# The Molecular Clock of Neutral Evolution Can Be Accelerated or Slowed by Asymmetric Spatial Structure 

## Citation

Allen, Benjamin, Christine Sample, Yulia Dementieva, Ruben C. Medeiros, Christopher Paoletti, and Martin A. Nowak. 2015. "The Molecular Clock of Neutral Evolution Can Be Accelerated or Slowed by Asymmetric Spatial Structure." PLoS Computational Biology 11 (2): e1004108. doi:10.1371/journal.pcbi.1004108. http://dx.doi.org/10.1371/journal.pcbi. 1004108.

## Published Version

doi:10.1371/journal.pcbi. 1004108

## Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:14351182

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http:// nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use\#LAA

## Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story.

Accessibility

Citation: Allen B, Sample C, Dementieva Y, Medeiros RC, Paoletti C, Nowak MA (2015) The Molecular Clock of Neutral Evolution Can Be Accelerated or Slowed by Asymmetric Spatial Structure. PLoS Comput Biol 11(2): e1004108. doi:10.1371/journal.pcbi. 1004108

Editor: Carl T. Bergstrom, University of Washington, UNITED STATES

Received: September 11, 2014
Accepted: January 2, 2015
Published: February 26, 2015
Copyright: © 2015 Allen et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: BA and MAN are supported by the Foundational Questions in Evolutionary Biology initiative of the John Templeton Foundation, RFP-1202. http://www.templeton.org/node/1521. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

# The Molecular Clock of Neutral Evolution Can Be Accelerated or Slowed by Asymmetric Spatial Structure 

Benjamin Allen ${ }^{1,2,3 *}$, Christine Sample ${ }^{1}$, Yulia Dementieva ${ }^{1}$, Ruben C. Medeiros ${ }^{1}$, Christopher Paoletti ${ }^{1}$, Martin A. Nowak ${ }^{2,4}$<br>1 Department of Mathematics, Emmanuel College, Boston, Massachusetts, United States of America, 2 Program for Evolutionary Dynamics, Harvard University, Cambridge, Massachusetts, United States of America, 3 Center for Mathematical Sciences and Applications, Harvard University, Cambridge, Massachusetts, United States of America, 4 Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, United States of America<br>* benjcallen@gmail.com


#### Abstract

Over time, a population acquires neutral genetic substitutions as a consequence of random drift. A famous result in population genetics asserts that the rate, $K$, at which these substitutions accumulate in the population coincides with the mutation rate, $u$, at which they arise in individuals: $K=u$. This identity enables genetic sequence data to be used as a "molecular clock" to estimate the timing of evolutionary events. While the molecular clock is known to be perturbed by selection, it is thought that $K=u$ holds very generally for neutral evolution. Here we show that asymmetric spatial population structure can alter the molecular clock rate for neutral mutations, leading to either $K<u$ or $K>u$. Our results apply to a general class of haploid, asexually reproducing, spatially structured populations. Deviations from $K=u$ occur because mutations arise unequally at different sites and have different probabilities of fixation depending on where they arise. If birth rates are uniform across sites, then $K \leq u$. In general, $K$ can take any value between 0 and $N u$. Our model can be applied to a variety of population structures. In one example, we investigate the accumulation of genetic mutations in the small intestine. In another application, we analyze over 900 Twitter networks to study the effect of network topology on the fixation of neutral innovations in social evolution.


## Author Summary

Evolution is driven by genetic mutations. While some mutations affect an organism's ability to survive and reproduce, most are neutral and have no effect. Neutral mutations play an important role in the study of evolution because they generally accrue at a consistent rate over time. This result, first discovered 50 years ago, allows neutral mutations to be used as a "molecular clock" to estimate, for example, how long ago humans diverged from chimpanzees and bonobos. We used mathematical modeling to study how the rates of these molecular clocks are affected by the spatial arrangement of a population in its
habitat. We find that asymmetry in this spatial structure can either slow down or speed up the rate at which neutral mutations accrue. This effect could potentially skew our estimates of past events from genetic data. It also has implications for a number of other fields. For example, we show that the architecture of intestinal tissue can limit the rate of genetic substitutions leading to cancer. We also show that the structure of social networks affects the rate at which new ideas replace old ones. Surprisingly, we find that most Twitter networks slow down the rate of idea replacement.

