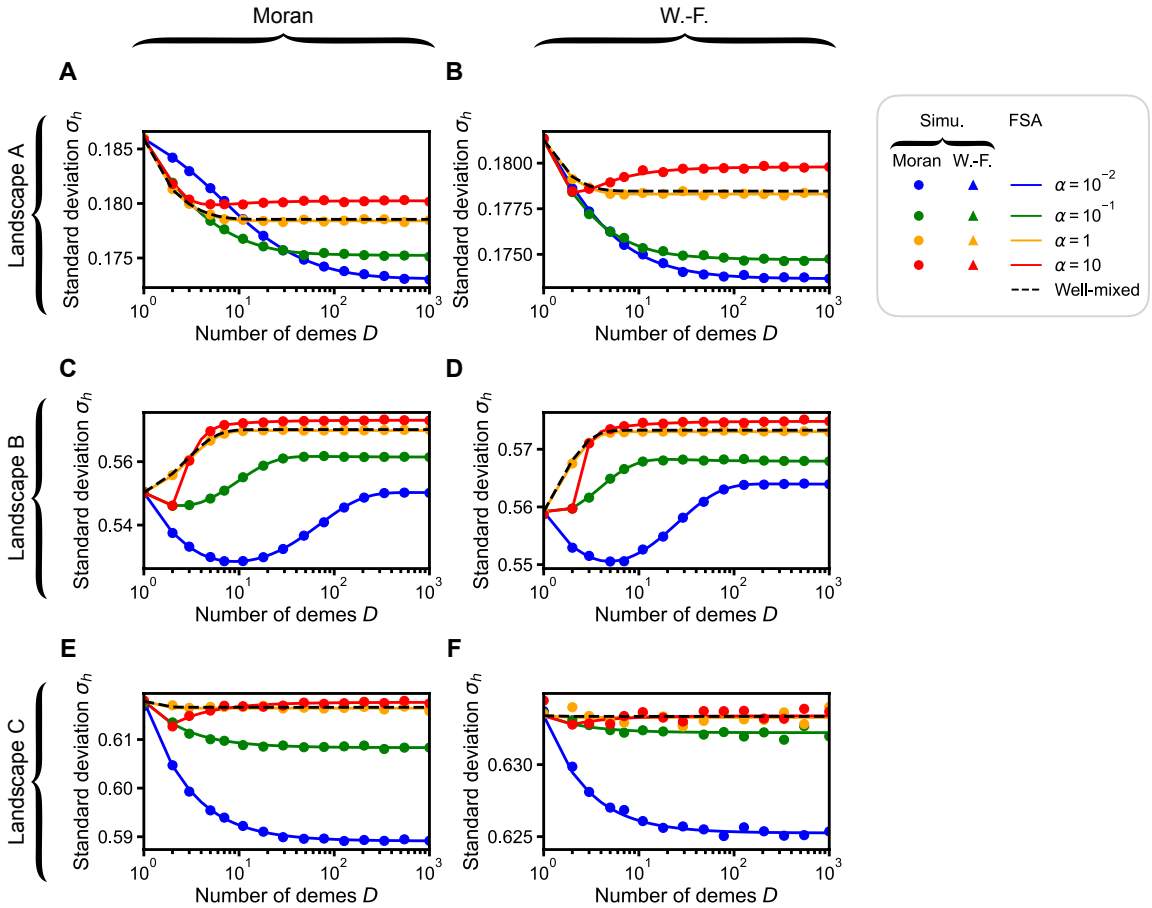

Figure S7: **Impact of the number of demes D of a star on the standard deviation of the height h of a walk, in LK landscapes with $L = 3$ and $K = 1$.** (A,B) The standard deviation of the height h of a walk, starting from a uniformly chosen initial genotype, is shown versus D for the Moran (left panel) and Wright-Fisher (right panel) star walk in landscape A. (C,D) and (E,F) Same as (A,B) but in landscapes B and C, respectively. In all cases, simulation results are averaged over at least 10^5 walks per starting genotype. Parameter values: $C = 20$ and $g = 0.01$.

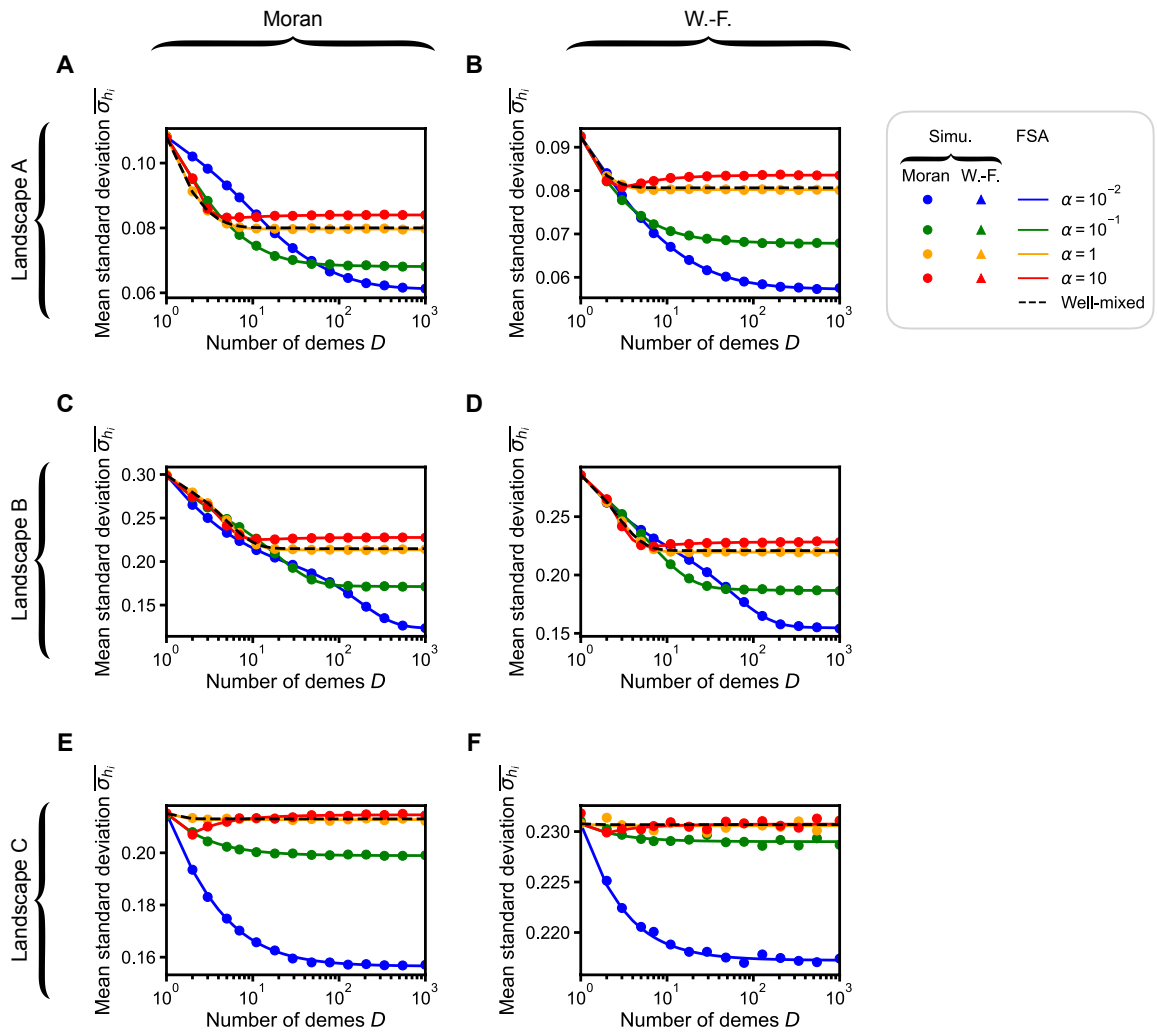

Figure S8: **Impact of the number of demes D of a star on the mean standard deviation of the height h_i of a walk, in LK landscapes with $L = 3$ and $K = 1$.** Same as Fig. S7 but for the mean standard deviation $\overline{\sigma_{h_i}}$ of the height h_i of the first fitness peak reached when starting from genotype i, averaged over all possible starting genotypes i.

