Population subdivision with migration can facilitate evolution on rugged fitness landscapes

Anne-Florence Bitbol and David J. Schwab
Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA
(Dated: May 11, 2014)

We show that subdivision of an asexual population into demes connected by migration significantly accelerates the crossing of fitness valleys and plateaus over a wide parameter range, both with respect to the non-subdivided population and with respect to a single deme. We predict the existence of a parameter range where valley or plateau crossing by the metapopulation is as fast as that of the fastest deme, and we verify this prediction using stochastic simulations. Our results hold not only for deleterious and neutral intermediates, but also for weakly beneficial ones. Finally, we extend our work to the case of a large population connected by migration to one or several smaller islands.

Natural selection drives organisms towards higher fitness (i.e. reproductive success), but crossing fitness valleys or plateaus may be necessary to make progress up a fitness landscape. Fitness valleys can arise from interactions between genes, called epistasis: for instance, two mutations together can yield a benefit while each alone is detrimental \cite{1,2}. While the high dimensionality of genotype space makes it challenging to probe the structure of fitness landscapes \cite{3,4}, evidence has been accumulating for frequent epistasis and landscape ruggedness \cite{11,12,13}. Additionally, fitness valleys often occur in the evolution of antibiotic resistance by bacteria \cite{9,10}. Population structure can play an important role in evolution \cite{11,12,13,14,15}. The time taken to cross a valley or a plateau depends on population size, since stochastic effects such as genetic drift have an increased importance in small populations, allowing neutral and deleterious mutations to fix with increased probability \cite{20,21}. Hence, while beneficial mutations spread faster in larger populations, population subdivision into demes can allow the maintenance of larger genetic diversity and facilitate valley or plateau crossing locally. Migration can then spread beneficial mutations. This idea was first discussed by Wright in his shifting balance theory \cite{22,23}. Here, we investigate subdivision with migration alone, without additional effects such as extinction and refounding of demes, which played an important part in Wright’s theory, specific geographic structure \cite{11,15}, or spatially heterogeneous environments \cite{16,17,18,19}, on which much recent work focuses.

Studying quantitatively the effect of subdivision on evolution in a rugged fitness landscape can help in inferring the statistics of epistasis from evolution experiments \cite{24}. Work on structured populations has been used as qualitative proofs of landscape ruggedness \cite{11}. Current experiments investigating the evolution of subdivided populations at various migration rates have produced mixed results, some finding evidence for faster adaptation of metapopulations \cite{25}, and others not \cite{26}.

It is therefore practically important to determine conditions under which subdivision is beneficial. Additionally, population subdivision is widespread and can play a crucial role in events such as speciation \cite{27}. For instance, evidence has recently been found for compartmentalization and subspecies of HIV in different organs of a single patient \cite{28,29}.

We investigate under what conditions demes of identical size coupled through migration cross fitness valleys and plateaus faster than a single population. We show that subdivision significantly accelerates fitness valley or plateau crossing over a wide parameter range, both with respect to a non-subdivided population and to a single deme. We predict the existence of a parameter range where valley or plateau crossing by a metapopulation is as fast as that of its fastest deme, and verify this prediction using stochastic simulations. Similar results hold for slightly beneficial (but effectively neutral) intermediates. Finally, we extend our work to a population connected by migration to one or several smaller population islands.

We consider asexual individuals, characterized by their genotype and associated fitness \(f\), which determines division rate. In our simulations, the carrying capacity \(K\) of a deme is fixed (see Supplementary Material), and each individual divides at a rate \(f(1-N/K)\), where \(N\) is the number of individuals in the deme to which the individual belongs: this corresponds to logistic growth. Simultaneously, each individual dies at a constant rate \(d\). Hence, \(N\) fluctuates weakly around its equilibrium value of \(K(1-d/f)\). In practice, we choose \(d=0.1\), and fitnesses of order one, thus the deme size is \(N\approx 0.9K\). In our analytic estimates, we treat \(N\) as constant, and denote the equal death and average division rates by \(d\). We consider an ensemble of \(\Pi\) demes of identical size, and treat migration as a random exchange of two individuals between two different demes, occurring at a rate of \(2m\) per individual (see Fig. \(1\)a)). There is no geographic structure in our model, i.e. exchange between any two demes is equally likely. We focus on the simplest fitness valley, involving three successive genotypes, denoted by 0, 1 and 2 (see Fig. \(1\)b)). The initial genotype is taken as reference for fitness: \(f_0 = 1\). The subsequent genotypes have fitness \(f_1 = 1-\delta\) and \(f_2 = 1+s\), with \(s>0\): the final

\(@afbitbol@princeton.edu@\)
\@dschwab@princeton.edu@
mutation is thus beneficial. For a fitness valley, $\delta > 0$, while for a plateau, $\delta = 0$. We also consider the case of a weakly beneficial intermediate ($\delta < 0$) that remains effectively neutral. We only allow forward mutations, and note that including back mutations only changes crossing times by a numerical factor \[21\]. Finally, for the sake of simplicity, we assume that all mutations have probability $\mu$ per division, but it is straightforward to generalize our work to different mutation rates.