## Introduction

A half-century ago, Zuckerkandl and Pauling [1] discovered that amino acid substitutions often occur with sufficient regularity as to constitute a "molecular clock". Theoretical support for this observation was provided by Kimura [2], who argued that observed rates of amino acid substitution could only be explained if the majority of substitutions are selectively neutral. Under simple models of evolution, a single neutral mutation has probability $1 / N$ of becoming fixed in a haploid population of size $N$. It follows that the rate $K$ of neutral substitution per gen-eration-given by the product of the population size $N$, the mutation probability $u$ per reproduction, and the fixation probability $\rho$-is simply equal to $u$. (A similar cancellation occurs in diploids, leading again to $K=u$.) In other words, for any neutral genetic marker, the rate of substitution at the population level equals the rate of mutation at the individual level [2].

A number of factors can alter the rate of neutral substitution [3, 4], including selection, changes in population demography over time, or mutation rates that vary systematically by demographic classes [5, 6] or sex [7]. The extent to which these factors compromise the applicability of the molecular clock hypothesis to biological sequence data has been intensely debated [3, 4, 8-14].

However, it is generally thought that spatial structure alone (without spatial variation in the mutation rate [7]) cannot alter the rate of neutral substitution. This consensus is based on analyses [15-31] of a variety of models of spatially structured populations. Each of these analyses found that the fixation probability of a neutral mutation is $1 / N$, and thus the rate of neutral substitution $K=N u \rho$ again simplifies to $u$.

Here we show that the absence of spatial effects on the rate of neutral substitution in these models does not represent a general principle of evolution. Rather, it is an artifact of two common modeling assumptions: (i) all spatial locations are equivalent under symmetry, and (ii) mutations are equally likely to arise in each location. While assumption (i) is relaxed in a number of models, assumption (ii) is made almost universally. Either of these assumptions alone is sufficient to guarantee $\rho=1 / N$ and $K=u$, as we will show. However, neither of these assumptions is necessarily satisfied in a biological population. In particular, assumption (ii) is violated if some spatial locations experience more turnover (births and deaths) than others. Since each birth provides an independent opportunity for mutation, the rate at which new mutations appear at a location is proportional to its turnover rate [32]. Thus, even with a constant probability of mutation per birth, mutations may appear with different frequency at different locations. If, in addition, fixation probability depends on a mutant's initial location, the rate of neutral substitution is altered.

Our goal is to identify conditions under which the molecular clock rate is maintained ( $K=$ $u)$, accelerated $(K>u)$, or slowed $(K<u)$ by spatial population structure. Our main results are as follows (see also Fig. 1):


Fig 1. How spatial structure affects the molecular clock rate $K$. Relative to the rate in a well-mixed population $(K=u)$, spatial structure can either accelerate $(K>u)$ or slow $(K<u)$ the accumulation of neutral substitutions, depending on how birth rates $b_{i}$ and death rates $d_{i}$ vary across sites. The rate is unchanged from that of a well-mixed population $(K=u)$ if either death rates are uniform across sites (Result 1), or the birth and death rates are equal at each site (Result 2). Almost all previous studies of neutral drift in spatially structured populations fall into one of these two categories; thus the effects of spatial structure on the molecular clock rate are unappreciated. We show that, in general, $K$ can take any non-negative value less than $N u$ (Result 4). If one adds the constraint that the birth rate is the same at each site, then the molecular clock rate cannot exceed that of a well-mixed population ( $K \leq u$; Result 3).
doi:10.1371/journal.pcbi.1004108.g001

Result 1 If the death rate is constant over sites, the molecular clock rate is identical to that of a well-mixed population: $K=u$.