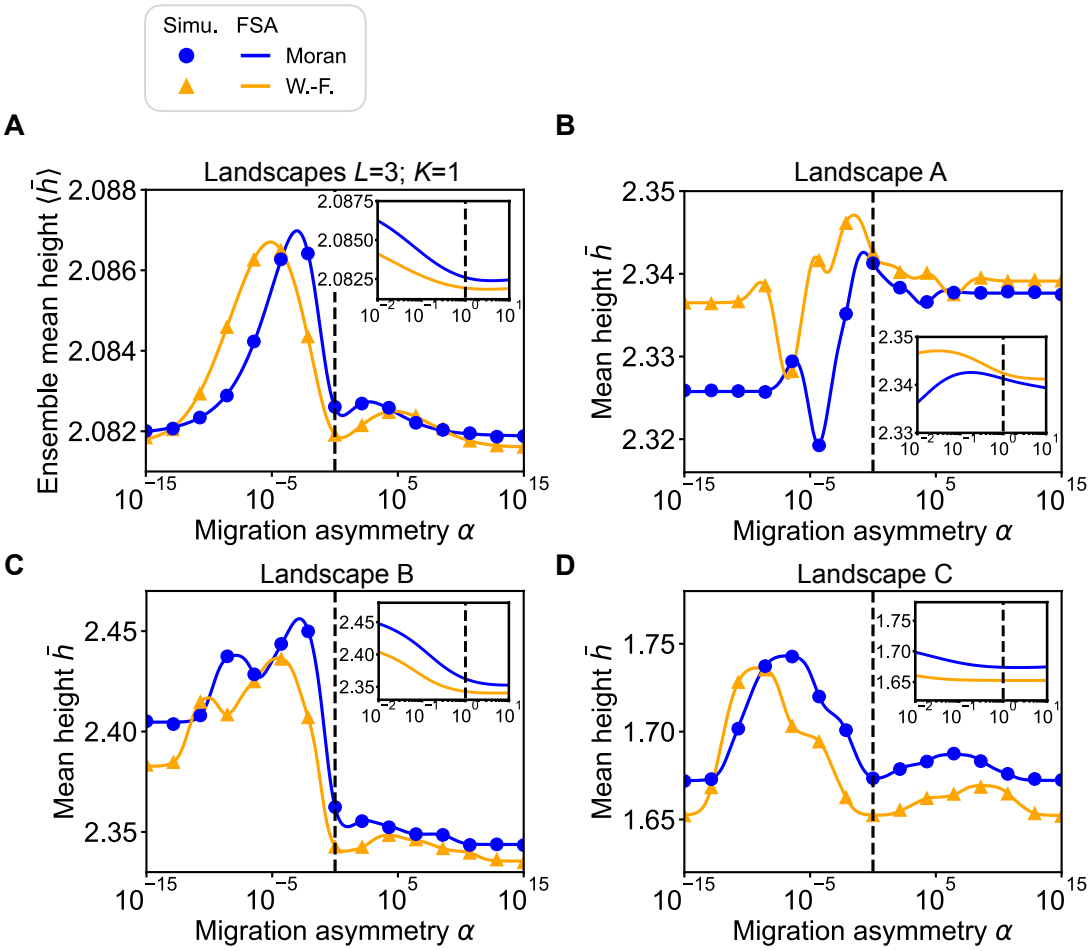

Figure S9: **Impact of the migration asymmetry α of a star on early adaptation in LK landscapes with $L = 3$ and $K = 1$.** (A) Ensemble mean height $\langle \bar{h} \rangle$ of the first fitness peak reached when starting from a uniformly chosen initial genotype versus α. (B-D) Mean height $\bar{h}$ versus α in landscapes A, B and C, shown in Fig. 2(I-K). Solid lines correspond to numerical resolutions of the FSA equations, while markers are simulation results averaged over 2×10^5 landscapes and 100 walks per starting genotype for (A), and 10^5 walks per starting genotype for (B-D). In all cases, parameter values are $C = 20$, $D = 5$ and $g = 0.01$.

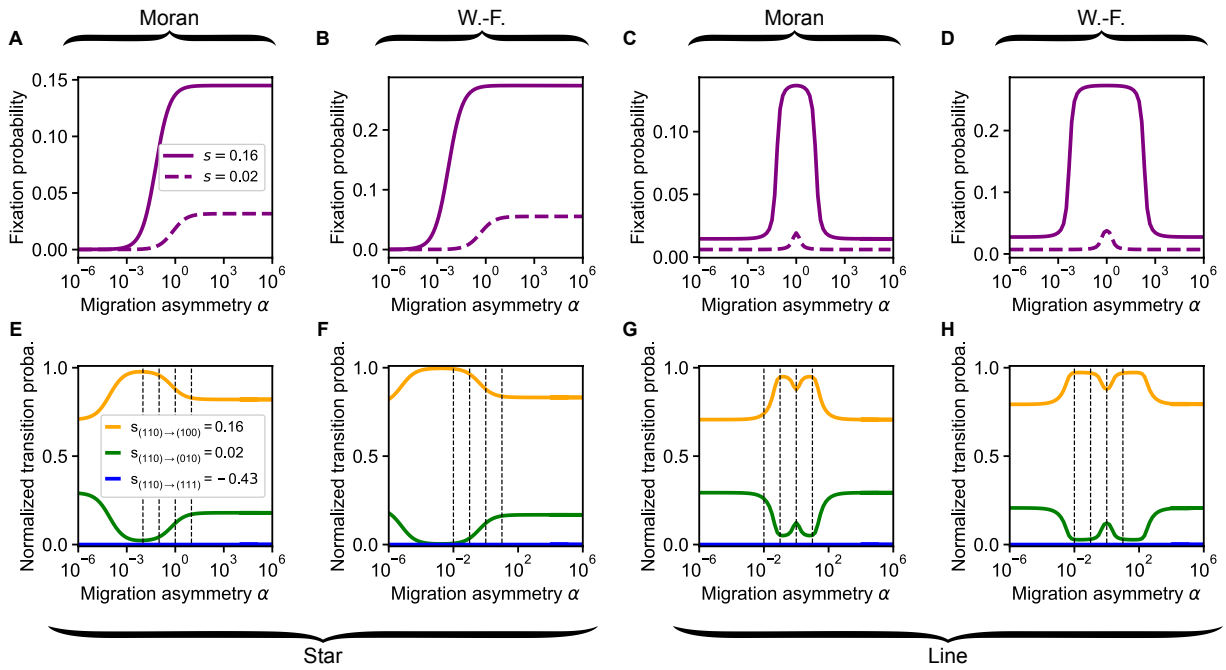

Figure S10: **Impact of the migration asymmetry α of a star and of a line on the fixation probability of different mutations.** (A) Fixation probability of mutations with respective relative fitness advantage values $s = 0.16$ and $s = 0.02$ in a Moran star for rare migrations versus migration asymmetry α, for $D = 10^3$. These relative fitness advantage values correspond respectively to the transitions $(110) \to (100)$ and $(110) \to (010)$ within landscape A. (B) Same as (A) but for the Wright-Fisher star. (C-D) Same as (A-B) but for the line with $D = 10$. (E) Probabilities of transitions $(110) \to (100)$, $(110) \to (010)$ and $(110) \to (111)$ within landscape A, normalized so that they sum to 1, versus α, for a Moran star with rare migrations, for $D = 10^3$. (B) Same as (A) but for the Wright-Fisher star. (C-D) Same as (A-B) but for the line with $D = 10$. In all cases, parameter values are $C = 20$ and $g = 0.01$.