We are interested in the average time $\tau_m$ required for the whole metapopulation to cross the fitness valley, i.e. to fix mutation ‘2’ in all demes, starting from an initial state where all individuals have genotype ‘0’. Fig. I(c) shows an example plot of $\tau_m$ from simulations, as a function of the migration rate parameter, $m$ (see Supplementary Material for simulation methods). We observe a broad range of migration rates for which the metapopulation crosses the valley faster than both (i) a single non-subdivided population with the same carrying capacity as the entire metapopulation, and (ii) a single isolated deme. This example shows that subdivision with migration can significantly accelerate fitness valley crossing.

Successful valley or plateau crossing can occur either through sequential fixation of mutations, or by stochastic tunneling, where the beneficial mutation arises in a small fluctuating minority population of first-mutants that never fixes. Sequential fixation dominates for small populations, while tunneling is faster for larger ones \[21\]. In the regime dominated by sequential fixation, valley (resp. plateau) crossing time increases (resp. remains constant) with the number $N$ of individuals in the population, while in the regime dominated by tunneling it decreases as $1/N$ \[21\]. Here, we assume that isolated demes are in the sequential fixation regime, but do not make any such assumption for the total population. We will show later that if isolated demes are in the tunneling regime, there is no benefit of subdivision.

In the best case, the crossing time of the whole metapopulation is determined by that of the best deme, i.e. the one that crosses the fitness valley or plateau fastest. We now determine the conditions under which this occurs, focusing on migration rates much smaller than division and death rates, $2m \ll d$, such that fixation or extinction of a mutant lineage in a deme is not perturbed by migration. First, migration should be rare enough for demes to remain effectively shielded from migration events while they have fixed the intermediate mutation, until the final beneficial mutation arises. Second, migration should also be frequent enough for the spreading time of the final beneficial mutation through the whole metapopulation to be negligible with respect to the crossing time of the best deme. These two criteria provide upper and lower bounds on the ratio of migration rate to mutation rate for which subdivision is most beneficial, which we now express explicitly.

(I) Upper bound on the ratio of migration rate to mutation rate. The average time $\tau_{12}$ required for a deme of 1-mutants to fix the beneficial mutation ‘2’ is $\tau_{12}^{-1} = N\mu p_{12} \[21\]$, where $N\mu$ is the total mutation rate in the deme, and $p_{12} = \left(1-e^{-(s+\delta)}\right)/\left(1-e^{-N(s+\delta)}\right)$ denotes the probability of fixation of a single mutant with genotype ‘2’ in a background of 1-mutants \[30\]. $\tau_{12}$ must be smaller than the average time, $t_d$, for a deme of 1-mutants to be wiped out by migration from the other demes, which still have genotype ‘0’. The total rate of migration events in the metapopulation is $\Pi N m$, so that $t_d = n_d/(\Pi N m)$, where $n_d$ is the average number of migration events required for the 1-mutants to be wiped out by migration.

To determine $n_d$, we study the evolution of the number $i \in \{0, \Pi\}$ of demes that have fixed genotype ‘1’, while other demes have genotype ‘0’. The possible outcomes of a migration event and their probabilities depend on the current value of $i$ but not on its history, and $i \in \{0, \Pi\}$ are absorbing states, so $i$ evolves according to a finite Markov chain. The only migration events that can affect $i$ are those that exchange individuals with different genotypes. As a result of one such relevant migration event, $i$ can increase by one with probability $p(1-p')$, where $p$ denotes the probability that the ‘1’ migrant fixes in the ‘0’ deme, and $p'$ the probability that the ‘0’ migrant fixes in the ‘1’ deme. Both $p$ and $p'$ can be expressed in a similar fashion as $p_{12}$ above \[30\]. Similarly, $i$ decreases by one with probability $p'(1-p)$, and otherwise does not change. These probabilities, multiplied by the probability that a migration event is relevant, $2i(\Pi - i)/[\Pi(\Pi - 1)]$, yield the transition matrix of our finite Markov chain, which is tri-diagonal (or continued). Hence, $n_d$, the average number of steps until absorption starting from $i = 1$, conditional on absorption at $i = 0$, can be expressed analytically \[30\]. Its full expression is provided in the Supplementary Material, for valleys ($\delta > 0$) and plateaus.
we expect the valley/plateau crossing time of the whole metapopulation to be dominated by the crossing time of the best deme, so that $\tau_m/\tau_{id} \approx 1/\Pi$.

With the parameter values from Fig. 1, the interval of Eq. (3) is $6.4 \times 10^{-9} \ll m \ll 6.2 \times 10^{-7}$. The broad minimum in Fig. 1 agrees remarkably well with this theoretically predicted optimal interval. Moreover, the simulation results in Fig. 1 yield a minimum $\tau_m = (4.9 \pm 0.8) \times 10^7$ (for $m = 10^{-7}$), while $\tau_{id} = (4.8 \pm 0.8) \times 10^8$: it satisfies $\tau_m/\tau_{id} \approx 1/10 = 1/\Pi$.

Our above analysis requires isolated demes to be in the sequential fixation regime. If instead they are in the tunneling regime, their crossing time involves a single rate $\delta$. Hence, in the ideal case where crossing by the metapopulation is dominated by that of the best deme, we again expect $\tau_m/\tau_{id} \approx 1/\Pi$. Since the average crossing time via tunneling is proportional to $1/\Pi$, this would yield $\tau_m/\tau_{ns} \approx 1$, where $\tau_{ns}$ is the average crossing time of a non-subdivided population (with IIN individuals). Thus, isolated demes being in the sequential fixation regime is a necessary condition for subdivision to accelerate fitness valley/plateau crossing.