Result 2 If the birth rate equals the death rate at each site, the molecular clock rate is identical to that of a well-mixed population: $K=u$.

Result 3 If the birth rate is constant over sites, the molecular clock rate is less than or equal to that of a well-mixed population: $K \leq u$.

Result 4 In general (with no constraints on birth or death rates), the molecular clock rate $K$ can take any non-negative value less than $N u$.

## Results

## Obtaining the neutral substitution rate

Our investigations apply to a class of evolutionary models (formally described in the Methods) in which reproduction is asexual and the population size and spatial structure are fixed. Specifically, there are a fixed number of sites, indexed $i=1, \ldots, N$. Each site is always occupied by a single individual. At each time-step, a replacement event occurs, meaning that the occupants of some sites are replaced by the offspring of others. Replacement events are chosen according to a fixed probability distribution-called the replacement rule-specific to the model in question. Since we consider only neutral mutations that have no phenotypic effect, the probabilities of replacement events do not depend on the current population state.

This class includes many established evolutionary models. One important subclass is spatial Moran processes [23, 33-35], in which exactly one reproduction occurs each time-step. This class also includes spatial Wright-Fisher processes, in which the entire population is replaced each time-step [36,37]. In general, any subset $R \subset\{1, \ldots, N\}$ of individuals may be replaced in a given replacement event. Parentage in a replacement event is recorded in an offspring-to-parent map $\alpha: R \rightarrow\{1, \ldots, N\}$ (see Methods, [32,38]), which ensures that each offspring has exactly one parent and allows us to trace lineages over time.

For a given model in this class, we let $e_{i j}$ denote the (marginal) probability that the occupant of site $j$ is replaced by the offspring of site $i$ in a single time-step. Thus the expected number of offspring of site $i$ over a single time-step is $b_{i}=\sum_{j=1}^{N} e_{i j}$. The probability that node $i$ dies (i.e., is replaced) in a time-step is $d_{i}=\sum_{j=1}^{N} e_{j i}$. The death rate $d_{i}$ can also be regarded as the rate of turnover at site $i$. The total expected number of offspring per time-step is denoted $B=\sum_{i=1}^{N} b_{i}=\sum_{i=1}^{N} d_{i}=\sum_{i, j} e_{i j}$. We define a generation to be $N / B$ time-steps, so that, on average, each site is replaced once per generation.

We use this framework to study the fate of a single neutral mutation, as it arises and either disappears or becomes fixed. The probability of fixation depends on the spatial structure and the initial mutant's location. We let $\rho_{i}$ denote the probability that a new mutation arising at site $i$ becomes fixed. ( $\rho_{i}$ can also be understood as the reproductive value of site $i$ [39].) We show in the Methods that the fixation probabilities $\rho_{i}$ are the unique solution to the system of equations

$$
\begin{gather*}
d_{i} \rho_{i}=\sum_{j=1}^{N} e_{i j} \rho_{j} \quad \text { for } i=1, \ldots, N  \tag{1}\\
\sum_{i=1}^{N} \rho_{i}=1 \tag{2}
\end{gather*}
$$

Equation (2) arises because $\rho_{i}$ equals the probability that the current occupant of site $i$ will become the eventual ancestor of the population, which is true for exactly one of the $N$ sites.

To determine the overall rate of substitution, we must take into account the likelihood of mutations arising at each site. The rate at which mutations arise at site $i$ is proportional to the turnover rate $d_{i}$, because each new offspring provides an independent chance of mutation. Specifically, if mutation occurs at rate $u \ll 1$ per reproduction, then new mutations arise at site $i$ at rate $N u d_{i} / B$ per generation [32]. Thus the fraction of mutations that arise at site $i$ is $d_{i} / B$.