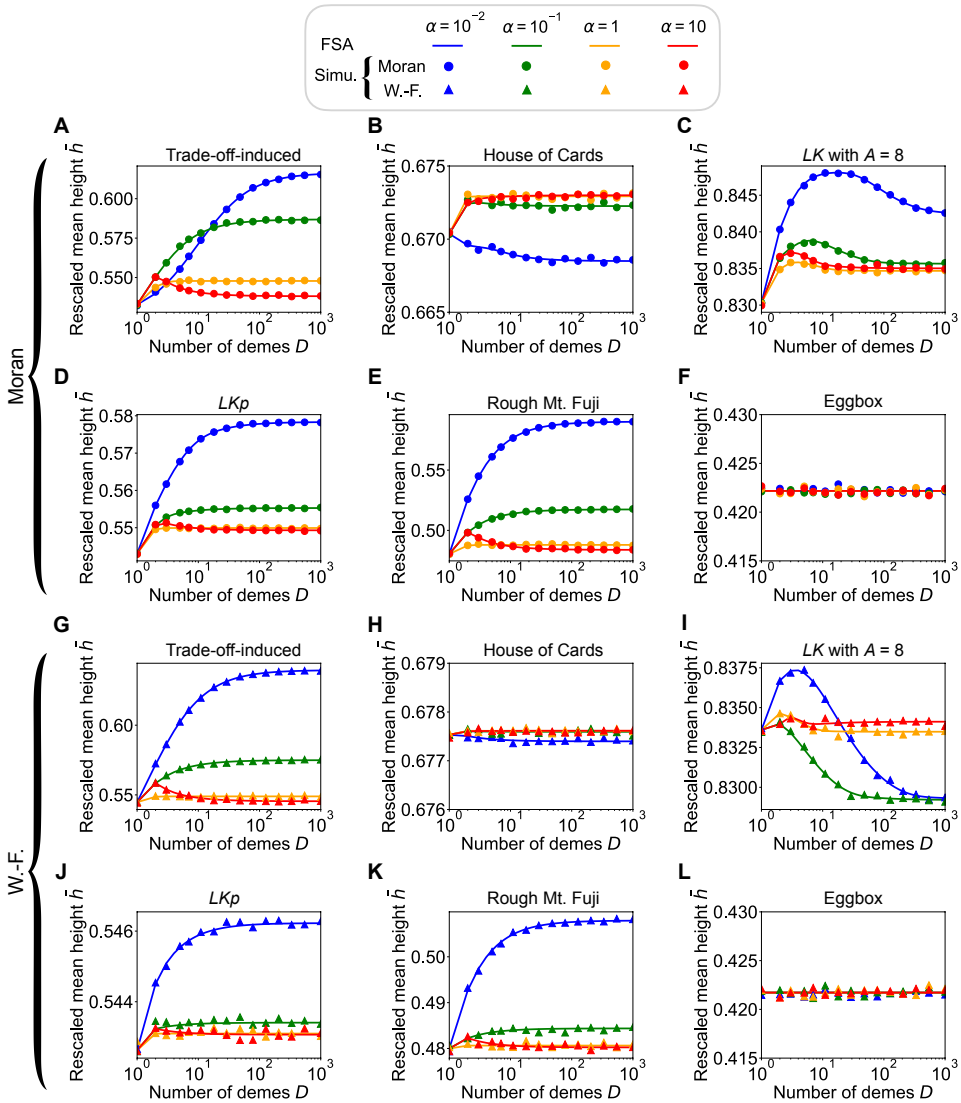

Figure S11: **Impact of the number of demes D of a star on early adaptation in various model landscapes with $L = 3$.** (A) The rescaled mean height $\bar{h}$ of the first fitness peak reached when starting from a uniformly chosen initial genotype is shown versus D for the Moran star for various values of α, in a landscape generated by the trade-off-induced landscape model [78, 79]. The rescaled mean height is defined as $(\bar{h}-h_{\min})/(h_{\max}-h_{\min})$, where $h_{\min}$ (resp. $h_{\max}$) denotes the fitness of the lowest peak (resp. highest peak). Recall that $h_{\min} < \bar{h} < h_{\max}$. Landscape parameters [78, 79]: dimensionless antibiotic concentration $c = 50$; exponent of the Hill function $a = 1$; r_i and m_i drawn from the distribution given in Eq. (S16) of the supplementary material of Ref. [54]. (B-F) Same as (A) but in a landscape generated by (B) the House of Cards model [72, 73], (C) the LK model with non-binary genotypes (with cardinal of the alphabet $A = 8$) [58], (D) the LKp-like model with $p = 0.5$ and $q = 0.1$ [74], (E) the Rough Mount Fuji model [32, 75] with $f_0' = 5$, $C = 0.8$ and epistatic contributions drawn from a standard normal distribution, (F) the Eggbox model [76] with fitnesses drawn from a Gaussian of mean $f_0 \pm \mu_E/2$, with $f_0 = 4$, $\mu_E = 4$, and a standard deviation of 1. See section S4 in the supplementary material of Ref. [54] for more details on these models. (G-L) Same as (A-F) but for the Wright-Fisher star walk. In all panels, lines are numerical resolution of the FSA equations, while markers are simulation results averaged over 10^5 to 10^7 walks per starting point (5×10^2 to 5×10^3 in panels (C) and (I) for the non-binary LK landscape that includes more genotypes than others). In all cases, $L = 3$, $C = 20$ and $g = 0.01$.

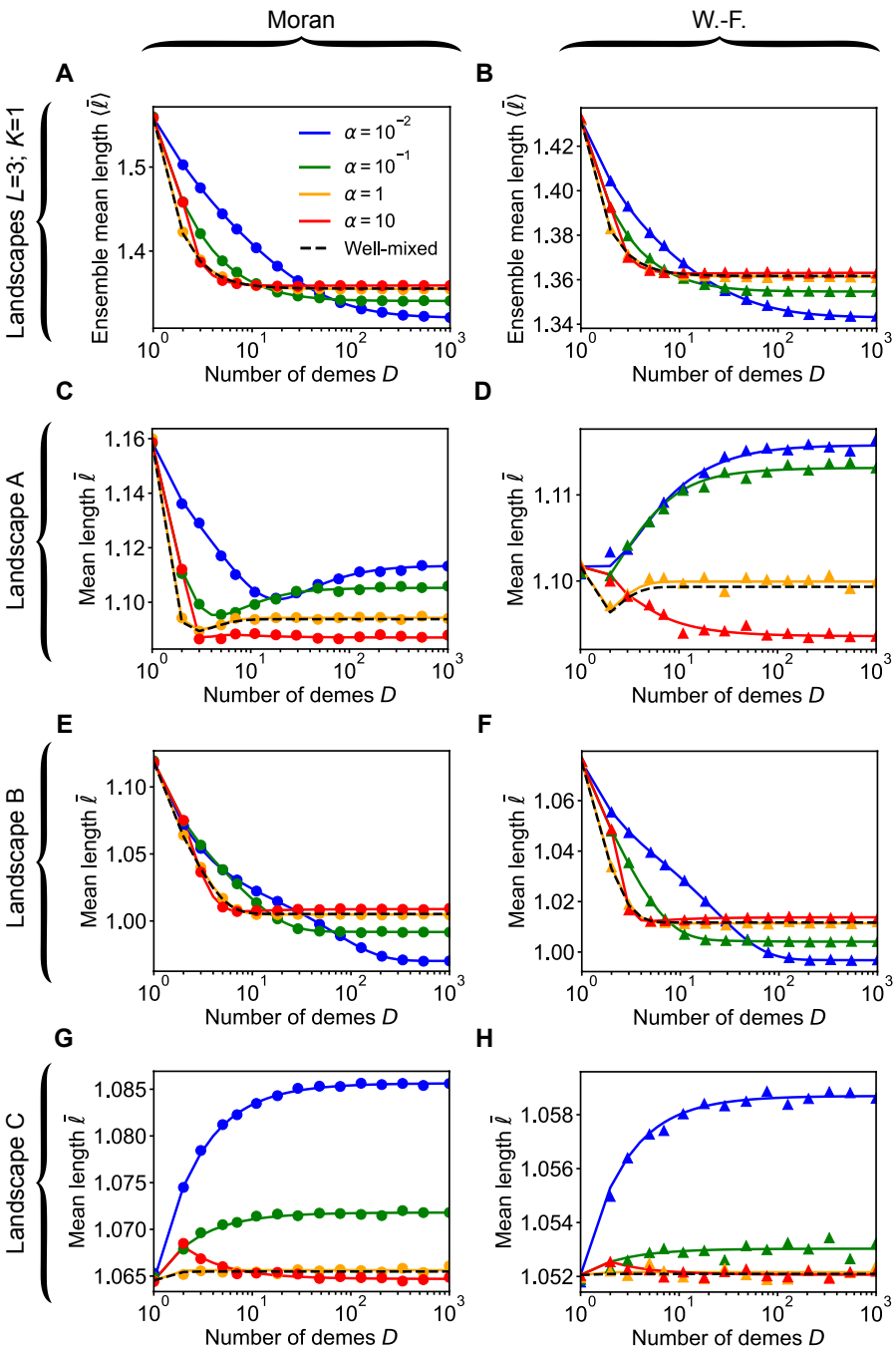

Figure S12: **Impact of the number of demes D of a star on the length ℓ of a walk, in LK landscapes with $L = 3$ and $K = 1$.** (A-B) The ensemble mean length $\langle \bar{\ell} \rangle$ of a walk (number of successful mutations before reaching the first fitness peak), when starting from a uniformly chosen initial genotype, is shown versus D for the Moran (left column) and Wright-Fisher (right column) star walks. Different values of migration asymmetry α are considered. In each case, the well-mixed case is shown for reference (dashed black line). Lines: numerical resolutions of the FSA equations for each landscape; markers: simulation results averaged over 10 walks per starting genotype in each landscape. The ensemble average is then performed over 2×10^5 landscapes. (C,E,G) The mean length $\bar{\ell}$ is shown versus D for the Moran star walk in landscapes A, B and C from Fig. 2(I-K). Same colors, markers and lines as in (A-B). (D,F,H) Same as (C,E,G) but for the Wright-Fisher star walk. Simulation results over landscapes A, B and C are averaged over at least 10^5 walks per starting genotype. In all cases, parameter values are $C = 20$ and $g = 0.01$.