![FIG. 2. (a) Heatmap of the ratio $\tau_m/\tau_{id}$ of the average valley crossing time of a metapopulation to that of an isolated deme, with $\Pi = 10$ and $K = 50$, as a function of valley depth $\delta$ and migration-to-mutation ratio $m/(\mu d)$. Parameter values same as Fig. 1 except that $\delta$ and $m$ are now varied. All numerical results averaged over 100 runs, and the heatmap is interpolated. Solid lines: bounds of the interval in Eq. (3)](image)
a metapopulation, compared to an isolated deme and to a non-subdivided population, as a function of the ratio of migration rate to mutation rate, \( m/(\mu d) \), and of the fitness valley depth, \( \delta \). Fig. 2(a) shows that the interval of Eq. 3 describes well the region where \( \tau_m/\tau_{id} \) is smallest and tends to 1/2. For lower migration rates, \( \tau_m/\tau_{id} \) increases. Indeed, if \( m = 0 \), \( \tau_m \) is determined by the valley crossing time of the slowest independent deme. Conversely, for high migration rates, \( \tau_m \) tends to the non-subdivided case, \( \tau_{ns} \) (see Fig. 1(c)). As \( \delta \) increases, the critical population size above which tunneling dominates over sequential fixation decreases \[21\]. Due to this, above a critical value of \( \delta \) (dashed line in Fig. 2), \( \tau_{ns} \) becomes smaller than \( \tau_{id} \), in which case large values of \( m \) give a low \( \tau_m/\tau_{id} \) (see Fig. 2(a)). Above another, larger critical value of \( \delta \) (dotted line in Fig. 2), isolated demes are in the tunneling regime. In Fig. 2(b), which plots the ratio \( \tau_m/\tau_{ns} \), we see that when \( \delta \) sufficiently exceeds the latter threshold, the metapopulation no longer crosses the valley faster than the non-subdivided one, as predicted above. Fig. 2 shows that, for the parameter values chosen, subdivision accelerates valley crossing over a large range of valley depths and migration rates, and that the metapopulation can cross the valley orders of magnitude faster than a single large population.

As mentioned previously, our results hold for fitness plateaus, such that mutation ‘1’ is neutral (\( \delta = 0 \)), of relevance for neutral mutation networks \[22\]. Moreover, for \( |\delta| < \max(\sqrt{8N},1/N) \), mutation ‘1’ is effectively neutral in a finite population with \( N \) individuals \[21\]. Hence, our arguments hold even for weakly beneficial intermediates, as demonstrated in Fig. 3(a). Thus, while sufficient magnitude epistasis is needed for subdivision to be beneficial, there is no stringent requirement for sign epistasis.

Thusfar, we focused on demes of equal size for simplicity, but demes of different sizes are relevant in practice. Let us consider a large population connected by migration to \( S \) smaller satellite populations of identical size, assumed to be in the sequential fixation regime. Migration only occurs between the large population and one of the smaller islands (chosen randomly at each migration event), and the total migration rate is denoted by \( M \). It is straightforward to adapt our work to this case (see Supplementary Material). We again find an interval of \( M/(\mu d) \) over which the crossing time for the large population is dominated by the crossing time of the best island. This is corroborated by our simulations (see Fig. 3(b)).

For subdivision to accelerate valley or plateau crossing, isolated demes must be small enough to be in the sequential fixation regime. Although this hypothesis looks stringent, the transition between sequential fixation and tunneling occurs at larger population sizes if the mutation probability per division \( \mu \) and the valley parameters \( s \) and \( \delta \) are smaller \[21\]. For \( E. Coli \), the mutation probability per base pair per division is \( \mu = 8.9 \times 10^{-11} \) \[33\]. Consider a total population size of \( 10^6 \) individuals subdivided into 100 demes, which is experimentally reasonable \[24\] \[31\]. For \( s = 10^{-2} \) and \( \delta = 10^{-3} \), isolated demes are in the sequential fixation regime. Provided that \( 2m < d \) (see above), Eq. 3 predicts an optimal effect of subdivision for \( 2.3 \times 10^{-3} \leq m/(\mu d) \leq 5.5 \), a wide range, where \( \tau_m \approx \tau_{id}/100 \) \((= 2.5 \times 10^{12})\). Comparing this to the valley crossing time of the non-subdivided population \( \tau_{ns} = 1.3 \times 10^{14} \) demonstrates a significant benefit of subdivision for valley crossing: \( \tau_m/\tau_{ns} = 2.0 \times 10^{-2} \).

Our quantitative assessment of the conditions under which subdivision significantly speeds up valley or plateau crossing can aid in inferring characteristics of fitness landscapes from evolution experiments, such as those in Refs. \[25\] \[26\], and help optimally design future experiments. Further directions include investigating the evolution of a metapopulation with a distribution of deme sizes on a more general rugged landscape, as well as assessing the impact of specific geographic structure. Our work could also be extended to sexual populations, where recombination plays an important role in valley or plateau crossing \[35\]. The interplay between recombination and subdivision, which respectively alleviate and exacerbate clonal interference, will be interesting to study.