The overall fixation probability $\rho$ of new mutations, taking into account all possible initial sites, is therefore

$$
\begin{equation*}
\rho=\frac{1}{B} \sum_{i=1}^{N} d_{i} \rho_{i} \tag{3}
\end{equation*}
$$

The molecular clock rate $K$ is obtained by multiplying the fixation probability $\rho$ by the total rate of mutation per generation:

$$
\begin{equation*}
K=N u \rho=\frac{N u}{B} \sum_{i=1}^{N} d_{i} \rho_{i} . \tag{4}
\end{equation*}
$$

The units of $K$ are substitutions per generation. Alternatively, the molecular clock can be expressed in units of substitutions per time-step, in which case the formula is
$K=B u \rho=u \sum_{i=1}^{N} d_{i} \rho_{i}$.


Fig 2. For spatial structures with symmetry, $\boldsymbol{K}=\boldsymbol{u}$. (a) For a well-mixed population, represented by a complete graph with uniform edge weights, a neutral mutation has a $1 / N$ chance of fixation, where $N$ is the population size. It follows that the rate $K$ of neutral substitution in the population equals the rate $u$ of neutral mutation in individuals. (b) The same result holds for spatial structures in which each site is a priori identical, such as the cycle with uniform edge weights.
doi:10.1371/journal.pcbi.1004108.g002

## Effects of spatial structure

How does spatial structure affect the rate of neutral substitution? In a well-mixed population, each individual's offspring is equally likely to replace each other individual, meaning that $e_{i j}$ is constant over all $i$ and $j$ (Fig. 2a). In this case, the unique solution to Eqs. (1)-(2) is $\rho_{i}=1 / \mathrm{N}$ for all $i$, and we recover Kimura's [2] result $K=N u(1 / N)=u$. Moreover, if each site is equivalent under symmetry, as in Fig. 2b, this symmetry implies that $\rho_{i}=1 / N$ for all $i$ and $K=u$ as in the well-mixed case.

However, asymmetric spatial structure can lead to faster $(K>u)$ or slower $(K<u)$ molecular clock rates than a well-mixed population, as shown in Fig. 3. From Eqs. (2) and (4) we can see that $K>u$ is equivalent to the condition $\overline{d \rho}>\bar{d} \bar{\rho}$, where the bars indicate averages over $i$ $=1, \ldots, N$. This means that the molecular clock is accelerated if and only if $d_{i}$ and $\rho_{i}$ are positively correlated over sites; that is, if and only if there is a positive spatial correlation between the arrival of new mutations and the success they enjoy.

These results led us to seek general conditions on the spatial structure leading to faster, slower, or the same molecular clock rates as a well-mixed population. We first find

Result 1. If the death rates $d_{i}$ are constant over all sites $i=1, \ldots, N$, then $\rho=1 / N$, and consequently $K=u$.

Thus the molecular clock rate is unaffected by spatial structure if each site is replaced at the same rate (Fig. 4a). This result can be seen by noting that if the $d_{i}$ are constant over $i$, then since $\sum_{i=1}^{N} d_{i}=B$, it follows that $d_{i}=B / N$ for each $i$. Substituting in Eq. (3) yields $\rho=1 / N$.

Another condition leading to $K=u$ is the following:
Result 2. If the birth rate equals the death rate at each site $\left(b_{i}=d_{i}\right.$ for all $\left.i=1, \ldots, N\right)$, then $\rho=$ $1 / N$, and consequently $K=u$. Moreover, $b_{i}=d_{i}$ for all $i=1, \ldots, N$ if and only if the fixation probability is the same from each site ( $\rho_{i}=1 / N$ for all $i=1, \ldots, N$ ).