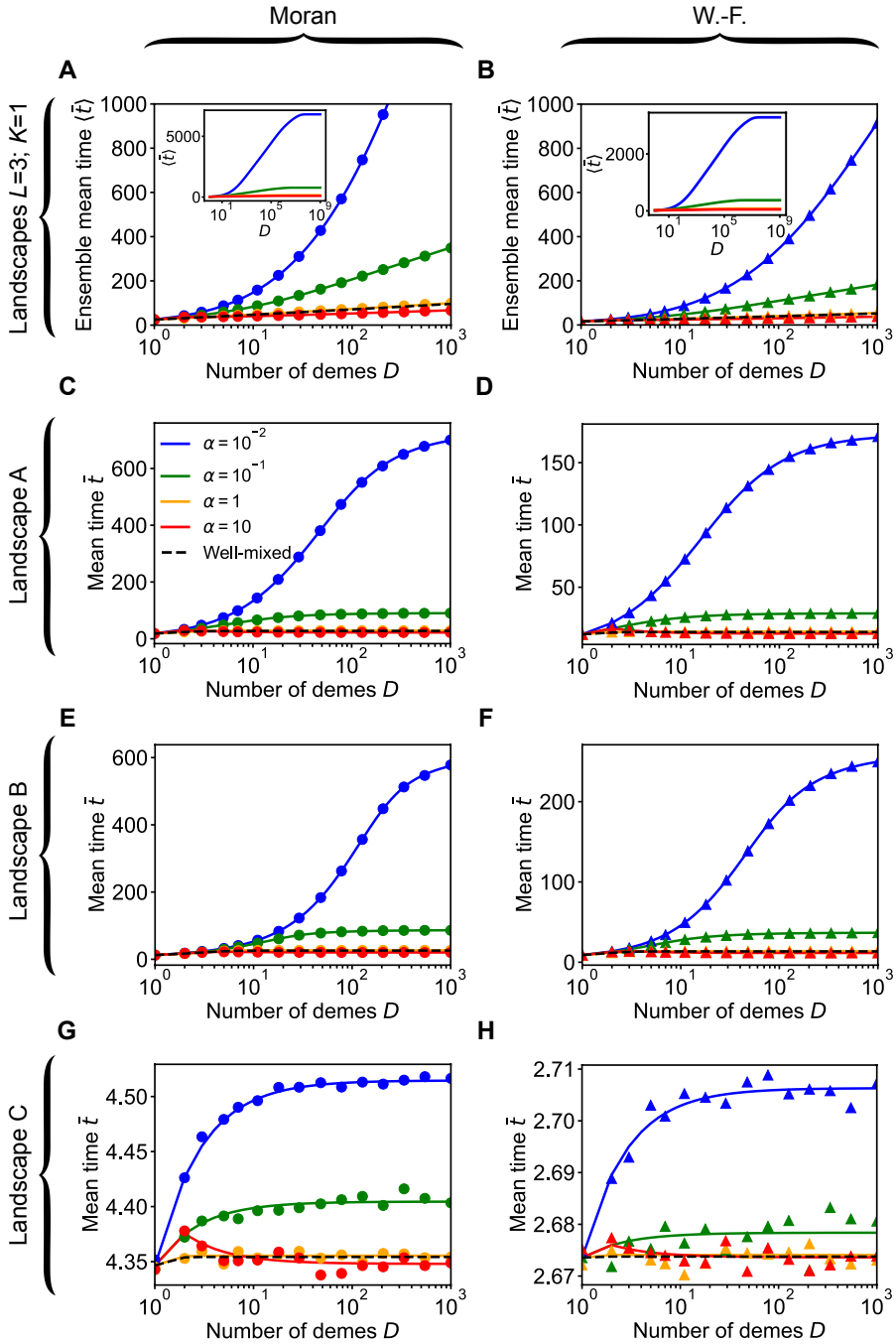

Figure S13: **Impact of the number of demes D of a star on the time t of a walk, in LK landscapes with $L = 3$ and $K = 1$.** Same as Fig. S12 but for the time t, which represents the number of all mutations (successful or not) before reaching a fitness peak (see main text).

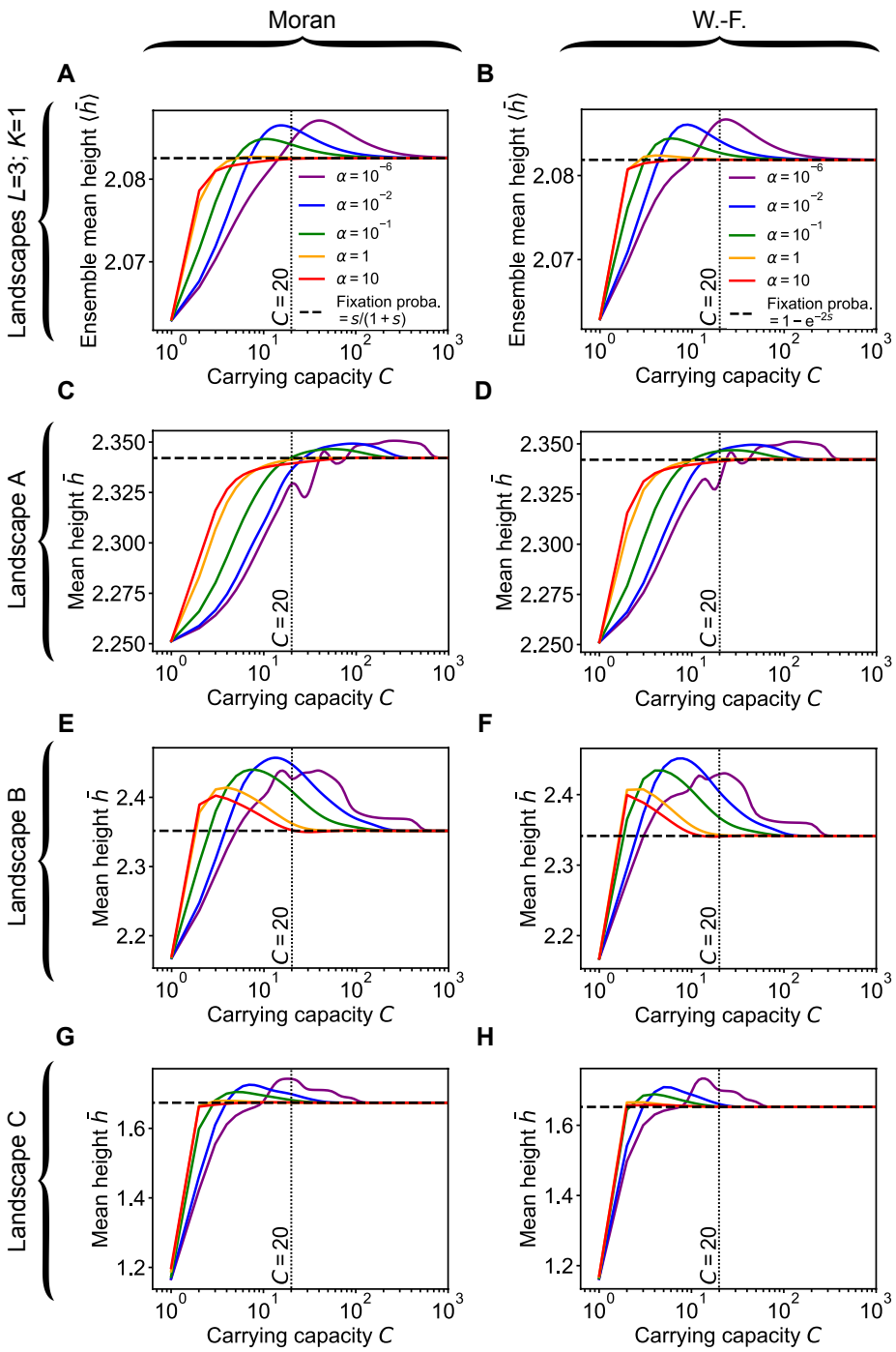

Figure S14: **Impact of the carrying capacity C of the demes of a star on early adaptation in LK landscapes with $L = 3$ and $K = 1$.** (A) Ensemble mean height $\langle \bar{h} \rangle$ of the first peak starting reached, starting from a uniformly chosen initial genotype, versus C for the Moran star walk with various migration asymmetry values α. Lines: numerical resolutions of the FSA equations for each landscape. The ensemble average is then performed over 2×10^5 landscapes. (B) Same but for the Wright-Fisher star walk. (C), (E) and (G) $\bar{h}$ versus C for the Moran star walk in landscapes A, B and C, respectively. (D), (F) and (H) Same as (C), (E) and (G) but for the Wright-Fisher star walk. In all panels, the vertical dotted line indicates $C = 20$, and parameter values are $D = 5$ and $g = 0.01$. The horizontal dashed lines correspond to $\bar{h}$ in the large-size limit of a well-mixed population either under the Moran model (left column) or the Wright-Fisher model (right column) (see Ref. [54]).