We thank Ned S. Wingreen for stimulating discussions. This work was supported in part by National Science Foundation Grant PHY-0957757 and National Institute of Health Grant R01 GM082938. DJS was supported by
I. NUMBER OF MIGRATION EVENTS FOR EXTINCTION OR SPREADING IN A METAPOPULATION

In the main text, we presented upper and lower bounds on the ratio of migration rate to mutation rate, defining an interval over which subdivision significantly reduces valley or plateau crossing time (see Eqs. (1-3)). The upper bound involves \( n_d \), the average number of migration events required for the ‘1’-mutants to be wiped out by migration, starting from a state where one deme has fixed genotype ‘1’, while all other demes have genotype ‘0’. Similarly, the lower bound involves \( n_s \), the average number of migration events required for the ‘2’-mutants to spread by migration to the whole metapopulation, starting from a state where one deme has fixed genotype ‘2’, while all other demes have genotype ‘0’. In this section, we provide explicit expressions for \( n_d \) and \( n_s \), both for fitness plateaus and for fitness valleys. Throughout this section, we consider a metapopulation of \( \Pi \) demes composed of \( N \) individuals each, and we assume that individual demes are in the sequential fixation regime (see the main text for a discussion of this point).

As mentioned in the main text, \( n_s \) and \( n_d \) can be determined from finite Markov chain theory. For this, we study the evolution of the number \( i \in [0, \Pi] \) of demes that have fixed the mutant genotype ‘1’ for the calculation of \( n_d \); ‘2’ for that of \( n_s \), while other demes have genotype ‘0’: \( i \) evolves according to a finite Markov chain, each step being a migration event, and \( i \in [0, \Pi] \) are the two absorbing states. The transition matrix of this Markov chain is contingent (or tri-diagonal), with

\[
P_{i \to i+1} = \frac{2(i(\Pi-i))}{\Pi(\Pi-1)}p(1-p'), \quad (4)\]
\[
P_{i \to i-1} = \frac{2((\Pi-i))}{\Pi(\Pi-1)}p'(1-p), \quad (5)\]
\[
P_{i \to 0} = 1 - P_{i \to i+1} - P_{i \to i-1}, \quad (6)\]

for \( i \in [1, \Pi - 1] \), and \( P_{0 \to 0} = P_{\Pi \to \Pi} = 1 \). Here, \( P_{j \to k} \) denotes the probability that \( i \) varies from \( j \) to \( k \) as the final outcome of one migration event (i.e., the outcome after fixation or extinction of both migrants’ lineages in the demes where they have migrated). In these expressions, \( 2i(\Pi-i)/\Pi(\Pi-1) \) is the probability that a migration event exchanges individuals with different genotypes, while \( p \) denotes the probability that the mutant migrant fixes in the ‘0’ deme, and \( p' \) the probability that the ‘0’ migrant fixes in the mutant deme.

Here, we do not account for independent mutations arising and fixing in other demes during the process of spreading (or extinction) of the mutant’s lineage in the metapopulation. Indeed, our aim is to compare the timescales of migration and mutation processes, and thus treat them separately. Note that, in practice, this hypothesis is reasonable if mutations that fix are sufficiently rarer than migration events. We also consider that the time between two successive migration events is large enough for fixation to occur in the demes affected by migration before the next migration event occurs, which is true in the low-migration rate regime \( (2m \ll d, \text{ where } \Pi = 2m \text{ is the migration rate per individual, while } d \text{ is the death rate and average division rate per individual}) \) that we study in our work.

\( n_s \) and \( n_d \) can be directly expressed as the average number of steps until absorption in a particular absorbing state. Let us present general expressions of these average numbers of steps, before using them to obtain explicit expressions of \( n_s \) and \( n_d \). These expressions of \( n_d \) and \( n_s \) will enable us to study the parameter dependence of the bounds of the interval of \( m/(\mu d) \) over which subdivision significantly accelerates valley or plateau crossing (see Eqs. (1-3) of the main text).

A. Some results regarding finite Markov chains with contingent probability matrices [30]

We are interested in the average number of steps \( \nu_a \) until the system reaches each of the absorbing states \( a \in \{0, \Pi\} \), starting from the state \( i = 1 \):

\[
\nu_a = \sum_{j=1}^{\Pi-1} s_{j,a}, \quad (7)
\]

where \( s_{j,a} \) is the average number of steps that the system spends in the state \( i = j \) before absorption, given that it starts in the state \( i = 1 \) and finally absorbs in state \( i = a \). It can be expressed as [30]

\[
s_{j,a} = \frac{\pi_{j,a}}{\pi_{1,a}} \nu_{j,a}, \quad (8)
\]

where \( s_j \) is the average number of steps the system spends in state \( i = j \) before absorption in either of the two absorbing states, given that it started in state \( i = 1 \), and \( \pi_{j,a} \) is the probability that the system finally absorbs in state \( i = a \) if it starts in state \( i = j \).