A


B

Fig 3. Asymmetric spatial structure affects the rate of neutral substitution. This is because the frequency of mutations and the probability of fixation differ across sites. Turnover rates are indicated by coloration, with red corresponding to frequent turnover and consequently frequent mutation. (a) A star consists of a hub and $n$ leaves, so that the population size is $N=n+1$. Edge weights are chosen so that the birth rates are uniform ( $b_{i}=1$ for all $i$ ). Solving Eqs. (1)-(2), we obtain site-specific fixation probabilities of $\rho_{\mathrm{H}}=$ $1 /\left(1+n^{2}\right)$ and $\rho_{\mathrm{L}}=n /\left(1+n^{2}\right)$ for the hub and each leaf, respectively. From Eq. (4), the molecular clock rate is $K=\frac{2 n}{1+n^{2}} u$, which equals $u$ for $n=1$ and is less than $u$ for $n \geq 2$. Thus the star structure slows down the rate of neutral substitution, in accordance with Result 3. Intuitively, the slowdown occurs because mutations are more likely to arise at the hub, where their chances of fixation are reduced. (b) A one-dimensional population with self-replacement only in site 1 . Solving Eqs. (1)-(2) we find $\rho_{1}=\frac{8}{15}, \rho_{2}=\frac{4}{15}, \rho_{3}=\frac{2}{15}$ and $\rho_{4}=\frac{1}{15}$. (The powers of two arise because there is twice as much gene flow in one direction as the other.) From Eq. (4), the molecular clock rate is $K=\frac{16}{13} u>u$, thus the molecular clock is accelerated in this case.
doi:10.1371/journal.pcbi.1004108.g003

Thus if births and deaths are balanced at each site, then all sites provide an equal chance for mutant fixation (Fig. 4b). In this case the molecular clock is again unchanged from the baseline value. In particular, if dispersal is symmetric in the sense that $e_{i j}=e_{j i}$ for all $i$ and $j$ then $K=u$. Result 2 can be obtained by substituting $\rho_{i}=1 / N$ for all $i$ into Eq. (1) and simplifying to obtain $b_{i}=d_{i}$ for all $i$ (details in Methods). Alternatively, Result 2 can be obtained as a corollary to the Circulation Theorem of Lieberman et al. [23].

Our third result reveals a "speed limit" to neutral evolution in the case of constant birth rates:

Result 3. If the birth rates $b_{i}$ are constant over all sites $i=1, \ldots, N$, then $\rho \leq 1 / N$, and consequently $K \leq u$, with equality if and only if the death rates $d_{i}$ are also constant over sites.
In other words, a combination of uniform birth rates and nonuniform death rates slows down the molecular clock. An instance of this slowdown in shown in Fig. 3a. Intuitively, the sites at which mutations occur most frequently are those with high death rates $d_{i}$; because of these high death rates, these sites on the whole provide a reduced chance of fixation. The proof of this result, however, is considerably more intricate than this intuition would suggest (see Methods).

Finally, we investigate the full range of possible values for $K$ with no constraints on birth and death rates. We find the following:

Result 4. For arbitrary spatial population structure (no constraints on $e_{i j}$ ) the fixation probability can take any value $0 \leq \rho<1$, and consequently, the molecular clock can take any rate $0 \leq$ $K<N u$.


Fig 4. Results 1 and 2 give conditions leading to $K=u$. (a) Our Result 1 states that the molecular clock has the same rate as in a well-mixed population, $K=u$, if the rate of turnover $d_{i}$ is uniform across sites, as in this example ( $d_{i}=0.2$ for all $\left.i\right)$. (b) Result 2 asserts that $\rho_{i}=1 / \mathrm{N}$ for all $i$-again implying $K=u$-if and only if each site has birth rate equal to death rate, $b_{i}=d_{i}$ for all $i$, as in this example. Nodes are colored according to their rates of turnover $d_{i}$.