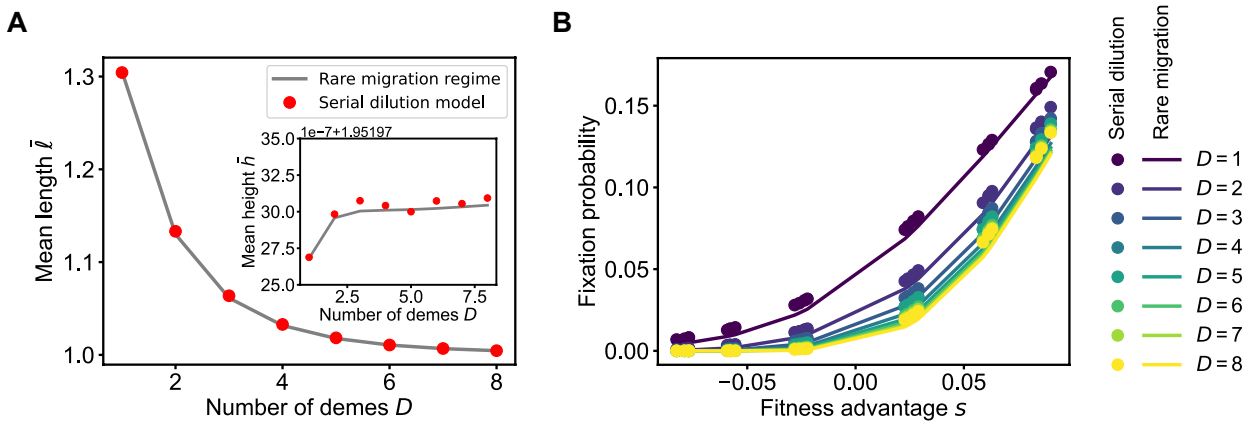

Figure S15: **Comparison between the Wright-Fisher star walk and the star walk in the serial dilution model, when migrations are rare.** (A) The mean length $\bar{\ell}$ (and mean height $\bar{h}$, see inset) of a walk before reaching a fitness peak when starting from a uniformly chosen initial genotype are shown versus D in a LK landscape with $L = 3$ and $K = 1$. This landscape was chosen among the generated ones because all mutations have small fitness effects (absolute relative selective advantages $|s| < 0.1$ for all possible mutations), so that the diffusion approximation is appropriate. Recall that this approximation is used in our treatment of the Wright-Fisher walk with rare migrations, see main text. Gray lines correspond to numerical resolutions of the FSA equations for the Wright-Fisher star walk with rare migrations (the main approach employed in this work), while red markers correspond to the walk under the serial dilution model of Ref. [16] with growth time $t = 1$, chosen to match the standard Wright-Fisher case. Specifically, the markers are obtained via numerical resolutions of the FSA equations, using direct simulations of the serial dilution model to estimate fixation probabilities (at least 10^6 replicates). Recall that in our usual approach, the analytical formula of the fixation probability in the structured population in the rare migration regime [13], using the diffusion approximation to analytically express the fixation probability in a deme under the Wright-Fisher model. (B) Fixation probabilities versus relative mutant fitnss advantage s, for different values of D. Lines and symbols are obtained as in (A). In addition, here, each marker corresponds to one value of s in the landscape considered in panel A. In all panels, parameter values are $\alpha = 0.1$ and $g = 0.01$. In the serial dilution model, we chose migration rates $m_I = 10^{-5}$ and $m_O = 10^{-4}$.

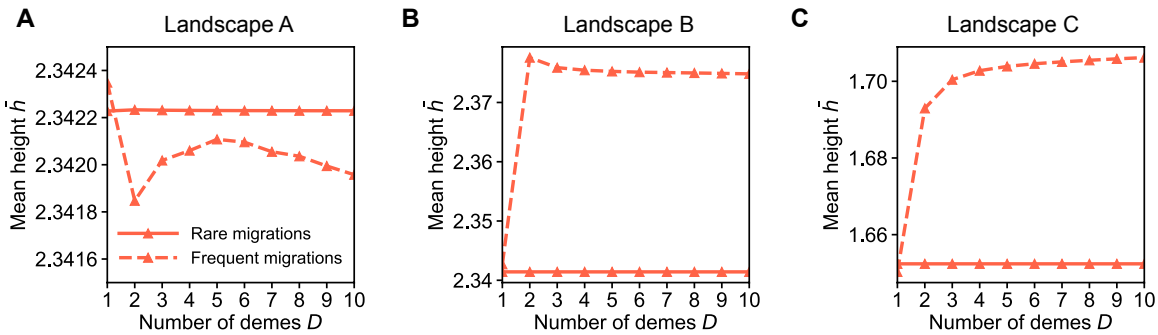

Figure S16: **Impact of the number of demes D of a star on early adaptation in three LK landscapes with $L = 3$ and $K = 1$ beyond the rare migration regime.** Same as Fig. 4 in the main text, but with larger demes: $C = 200$ instead of $C = 20$, and larger asymmetry: $\alpha = 10$.

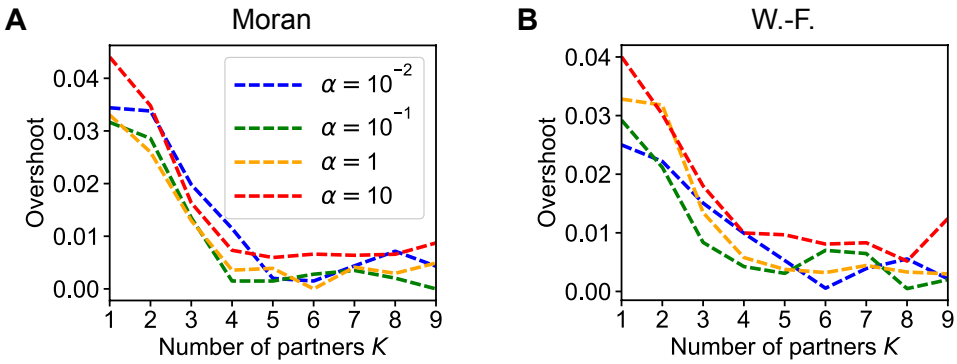

Figure S17: **Impact of the number of partners K on early adaptation in LK landscapes with $L = 10$.** (A) The overshoot of the ensemble mean height $\langle \bar{h} \rangle$ of the first fitness peak reached when starting from a uniformly chosen initial genotype is shown versus K for the Moran star walk, with various values of α. The overshoot of $\langle \bar{h} \rangle$ is defined as $\langle \bar{h} \rangle_{\max} - \langle \bar{h} \rangle_{D\to\infty}$, where $\langle \bar{h} \rangle_{\max}$ is the maximum value of $\langle \bar{h} \rangle$ versus D, while $\langle \bar{h} \rangle_{D\to\infty}$ is the large-D limit of $\langle \bar{h} \rangle$ (evaluated at $D = 10^5$ in practice), see [54]. (B) Same as (A) but for the Wright-Fisher star walk. In both cases, $L = 10$, $C = 20$ and $g = 0.01$.