Using the explicit expressions given in Ref. [30] for \( s_j \) and \( \pi_{j,a} \) for the case of a contingent probability matrix, we obtain:

\[
\nu_0 = \sum_{j=1}^{\Pi-1} \left( \frac{\sum_{k=0}^{j-1} \rho_k}{\sum_{k=0}^{j-1} \rho_k} \right) \left( \frac{\sum_{k=1}^{\Pi-1} \rho_k}{\sum_{k=1}^{\Pi-1} \rho_k} \right) \rho_j P_{j \to j+1}, \quad (9)
\]
\[
\nu_\Pi = \sum_{j=1}^{\Pi-1} \left( \frac{\sum_{k=0}^{j-1} \rho_k}{\sum_{k=0}^{j-1} \rho_k} \right) \rho_j P_{j \to j+1}, \quad (10)
\]
where we have introduced

$$\rho_k = \prod_{i=1}^{k} \frac{P_{i \rightarrow i-1}}{P_{i \rightarrow i+1}}. \quad (11)$$

### B. Explicit expression of $$n_d$$

$$n_d$$ corresponds to $$\nu_0$$, where $$p$$ is the probability that a '1'-mutant fixes in a deme of '0' individuals and $$p'$$ is the probability that a '0'-individual fixes in a deme of '1'-mutants. Hence, it can be expressed explicitly from Eqs. (4), (1), and (5). Once the expressions of $$p$$ and $$p'$$ depend whether mutation '1' is neutral or deleterious, we obtain different expressions for the fitness plateau and for the fitness valley.

#### 1. Fitness plateau

For a fitness plateau (i.e. a neutral intermediate '1'), $$p = p' = 1/N$$, where $$N$$ is the number of individuals per deme. Hence,

$$P_{i \rightarrow i+1} = P_{i \rightarrow i-1} = \frac{2i(N-i)(N-1)}{N(N-1)2^N}, \quad (12)$$

which implies that $$\rho_k = 1$$ for all $$k$$ (see Eq. (11)). Thus, Eq. (9) yields

$$n_d = \frac{N^2 \Pi}{2(N-1)} \sum_{j=2}^{\Pi} \frac{1}{j}. \quad (13)$$

#### 2. Fitness valley

For a fitness valley, $$p$$ is the probability of fixation of the deleterious '1'-mutant, with fitness $$1 - \delta$$, in a deme where all other individuals have genotype '0' and fitness 1. It can readily be shown [21] that $$p$$ takes the form

$$p = p_{01} = \frac{1 - e^{\delta}}{1 - e^{\Pi \delta}}. \quad (14)$$

Similarly,

$$p' = p_{10} = \frac{1 - e^{-\delta}}{1 - e^{-\Pi \delta}}. \quad (15)$$

Hence, Eqs. (11) and (4, 5) yield $$\rho_k = r^k$$, with

$$r = \frac{(1 - p)p'}{(1 - p')p}. \quad (16)$$

and Eq. (9) gives:

$$n_d = \frac{\Pi(\Pi - 1)}{2(r - r^\Pi)} \frac{\sum_{j=1}^{\Pi-1} (r^j - r^\Pi)^2}{(1 - r^\Pi)(1 - r^j)(\Pi - j)}. \quad (17)$$

### C. Explicit expression of $$n_s$$

$$n_s$$ corresponds to $$\nu_1$$, where $$p$$ is the probability that a '2'-mutant fixes in a deme of '0' individuals and $$p'$$ is the probability that a '0'-individual fixes in a deme of '2'-mutants. Hence, it can be expressed explicitly from Eqs. (10), (4), and (5). For $$n_s$$, there is no difference between the valley and the plateau, since genotype '1' is not involved.

Since the fitness of '2'-mutants is 1 + $$s$$ while that of genotype-0' individuals is 1, we have

$$p = p_{02} = \frac{1 - e^{-s}}{1 - e^{-N_s}}, \quad (18)$$

and

$$p' = p_{20} = \frac{1 - e^{-s}}{1 - e^{-N_s}}. \quad (19)$$

As above, Eqs. (11), (4), and (5) yield $$\rho_k = r^k$$, with $$r$$ defined in Eq. (16). Thus, Eq. (11) gives

$$n_s = \frac{\Pi(\Pi - 1)}{2(1 - r)(1 - r^\Pi)(1 - r^j)p} \sum_{j=1}^{\Pi-1} \frac{(1 - r^j)(1 - r^\Pi - j)}{(\Pi - j)}. \quad (20)$$

### D. Simplified expressions for upper and lower bounds of $$m/(\mu d)$$

In the main text, we have shown that the benefit of subdivision is highest when $$m/(\mu d)$$ is situated between a lower bound,

$$L = n_s p_{01}, \quad (21)$$

and an upper bound,

$$U = \frac{n_d p_{12}}{\Pi}, \quad (22)$$

(see Eqs. (1-3) of the main text), where

$$p_{12} = \frac{1 - e^{-\Pi(s+\delta)}}{1 - e^{-N(s+\delta)}}. \quad (23)$$

denotes the probability of fixation of a single mutant with genotype '2' in a background of '1'-mutants. Now that we have given general explicit expressions for $$n_s$$ and $$n_d$$, we present simplifications of these expressions in relevant parameter regimes. This will yield better insight into the dependence of $$L$$ and $$U$$ on the various parameters.