## $K=N[1-(N-1) a] u$



Fig 5. Result 4 shows that $\boldsymbol{K}$ can achieve any value $\mathbf{0} \leq \boldsymbol{K}<\boldsymbol{N} \boldsymbol{u}$. This is proven by considering a population structure with unidirectional gene flow from a hub $(\mathrm{H})$ to $N-1$ leaves (L). Fixation is guaranteed for mutations arising in the hub ( $\rho_{\mathrm{H}}=1$ ) and impossible for those arising in leaves $\left(\rho_{\mathrm{L}}=0\right)$. The overall fixation probability is equal by Eq. (3) to the rate of turnover at the hub: $\rho=d_{\mathrm{H}}=1-(N-1) a$. The molecular clock rate is therefore $K=N u \rho=N[1-(N-1) a] u$. It follows that $K>u$ if and only if $a<1 / N$. Intuitively, the molecular clock is accelerated if the hub experiences more turnover (and hence more mutations) than the other sites. Any value of $\rho$ greater than or equal to 0 and less than 1 can be achieved through a corresponding positive choice of a less than or equal to $1 /(N-1)$. For $a=1 /(N-1)$ we have $K=0$, because mutations arise only at the leaves where there is no chance of fixation. At the opposite extreme, in the limit $a \rightarrow 0$, we have $K \rightarrow N u$.
doi:10.1371/journal.pcbi.1004108.g005

This result is especially surprising, in that it implies that the probability of fixation of a new mutation can come arbitrarily close to unity. Result 4 can be proven by considering the hypothetical spatial structure illustrated in Fig. 5. Any non-negative value of $\rho$ less than 1 can be obtained by an appropriate choice of parameters (details in Methods).

## Application to upstream-downstream populations

To illustrate the effects of asymmetric dispersal on the molecular clock, we consider a hypothetical population with two subpopulations, labeled "upstream" and "downstream" (Fig. 6). The sizes of these subpopulations are labeled $N_{\uparrow}$ and $N_{\downarrow}$, respectively. Each subpopulation is well-mixed, with replacement probabilities $e_{\uparrow}$ for each pair of upstream sites and $e_{\downarrow}$ for each pair of downstream sites. Dispersal between the subpopulations is represented by the replacement probabilities $e_{\rightarrow}$ from each upstream site to each downstream site, and $e_{\leftarrow}$ from each downstream site to each upstream site. We assume there is net gene flow downstream, so that $e_{\rightarrow}>e_{\leftarrow}$.

Solving Eqs. (1)-(2), we find that the fixation probabilities from each upstream site and each downstream site, respectively, are

$$
\begin{align*}
& \rho_{\uparrow}=\frac{e_{\rightarrow}}{N_{\uparrow} e_{\rightarrow}+N_{\downarrow} e_{\leftarrow}}, \\
& \rho_{\downarrow}=\frac{e_{\leftarrow}}{N_{\uparrow} e_{\rightarrow}+N_{\downarrow} e_{\leftarrow}} . \tag{5}
\end{align*}
$$

These fixation probabilities were previously discovered for a different model of a subdivided population [40]. Substituting these fixation probabilities into Eq. (4) yields the molecular clock rate:

$$
\begin{equation*}
K=\frac{N}{B} \frac{N_{\uparrow} d_{\uparrow} e_{\rightarrow}+N_{\downarrow} d_{\downarrow} e_{\leftarrow}}{N_{\uparrow} e_{\rightarrow}+N_{\downarrow} e_{\leftarrow}} u . \tag{6}
\end{equation*}
$$

$$
K=\frac{N}{B} \frac{N_{\uparrow} d_{\uparrow} e_{\rightarrow}+N_{\downarrow} d_{\downarrow} e_{\leftarrow}}{N_{\uparrow} e_{\rightarrow}+N_{\downarrow} e_{\leftarrow}} u
$$


upstream

downstream

Fig 6. A model of a population divided into upstream and downstream subpopulations. Each subpopulation is well-mixed. The replacement probability $e_{i j}$ equals $e_{\uparrow}$ if sites $i$ and $j$ are both upstream, $e_{\downarrow}$ if $i$ and $j$ are both downstream, $e_{\rightarrow}$ if $j$ is upstream and $j$ is downstream, and $e_{\leftarrow}$ if $i$ is downstream and $j$ is upstream. We suppose there is net gene flow downstream, so that $e_{\rightarrow}>e_{\leftarrow}$. We find that the molecular clock is accelerated, relative to the well-mixed case, if and only if the upstream subpopulation experiences more turnover than the downstream subpopulation: $K>u$ if and only if $d_{\uparrow}>d_{\downarrow}$.
doi:10.1371/journal.pcbi.1004108.g006