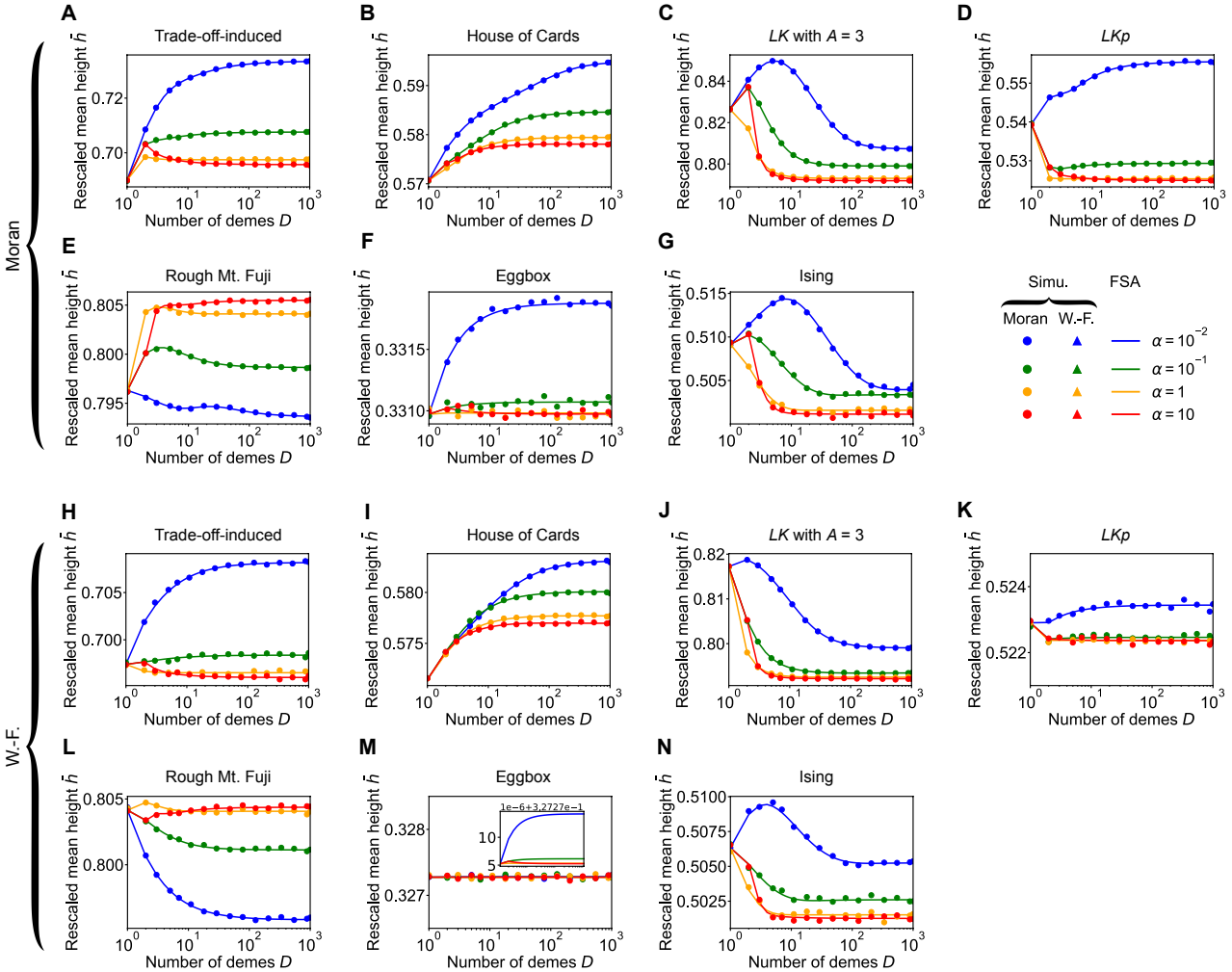

Figure S18: **Impact of the number of demes D and migration asymmetry α of a star on early adaptation in various model landscapes.** (A) The rescaled mean height of the first fitness peak reached when starting from a uniformly chosen initial genotype is shown versus D for the Moran star for various values of α, in a tradeoff-induced landscape [78, 79]. The rescaled mean height is defined as $(\bar{h} - h_{\min})/(h_{\max} - h_{\min})$, where $h_{\min}$ (resp. $h_{\max}$) denotes the fitness of the lowest peak (resp. highest peak). Recall that $h_{\min} < \bar{h} < h_{\max}$. As in Fig. S11, the tradeoff-induced landscape was generated with $c = 50$, $a = 1$, and r_i and m_i drawn from the same distribution as in [54]. (B-G) Same as (A) but in a landscape generated by (B) the House of Cards model [72, 73], (C) the LK model with non-binary genotypes (with alphabet size $A = 3$) [58], (D) the modified LKp model with $p = 0.5$ and $q = 0.1$ [74], (E) the Rough Mount Fuji model [32, 75] with $f_0' = 6$, $C = 0.5$, and where the epistatic contributions to fitness are drawn from a standard normal distribution, (F) the Eggbox model [76] where fitness are drawn from a Gaussian with mean 6 ± 2 and a standard deviation of 1, and (G) the Ising model [76, 77] with $B = 6$ [54] and where the couplings are drawn from a standard normal distribution. See section S4 in the supplementary material of Ref. [54] for more details on these models. (H-N) Same as (A-G) but for the Wright-Fisher star walk. In all panels, lines are numerical resolution of the FSA equations, while markers are simulation results averaged over at least 10^5 walks per starting point. Parameter values in all panels: $L = 6$, $C = 20$ and $g = 0.01$.

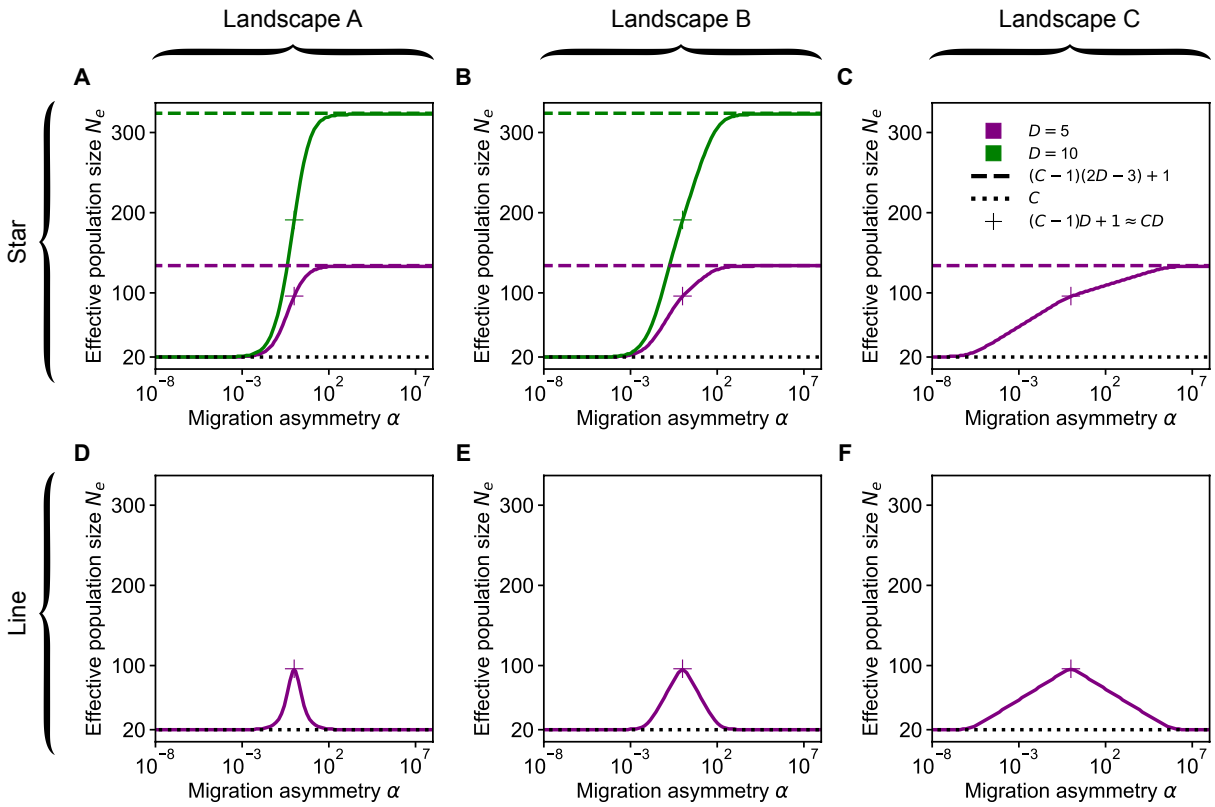

Figure S19: **Impact of migration asymmetry on steady-state effective population size in three LK landscapes with $L = 3$ and $K = 1$.** (A-C) Steady-state effective population size N_e of the star versus migration asymmetry α in landscape A (A), B (B) and C (C) (see Fig. 2(I-K)) with $D = 5$ (green) and $D = 10$ (purple). (D-F) Same as (A-C) but for the line with $D = 5$. Horizontal dashed and dotted lines: analytical asymptotic values; plus markers: analytical prediction for circulations (see Table 1 in the main text). Solid curves: numerical results. To obtain them, we developed the following method. First, in the fitness landscape of interest, we compute the steady-state distribution, using the eigenvector associated with eigenvalue 1 of the transition matrix of the Markov process describing the population's evolution [97]. We do this (i) for the structured population we consider, and (ii) for well-mixed populations with various population sizes N. Second, we compare the distributions obtained for (i) and (ii), using the Jensen-Shannon divergence. Finally, we estimate the steady-state effective population size N_e as the value of N that minimizes the Jensen-Shannon divergence. Note that in landscape C, for the star with $D = 10$ and large α, the population is almost always on the large peak, which makes it numerically challenging to estimate N_e. This is why this case is not shown in panel (C). Parameter values (in all panels): $C = 20$ and $g = 0.01$.

4 Supplementary table

Genotype	Fitness (landscape A)	Fitness (landscape B)	Fitness (Landscape C)
(0, 0, 0)	1.555	2.199	0.377
(0, 0, 1)	1.674	0.706	0.328
(0, 1, 0)	2.051	2.767	0.213
(0, 1, 1)	1.181	1.273	0.613
(1, 0, 0)	2.332	1.524	0.299
(1, 0, 1)	2.452	1.569	0.988
(1, 1, 0)	2.004	1.481	0.828
(1, 1, 1)	1.134	1.527	1.967

Table S1: **Specific LK fitness landscapes with $L = 3$ and $K = 1$ used throughout this work.** The first column shows genotypes, while the next ones display the corresponding fitness values (rounded to 3 decimal places) in the three landscapes A, B and C, shown in Fig. 2(I-K) and Fig. S4. These three landscapes were generated within the LK model with $L = 3$ and $K = 1$. Note that landscapes A and B are the same as those considered in Ref. [54]. While landscapes A and B were picked at random among LK fitness landscapes with $L = 3$ and $K = 1$, landscape C was selected because of its particular topography. To generate LK landscapes, epistatic partners were chosen randomly (in the random neighborhood scheme [42]) and fitness contributions were drawn from a uniform distribution between 0 and 1 (see Eq. (7) in the main text).