Throughout this section, we will focus on the regime where $$N_s \gg 1$$ but $$s \ll 1$$, such that mutation '2' is substantially, but not overwhelmingly, beneficial [21]. We then have $$p_{20} \ll 1$$ (see Eq. (19)). To leading (i.e. zeroth) order in $$p_{20}$$, we obtain from Eq. (11) that

$$n_s = \frac{\Pi - 1}{p_{02}} \sum_{j=1}^{\Pi-1} \frac{1}{j}, \quad (24)$$

where $$p_{02}$$ is given by Eq. (18).
1. *Fitness plateau*

For a fitness plateau, combining Eqs. (13) and (22), the upper bound $U$ reads

$$U = \frac{N^2 p_{12}}{2(N - 1)} \sum_{j=2}^{\Pi} \frac{1}{j}.$$  \hspace{1cm} (25)

Additionally, Eqs. (24) and (21) can be combined to write the lower bound $L$ as

$$L = \frac{\Pi - 1}{N \rho_{02}} \sum_{j=1}^{\Pi - 1} \frac{1}{j}.$$  \hspace{1cm} (26)

again to lowest order in $p_{20}$. Here, we have used $p_{01} = 1/N$, since in the case of the plateau, mutation ‘1’ is neutral.

For $N \gg 1$ and $\Pi \gg 1$, Eqs. (25) and (26) give

$$\frac{U}{L} \sim \frac{p_{02}^2}{\Pi} N^2 \approx \frac{s^2}{\Pi},$$  \hspace{1cm} (27)

where we have used $p_{02} = p_{12}$, since mutation ‘1’ is neutral. The last expression follows from $p_{02} \approx s$, since we have assumed $N s \gg 1$. This simple expression of the ratio $U/L$ demonstrates that the range over which subdivision significantly accelerates plateau crossing increases as the deme size $N$ becomes larger, and that this range is quite wide as long as the number of demes satisfies $\Pi \ll (N s)^2$, which is realistic (recall that we are in the regime $N s \gg 1$).

2. *Fitness valley*

Let us focus on valleys such that $\delta \ll 1$ but $N \delta \gg 1$. (Note that, in the opposite limit $N \delta \ll 1$, mutation ‘1’ is effectively neutral, and the above discussion on the fitness plateau applies.) Then, $p_{01} \sim \delta e^{-N \delta} \ll 1$ (see Eq. (14)). To lowest (i.e. zeroth) order in $p_{01}$, Eq. (17) becomes

$$n_d \sim \frac{\Pi}{2p_{10}},$$  \hspace{1cm} (28)

so that from Eq. (22), the upper bound $U$ is

$$U \sim \frac{p_{12}}{2p_{10}} \approx \frac{1}{2} \left(1 + \frac{s}{\delta}\right),$$  \hspace{1cm} (29)

where we used the conditions $\delta \ll 1$, $N \delta \gg 1$, $s \ll 1$ and $N s \gg 1$ to simplify the expressions of $p_{12}$ and $p_{10}$.

Meanwhile, from Eq. (21) and (24), the lower bound $L$ takes the form

$$L = \frac{p_{01}}{p_{02}} (\Pi - 1) \sum_{j=1}^{\Pi - 1} \frac{1}{j} \sim \frac{\delta e^{-N \delta}}{s + \delta} (\Pi - 1) \sum_{j=1}^{\Pi - 1} \frac{1}{j},$$  \hspace{1cm} (30)

where we used the conditions $\delta \ll 1$, $N \delta \gg 1$, $s \ll 1$ and $N s \gg 1$ to simplify the expressions of $p_{01}$ and $p_{02}$.

Eq. (29) shows that $U$ is independent of $N$ and $\Pi$ in this regime, while Eq. (30) shows that $L$ decreases exponentially with $N$ (this dependence on $N$ comes from that of $p_{01}$). Hence, for valleys as well, the interval of $m/(\mu d)$ where subdivision most accelerates crossing becomes wider as $N s \gg 1$ increases. Note that, conversely, $L$ decreases when $\Pi$ increases.

In Fig. 2 of the main text, we have plotted $U$ and $L$ versus $\delta$ (solid lines). It can be seen that $U$ decreases weakly when $\delta$ increases (upper solid line), while $L$ decreases faster with $\delta$ (lower solid line), so that the interval of $m/(\mu d)$ where subdivision most accelerates crossing becomes wider when $\delta$ increases. This behavior is consistent with the $\delta$ dependences of Eqs. (29-30). Note, however, that the conditions $\delta \ll 1$, $N \delta \gg 1$, $s \ll 1$ and $N s \gg 1$ are not sufficiently well satisfied in Fig. 2 for Eqs. (29-30) to be very precise approximations of $U$ and $L$, and therefore the exact expressions were used to plot the lines in Fig. 2.

II. CROSSING TIME OF THE BEST DEME

In this section, we give more details on the calculation of the average valley or plateau crossing time $\tau_b$ by the best deme amongst $\Pi$ independent ones. We show in the main text that the crossing time of the whole metapopulation can be determined by this time.

$\tau_b$ is the average shortest crossing time of $\Pi$ independent demes. This minimum crossing time, which we denote by $t_b$, is also called the smallest (or first) order statistic of the deme crossing time among a sample of size $\Pi$ [31].