Above, $d_{\uparrow}$ and $d_{\downarrow}$ are the turnover rates in the upstream and downstream populations, respectively, and $B=N_{\uparrow} d_{\uparrow}+N_{\downarrow} d_{\downarrow}$ is the total birth rate per time-step. In Methods, we show that $K>$ $u$ if and only if $d_{\uparrow}>d_{\downarrow}$; that is, the molecular clock is accelerated if and only if there is more turnover in the upstream population than in the downstream population.

In the case of unidirectional gene flow, $e_{\leftarrow}=0$, the molecular clock rate is simply $K=\left(d_{\uparrow} / B\right)$ $N u$. The quantity $d_{\uparrow} / B$ represents the relative rate of turnover in the upstream population, and can take any value in the range $0 \leq d_{\uparrow} / B<1 / N_{\uparrow}$; thus $K$ takes values in the range $0 \leq K<(N /$ $\left.N_{\uparrow}\right) u$. We note that the upper bound on $K$ is inversely proportional to the size $N_{\uparrow}$ of the upstream population. The largest possible values of $K$ are achieved when $N_{\uparrow}=1$, in which case $K$ can come arbitrarily close to Nu . These bounds also hold if there are multiple downstream subpopulations, since for unidirectional gene flow, the spatial arrangement of downstream sites does not affect the molecular clock rate. In particular, if the hub and leaves in Fig. 5 are each replaced by well-mixed subpopulations, then $K$ is bounded above by $\left(N / N_{\mathrm{H}}\right) \mathcal{u}$, where $N_{\mathrm{H}}$ is the size of the hub subpopulation.

## Application to epithelial cell populations

Our results are also applicable to somatic evolution in self-renewing cell populations, such as the crypt-like structures of the intestine. Novel labeling techniques have revealed that neutral mutations accumulate in intestinal crypts at a constant rate over time [41]. The cell population in each crypt is maintained by a small number of stem cells that reside at the crypt bottom and continuously replace each other in stochastic manner (Fig. 7; [42-44]). We focus on the proximal small intestine in mice, for which recent studies $[41,45]$ suggest there are $\sim 5$ active stem cells per crypt, each replaced $\sim 0.1$ times per day by one of its two neighbors. In our framework, this corresponds to a cycle-structured population of size 5 with replacement rates 0.05/ day between neighbors, so that $d_{i}=0.1 /$ day for all $i$.

Only mutations that arise in stem cells can become fixed within a crypt; thus we need only consider the fixation probabilities and turnover rates among stem cells. By symmetry among the stem cells, $\rho_{i}=1 / 5$ for each of the five stem cell sites. The molecular clock rate is therefore $K=u \sum_{i=1}^{5} d_{i} \rho_{i}=0.1 u$ substitutions per day. This accords with the empirical finding that, for a neutral genetic marker with mutation rate $u \approx 1.1 \times 10^{-4}$, substitutions accumulate at a rate $\tilde{K} \approx 1.1 \times 10^{-5}$ per crypt per day [41].

Does crypt architecture limit the rate of genetic change in intestinal tissue? Intestinal crypts in mice contain $\sim 250$ cells and replace all their cells about once per day [46]. If each crypt were a well-mixed population, the molecular clock rate would be $\tilde{K}=B u / N \approx u$ substitutions per day. Thus the asymmetric structure of these epithelial crypts slows the rate of neutral genetic substitution tenfold.