References

[1] F. Bertels, C. Leemann, K. J. Metzner, and R. R. Regoes. Parallel Evolution of HIV-1 in a Long-Term Experiment. *Mol Biol Evol*, 36(11):2400–2414, Nov 2019.

[2] G. P. Donaldson, S. M. Lee, and S. K. Mazmanian. Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol*, 14(1):20–32, Jan 2016.

[3] H. K. Mod, A. Buri, E. Yashiro, N. Guex, L. Malard, E. Pinto-Figueroa, M. Pagni, H. Niculita-Hirzel, J. R. van der Meer, and A. Guisan. Predicting spatial patterns of soil bacteria under current and future environmental conditions. *ISME J*, 15(9):2547–2560, Sep 2021.

[4] S. Wright. Evolution in Mendelian populations. *Genetics*, 16(2):97–159, 1931.

[5] M. Kimura and G. H. Weiss. The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. *Genetics*, 49(4):561–576, Apr 1964.

[6] T. Nagylaki. The strong-migration limit in geographically structured populations. *J Math Biol*, 9(2):101–114, Apr 1980.

[7] M. C. Whitlock and N. H. Barton. The effective size of a subdivided population. *Genetics*, 146(1):427–441, May 1997.

[8] M. Nordborg and S. M. Krone. Separation of time scales and convergence to the coalescent in structured populations. In *Modern Developments in Theoretical Population Genetics: The Legacy of Gustave Malécot*, page 194–232. Oxford University Press, Oxford, 2002.

[9] P. Sjödin, I. Kaj, S. Krone, M. Lascoux, and M. Nordborg. On the meaning and existence of an effective population size. *Genetics*, 169(2):1061–1070, Feb 2005.

[10] B. Houchmandzadeh and M. Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New Journal of Physics*, 13(7):073020, Jul 2011.

[11] B. Houchmandzadeh and M. Vallade. Exact results for fixation probability of bithermal evolutionary graphs. *Biosystems*, 112(1):49–54, Apr 2013.

[12] G. W. Constable and A. J. McKane. Population genetics on islands connected by an arbitrary network: an analytic approach. *J Theor Biol*, 358:149–165, Oct 2014.

[13] L. Marrec, I. Lamberti, and A.-F. Bitbol. Toward a universal model for spatially structured populations. *Physical Review Letters*, 127:218102, 2021.

[14] S. Yagoobi and A. Traulsen. Fixation probabilities in network structured meta-populations. *Scientific Reports*, 11(1):17979, 2021.

[15] S. Yagoobi, N. Sharma, and A. Traulsen. Categorizing update mechanisms for graph-structured metapopulations. *J R Soc Interface*, 20(200):20220769, Mar 2023.

[16] Alia Abbara and Anne-Florence Bitbol. Frequent asymmetric migrations suppress natural selection in spatially structured populations. *PNAS Nexus*, 2:2023–06, pgad392.

[17] T. Maruyama. On the fixation probability of mutant genes in a subdivided population. *Genetics Research*, 15(2):221–225, 1970.

[18] T. Maruyama. A simple proof that certain quantities are independent of the geographical structure of population. *Theoretical Population Biology*, 5(2):148–154, 1974.

[19] N. Barton. The probability of fixation of a favoured allele in a subdivided population. *Genet. Res.*, 62:149–157, 1993.

[20] M. C. Whitlock. Fixation Probability and Time in Subdivided Populations. *Genetics*, 164(2):767–779, 06 2003.

[21] Vitor Sudbrack and Charles Mullon. Fixation times of de novo and standing beneficial variants in subdivided populations. *Genetics*, 227(2):iyae043, 03 2024.

[22] E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, Jan 2005.

[23] T. Antal, S. Redner, and V. Sood. Evolutionary dynamics on degree-heterogeneous graphs. *Phys Rev Lett*, 96(18):188104, May 2006.

[24] K. Kaveh, N. L. Komarova, and M. Kohandel. The duality of spatial death-birth and birth-death processes and limitations of the isothermal theorem. *Royal Society Open Science*, 2(4):140465, 2015.

[25] L. Hindersin and A. Traulsen. Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics. *PLOS Computational Biology*, 11(11):1–14, 11 2015.

[26] K. Pattni, M. Broom, J. Rychtář, and L. J. Silvers. Evolutionary graph theory revisited: when is an evolutionary process equivalent to the Moran process? *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2182):20150334, 2015.

[27] S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics*, 1:356–366, 1932.

[28] J. Maynard Smith. Natural selection and the concept of a protein space. *Nature*, 225:563–564, 1970.

[29] A. Dawid, D. J. Kiviet, M. Kogenaru, M. de Vos, and S. J. Tans. Multiple peaks and reciprocal sign epistasis in an empirically determined genotype-phenotype landscape. *Chaos*, 20:026105, 2010.

[30] J. D. Bloom, L. I. Gong, and D. Baltimore. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science*, 328(5983):1272–1275, 2010.

[31] S. Kryazhimskiy, J. Dushoff, G.A. Bazykin, and J.B. Plotkin. Prevalence of epistasis in the evolution of influenza a surface proteins. *PLoS Genetics*, 7(2):e1001301, 2011.

[32] I. G. Szendro, M. F. Schenk, J. Franke, J. Krug, and J. A. G. M. de Visser. Quantitative analyses of empirical fitness landscapes. *J. Stat. Mech.: Theory Exp.*, 2013(1):P01005, 2013.

[33] J. A. Draghi and J. B. Plotkin. Selection biases the prevalence and type of epistasis along adaptive trajectories. *Evolution*, 67(11):3120–3131, 2013.

[34] A. W. Covert, R. E. Lenski, C. O. Wilke, and C. Ofria. Experiments on the role of deleterious mutations as stepping stones in adaptive evolution. *Proc. Natl. Acad. Sci. USA*, 110(34):E3171–E3178, 2013.

[35] C. Bank, S. Matuszewski, R. T. Hietpas, and J. D. Jensen. On the (un)predictability of a large intragenic fitness landscape. *Proceedings of the National Academy of Sciences*, 113(49):14085–14090, 2016.

[36] J. A. G. M. de Visser, S. F. Elena, I. Fragata, and S. Matuszewski. The utility of fitness landscapes and big data for predicting evolution. *Heredity*, 121:401–405, 2018.

[37] I. Fragata, A. Blanckaert, M. A. Dias Louro, D. A. Liberles, and C. Bank. Evolution in the light of fitness landscape theory. *Trends Ecol Evol*, 34(1):69–82, 01 2019.

[38] J. A. G. M. de Visser, T. F. Cooper, and S. F. Elena. The causes of epistasis. *Proceedings of the Royal Society B*, 278(1725):3617–3624, 2011.

[39] F. J. Poelwijk, S. Tănase-Nicola, D. J. Kiviet, and S. J. Tans. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *Journal of Theoretical Biology*, 272:141–144, 2011.

[40] J. A. G. M. de Visser and J. Krug. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics*, 15(7):480–490, 2014.

[41] J. A. G. M. de Visser, S.-C. Park, and J. Krug. Exploring the effect of sex on empirical fitness landscapes. *The American Naturalist*, 174(S1):S15–S30, 2009.

[42] S. Nowak and J. Krug. Analysis of adaptive walks on NK fitness landscapes with different interaction schemes. *Journal of Statistical Mechanics*, 2015(6):P06014, 2015.

[43] H. A. Orr. The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, 6:119–127, 2005.

[44] J. H. Gillespie. A simple stochastic gene substitution model. *Theoretical Population Biology*, 23(2):202–215, 1983.

[45] D. L. Hartl and E. W. Jones. *Genetics: principles and analysis.* Jones and Bartlett Publishers, 4th edition, 1998.

[46] H Allen Orr. A minimum on the mean number of steps taken in adaptive walks. *Journal of Theoretical Biology*, 220(2):241–247, 2003.

[47] K. Jain and J. Krug. Evolutionary trajectories in rugged fitness landscapes. *J. Stat. Mech.*, page P04008, 2005.

[48] Sijmen E Schoustra, Thomas Bataillon, Danna R Gifford, and Rees Kassen. The properties of adaptive walks in evolving populations of fungus. *PLoS Biology*, 7(11):e1000250, 2009.

[49] J. Franke, A. Klözer, J. A. G. M. de Visser, and J. Krug. Evolutionary accessibility of mutational pathways. *PLoS Comput. Biol.*, 7:e1002134, 2011.