Let us denote by $p(t)$ the probability density function of valley or plateau crossing time for a single deme, and let us introduce $P(t) = \int_{-t}^{\infty} p(t') dt'$ (it satisfies $P(t) = 1 - C(t)$ where $C(t)$ is the cumulative distribution function of valley or plateau crossing by a single deme). The probability that $t_b$ is larger than $t$ is equal to the probability that the crossing times of each of the $\Pi$ independent demes are all larger than $t$: $P(t_b > t) = [P(t)]^\Pi$. By differentiating this expression, one obtains the probability density function $p_b(t_b)$ of the crossing time $t_b$ by the best deme (see e.g. Ref. [31]):

$$p_b(t_b) = \Pi [P(t_b)]^{\Pi - 1} p(t_b).$$  \hspace{1cm} (31)

We now express $p(t)$ explicitly. Since demes are assumed to be in the sequential fixation regime, valley or plateau crossing involves two successive steps. The first step, fixation of a ‘1’-mutant, occurs with rate $r_{01}$, and the second step, fixation of a ‘2’-mutant, occurs with rate $r_{12}$ (see the main text for expressions of these rates). The total crossing time is thus a sum of two independent exponential random variables, with probability density function given by a two-parameter hypoexponential distribution [31]:

$$p(t) = \frac{r_{01} r_{12}}{r_{12} - r_{01}} (e^{-r_{01} t} - e^{-r_{12} t}).$$  \hspace{1cm} (32)
Combining Eqs. (31) and (32), we obtain
\[ p_b(t_b) = \pi \left( \frac{r_{12} e^{-r_{01} t_b} - r_{01} e^{-r_{12} t_b}}{r_{12} - r_{01}} \right)^{\pi - 1} p(t_b), \]
(33)
with \( p(t_b) \) given by Eq. (32). \( t_b \) can then be obtained for any value of the parameters by computing the average value of \( t_b \) over this distribution.

As explained in the main text, \( r_{01} \ll r_{12} \) over a broad range of parameter values. In this limit, we can approximate \( p(t) \) with a simple exponential distribution,
\[ p(t) = r_{01} e^{-r_{01} t}. \]
(34)
Eq. (31) then yields
\[ p_b(t_b) = \pi r_{01} e^{-\pi r_{01} t}, \]
(35)
i.e. \( t_b \) is distributed exponentially with rate \( \pi r_{01} \). In this case, we simply have \( \tau_b = 1/(\pi r_{01}) \), which can be written as \( \tau_b \approx \tau_{id}/\pi \), where \( \tau_{id} \approx 1/r_{01} \) is the average crossing time for an isolated deme in the limit where \( r_{01} \ll r_{12} \).

Hence, in this limit, the best deme crosses the valley \( \Pi \) times faster on average than an isolated deme.

III. A LARGE POPULATION CONNECTED BY MIGRATION TO SMALLER POPULATION ISLANDS

Let us consider a population of \( N \) individuals connected by migration to \( S \) smaller population islands with \( N < S \) individuals each. These islands of identical size are assumed to be in the sequential fixation regime. For the sake of simplicity, we consider that migration only occurs between the large population and the islands: a migration step is a random exchange of two individuals between the large population and one of the islands (chosen at random at each migration event), and the total migration rate is denoted by \( M \). Here, we focus on the valley or plateau crossing time of the large population. We demonstrate that the evolution of a large population can be driven by that of satellite islands.

In the optimal case, the crossing time of the large population is determined by that of the best island, i.e., that which crosses the fitness valley or plateau fastest. We now determine the conditions under which this optimum is achieved, focusing on migration rates much smaller than division and death rates, \( M < \min(dNS, dN) \), such that fixation or extinction of a mutant lineage in either the large population or an island is not significantly perturbed by migration. Again, migration should be rare enough for islands to remain effectively shielded from migration events while they have fixed the intermediate mutation, until the final beneficial mutation arises. Second, migration should also be frequent enough for the spreading time of the final beneficial mutation from the best island to the large population to be negligible with respect to the crossing time of the best island. These two criteria again provide upper and lower bounds on \( M/\mu d \).

The average time \( \tau_{12} = 1/(N \mu dp_{12}) \) (with \( p_{12} \) from Eq. (23)) required for an island of ‘1’-mutants to fix the beneficial mutation ‘2’ must be smaller than the average time, \( t_d \), for an island of ‘1’-mutants to be wiped out by migration from the large population, which still exhibits genotype ‘0’. The rate of migration events between the island of ‘1’-mutants and the large population is \( M/S \). Hence, \( t_d = S/(M p_{10}) \), where \( p_{10} \) is the probability of fixation of the lineage of a single migrant with genotype ‘0’ in an island where all other individuals are ‘1’-mutants: for valleys, it is given by Eq. (15), while for plateaus, it is equal to \( 1/N \). The first condition, \( \tau_{12} < t_d \), thus yields
\[ \frac{M}{\mu d} < \frac{N S p_{12}}{p_{10}}. \]
(36)
The second condition is that the average spreading time, \( t_s \), for the final beneficial mutation to fix in a large population after it has fixed in the best island, must be smaller than the average valley/plateau crossing time, \( \tau_0 \), of the best island. Similar to \( t_d \) previously, we obtain \( t_s = S/(M p_{02}) \), where \( p_{02} = (1 - e^{-s})/(1 - e^{-N s}) \) is the probability of fixation of a migrant with genotype ‘2’ in the large population (assumed to exhibit genotype ‘0’ before migration). \( \tau_0 \) is the average of the minimum crossing time among \( S \) independent islands. We again focus, for simplicity, on the limit where the first step of valley/plateau crossing, which occurs at rate \( r_{01} \), is much longer than the second. Then, we simply have \( \tau_b \approx \tau_{ii}/S \) (see Section II). In this expression, \( \tau_{ii} \approx r_{01}^{-1} = 1/(N \mu dp_{01}) \) (with \( p_{01} \) in Eq. (14)) is the average crossing time for an isolated island. Hence, the best island crosses the valley \( S \) times faster on average than a single isolated island. The second condition, \( t_s < \tau_0 \), finally yields
\[ \frac{S^2 N p_{01}}{p_{02}} < \frac{M}{\mu d}. \]
(37)
Together, Eqs. (36) and (37) yield the interval of \( M/\mu d \) over which we expect subdivision to be most beneficial:
\[ \frac{S^2 N p_{01}}{p_{02}} < \frac{M}{\mu d} < \frac{NS p_{12}}{p_{10}}. \]
(38)
In this range, we expect the valley/plateau crossing time \( \tau_i \) of the large population to be dominated by the crossing time of the best island, so that \( \tau_i/\tau_{ii} \approx 1/S \). This prediction is confirmed by our simulations (see Fig. 3(b) of the main text).

IV. SIMULATION METHODS

Our simulations are based on a Gillespie algorithm [36, 37] that we coded in the C language. Here we will describe our algorithm for the case of a metapopulation of
Π demes of identical size, which is the primary situation discussed in our work. In our simulations, each deme has a fixed carrying capacity $K$—we discuss this choice in Sec. B.

A. Algorithm

A number of different events occur in our simulations, each with an independent rate:

- Each individual divides at rate $f_g(1-N_i/K)$, where $f_g$ is the fitness associated with the genotype $g \in \{0,1,2\}$ of the individual, and $N_i$ is the current total number of individuals in the deme $i \in [1,\Pi]$ to which the individual belongs.

- If a dividing cell has $g < 2$, upon division, its offspring (i.e., one of the two individuals resulting from the division) mutates with probability $\mu$, to have genotype $g + 1$ instead of $g$.

- Each individual dies at rate $d$.

- Migration occurs at total rate $m \sum_{i=1}^{\Pi} N_i$. Two different demes are chosen at random, an individual is chosen at random from each of these two demes, and the two individuals are exchanged. There is no geographic structure in our model, i.e. exchange between any two demes is equally likely.

In practice, the number of individuals with each genotype in each deme is stored, as well as the corresponding division rate. This data fully describes the state of the metapopulation, and allows determination of the rates of all events. For each event in the simulation, the following steps are performed:

- A timestep $dt$ is drawn from an exponential distribution with rate equal to the total rate $R_t$ of events (i.e., the sum of all rates), and time is increased from its previous value, $t$, to $t + dt$. In other words, the next event occurs at time $t + dt$.

- The event that occurs at $t + dt$ is chosen randomly, such that the probability of an event with rate $r$ is equal to $r/R_t$: either a cell divides, or a cell dies, or a migration event occurs.

- The event is performed, and the relevant data is updated. Since we store the number of individuals with each genotype in each deme, only one or two of these numbers need to be updated at each step. In addition, the division rates of the affected deme $i$ must be updated upon division and death because $N_i$ is modified. Note, however, that this represents only three numbers at most (one for each genotype).

The advantage of the Gillespie algorithm is that it is exact, and does not involve any artificial discretization of time.

B. Working at fixed carrying capacity

In our simulations, demes have a fixed carrying capacity, and the number of individuals per deme fluctuates weakly around its equilibrium value. This approach, also used in, e.g., Ref. [17] has the advantage of realism. Alternatively, we could impose a constant number of individuals per deme.

(i) First, we could choose a dividing individual in the whole metapopulation with probability proportional to its fitness, and simultaneously suppress another individual, chosen at random in the same deme. However, in this case, individuals in demes of higher fitness would exhibit shorter lifespans, which is not realistic and may introduce a bias.

(ii) A second possibility would be to choose a dividing individual (according to fitness) in each of the demes, and to simultaneously suppress another individual, chosen at random, in each deme. However, in this case, unless migration events are far less frequent than these collective division-death events (i.e., these Π division-death events), the time interval between them becomes artificially discretized. This introduces biases unless the total migration rate $m\Pi N$ is much smaller than $Nd$, i.e. unless $m \ll d/\Pi$.

Consequently, while imposing a constant number of individuals is a good simulation approach for a non-subdivided population (see, e.g., Ref. [21]), it tends to introduce biases in the study of metapopulations. While we chose to perform simulations with fixed carrying capacities in order to avoid any of these biases, we checked that, for small enough migration rates, our results are completely consistent with simulation scheme (ii) described above. This consistency check also demonstrates that it is legitimate to compare our simulation results obtained with fixed carrying capacities to our analytical work carried out with constant population size per deme.