[50] Su-Chan Park, Johannes Neidhart, and Joachim Krug. Greedy adaptive walks on a correlated fitness landscape. *Journal of Theoretical Biology*, 397:89–102, 2016.

[51] D. M. McCandlish. Visualizing fitness landscapes. *Evolution*, 65(6):1544–1558, 2011.

[52] D. M. McCandlish and A. Stoltzfus. Modeling evolution using the probability of fixation: history and implications. *The Quarterly Review of Biology*, 89(3):225–252, 2014.

[53] W. J. Ewens. *Mathematical population genetics: theoretical introduction.* Springer, 2004.

[54] Richard Servajean and Anne-Florence Bitbol. Impact of population size on early adaptation in rugged fitness landscapes. *Phil. Trans. R. Soc. B*, 20220045, 2023.

[55] Nikhil Sharma and Arne Traulsen. Suppressors of fixation can increase average fitness beyond amplifiers of selection. *Proceedings of the National Academy of Sciences*, 119(37):e2205424119, September 2022.

[56] Nikhil Sharma, Suman G. Das, Joachim Krug, and Arne Traulsen. Graph-structured populations elucidate the role of deleterious mutations in long-term evolution. *bioRxiv*, July 2024.

[57] Yang Ping Kuo and Oana Carja. Evolutionary graph theory beyond single mutation dynamics: on how network-structured populations cross fitness landscapes. *Genetics*, 227(2):iyae055, June 2024.

[58] M. Zagorski, Z. Burda, and B. Waclaw. Beyond the hypercube: evolutionary accessibility of fitness landscapes with realistic mutational networks. *PLoS Computational Biology*, 12(12):e1005218, 2016.

[59] S. F. Elena and R. E. Lenski. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, 4:457–469, 2003.

[60] K. Jain, J. Krug, and S. C. Park. Evolutionary advantage of small populations on complex fitness landscapes. *Evolution*, 65(7):1945–1955, Jul 2011.

[61] M. Lässig, V. Mustonen, and A. M. Walczak. Predicting evolution. *Nat Ecol Evol*, 1(3):77, Feb 2017.

[62] B. H. Good, M. J. McDonald, J. E. Barrick, R. E. Lenski, and M. M. Desai. The dynamics of molecular evolution over 60,000 generations. *Nature*, 551:45–50, 2017.

[63] P. A. P. Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

[64] M. Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713–719, 1962.

[65] J. F. Crow and M. Kimura. *An Introduction to Population Genetics Theory.* Blackburn, 2009 (first published in 1970).

[66] J. R. Norris. *Markov chains.* Cambridge University Press, 1997.

[67] D. Aldous and J. A. Fill. *Reversible Markov chains and random walks on graphs.* 2002 (recompiled version, 2014).

[68] G. Sella and A. E. Hirsh. The application of statistical physics to evolutionary biology. *Proceedings of the National Academy of Sciences*, 102(27):9541–9546, 2005.

[69] D. M. McCandlish. Long-term evolution on complex fitness landscapes when mutation is weak. *Heredity*, 121:449–465, 2018.

[70] Michael Manhart, Allan Haldane, and Alexandre V Morozov. A universal scaling law determines time reversibility and steady state of substitutions under selection. *Theoretical population biology*, 82(1):66–76, 2012.

[71] S. A. Kauffman and E. D. Weinberger. The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology*, 141(2):211–245, 1989.

[72] S. Kauffman and S. Levin. Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128(1):11–45, 1987.

[73] J. F. C. Kingman. A simple model for the balance between selection and mutation. *Journal of Applied Probability*, 15(1):1–12, 1978.

[74] L. Barnett. Ruggedness and neutrality - the NKp family of fitness landscapes. In *Artificial Life VI: Proceedings of the sixth international conference on Artificial life*, pages 18–27. MIT Press, 1998.

[75] T. Aita, H. Uchiyama, T. Inaoka, M. Nakajima, T. Kokubo, and Y. Husimi. Analysis of a local fitness landscape with a model of the rough Mt. Fuji-type landscape: application to prolyl endopeptidase and thermolysin. *Biopolymers*, 54(1):64–79, 2000.

[76] L. Ferretti, B. Schmiegelt, D. Weinreich, A. Yamauchi, Y. Kobayashi, F. Tajima, and G Achaz. Measuring epistasis in fitness landscapes: The correlation of fitness effects of mutations. *Journal of Theoretical Biology*, 396:132–143, 2016.

[77] B. Diu, D Lederer, and B. Roulet. *Physique statistique.* Hermann, 1989.

[78] S. G. Das, S. O. L. Direito, B. Waclaw, R. J. Allen, and J. Krug. Predictable properties of fitness landscapes induced by adaptational tradeoffs. *eLife*, 9:e55155, 2020.

[79] Suman G Das, Joachim Krug, and Muhittin Mungan. Driven disordered systems approach to biological evolution in changing environments. *Physical Review X*, 12(3):031040, 2022.

[80] Kelsi R Hall, Katherine J Robins, Elsie M Williams, Michelle H Rich, Mark J Calcott, Janine N Copp, Rory F Little, Ralf Schwörer, Gary B Evans, Wayne M Patrick, et al. Intracellular complexities of acquiring a new enzymatic function revealed by mass-randomisation of active-site residues. *Elife*, 9, 2020.

[81] E. R. Lozovsky, T. Chookajorn, K. M. Brown, M. Imwong, P. J. Shaw, S. Kamchonwongpaisan, D. E. Neafsey, D. M. Weinreich, and D. L. Hartl. Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proceedings of the National Academy of Sciences*, 106(29):12025–12030, 2009.

[82] S Brouillet, H Annoni, L Ferretti, and G Achaz. MAGELLAN: a tool to explore small fitness landscapes. *bioRxiv*, 2015.

[83] S. Kryazhimskiy, D. P. Rice, and M. M. Desai. Population subdivision and adaptation in asexual populations of Saccharomyces cerevisiae. *Evolution*, 66(6):1931–1941, Jun 2012.

[84] P. P. Chakraborty, L. R. Nemzer, and R. Kassen. Experimental evidence that network topology can accelerate the spread of beneficial mutations. *Evol Lett*, 7:447–456, 2023.

[85] M. Weigt, R. A. White, H. Szurmant, J. A. Hoch, and T. Hwa. Identification of direct residue contacts in protein-protein interaction by message passing. *Proc. Natl. Acad. Sci. U.S.A.*, 106(1):67–72, Jan 2009.

[86] D. S. Marks, L. J. Colwell, R. Sheridan, T. A. Hopf, A. Pagnani, R. Zecchina, and C. Sander. Protein 3D structure computed from evolutionary sequence variation. *PLoS ONE*, 6(12):e28766, 2011.

[87] F. Morcos, A. Pagnani, B. Lunt, A. Bertolino, D. S. Marks, C. Sander, R. Zecchina, J. N. Onuchic, T. Hwa, and M. Weigt. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc. Natl. Acad. Sci. U.S.A.*, 108(49):E1293–1301, Dec 2011.

[88] Francisco J Ayala and Cathryn A Campbell. Frequency-dependent selection. *Annual review of Ecology and Systematics*, 5(1):115–138, 1974.

[89] Ken-Ichi Kojima. Is there a constant fitness value for a given genotype? No! *Evolution*, 25(2):281–285, 1971.

[90] Alix Moawad, Alia Abbara, and Anne-Florence Bitbol. Evolution of cooperation in deme-structured populations on graphs. *Physical Review E*, 109(2):024307, 2024.

[91] Loïc Marrec. Quantifying the impact of genotype-dependent gene flow on mutation fixation in subdivided populations. *bioRxiv*, 2023.

[92] A. F. Bitbol and D. J. Schwab. Quantifying the role of population subdivision in evolution on rugged fitness landscapes. *PLoS Comput. Biol.*, 10(8):e1003778, Aug 2014.

[93] I. G. Szendro, J. Franke, J. A. G. M. de Visser, and J. Krug. Predictability of evolution depends nonmonotonically on population size. *Proceedings of the National Academy of Sciences*, 110(2):571–576, 2013.

[94] D. B. Weissman, M. M. Desai, D. S. Fisher, and M. W. Feldman. The rate at which asexual populations cross fitness valleys. *Theor. Pop. Biol.*, 75:286–300, 2009.

[95] Mark Broom and Jan Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 464(2098):2609–2627, 2008.

[96] James W Miller. A matrix equation approach to solving recurrence relations in two-dimensional random walks. *Journal of Applied Probability*, 31(3):646–659, 1994.

[97] Dianne P. O'Leary. Iterative methods for finding the stationary vector for Markov chains. In Carl D. Meyer and Robert J. Plemmons, editors, *Linear Algebra, Markov Chains, and Queueing Models*, pages 125–136, New York, NY, 1993. Springer New York.