## Application to the spread of ideas

Our results can also be applied to ideas that spread by imitation on social networks. In this setting, a mutation corresponds to a new idea that could potentially replace an established one. Neutrality means that all ideas are equally likely to be imitated.

To investigate whether human social networks accelerate or slow the rate of idea substitution, we analyzed 973 Twitter networks from the Stanford Large Network Dataset Collection [47]. Each of these "ego networks" represents follower relationships among those followed by a single "ego" individual (who is not herself included in the network). We oriented the links in each network to point from followee to follower, corresponding to the presumed direction of information flow. Self-loops were removed. To ensure that fixation is possible, we eliminated

COMPUTATIONAL
BIOLOGY


Fig 7. A simple model of cell replacement structure in epithelial crypts of the small intestine, based on results of [41] and [45]. A small number of stem cells $\left(N_{\uparrow} \sim 5\right)$ residing at the bottom of the intestinal crypt and are replaced at rate $d_{\uparrow} \sim 0.1$ per stem cell per day. Empirical results [41, 45] suggest a cycle structure for stem cells. To achieve the correct replacement rate we set $e_{i j}=0.05 /$ day for each neighboring pair. Stem cells in an individual crypt replace a much larger number of progenitor and differentiated cells ( $\sim 250$; [46]). These downstream progenitor and differentiated cells are replaced about every day [46]. The hierarchical organization of intestinal crypts, combined with the low turnover rate of stem cells, limits the rate of neutral genetic substitutions ( $\kappa \approx 0.1 u$ substitutions per day), since only mutations that arise in stem cells can fix.
doi:10.1371/journal.pcbi.1004108.g007
individuals that could not be reached, via outgoing links, from the node with greatest eigenvector centrality. The resulting networks varied in size from 3 to 241 nodes.

To model the spread of ideas on these networks, we set $e_{i j}=1 / L$ if $j$ follows $i$ and zero otherwise, where $L$ is the total number of links. This can be instantiated by supposing that at each time-step, one followee-follower link is chosen with uniform probability. The follower either adopts the idea of the followee, with probability $1-u$, or innovates upon it to create a new idea, with probability $u$, where $u \ll 1$. With these assumptions, the resulting rate of idea substitution (as a multiple of $u$ ) depends only on the network topology and not on any other parameters.

We found that the mean value of $K$ among these ego networks is $0.557 u$, with a standard deviation of $0.222 u$. 19 of the 973 networks (2\%) have $K>u$. Two networks have $K=u$ exactly; each of these has $N=3$ nodes and uniform in-degree $d_{i}$, thus $K=u$ follows from Result 1 for these networks. We found a weak but statistically significant negative relationship between the


Fig 8. The rate of neutral substitution $K$ on Twitter "ego networks". (a-e) Five of the 973 networks analyzed, including those with (a) the largest value of $K$, (b) the smallest value of $K$, and (c) the fewest nodes. (f) A scatter plot of $K / u$ versus $N$ reveals a weak negative correlation (slope $\approx-0.00164$ with $95 \%$ confidence interval ( $-0.0023,-0.001$ ) based on the bootstrap method; $R \approx-0.45$ ). The colored dots on the scatter plot correspond to the networks shown in (a-e). The dashed line corresponds to $K / u=1$, above which network topology accelerates neutral substitution.

[^0]

Fig 9. Illustration of idea generation and spreading on Twitter "ego networks". Network structure accelerates idea substitution ( $K>u$ ) if and only if there is a positive spatial correlation between the generation of new ideas (which for our model occurs proportionally to the rate $d_{i}$ of incoming ideas) and the probability of fixation $\rho_{i}$. Panels (a) and (b) show the networks with the slowest ( $K \approx 0.667 u$ ) and fastest ( $K \approx 1.085 u$ ) rates of idea substitution, respectively, among networks of size 13. The coloration of nodes corresponds to their rate of turnover $d_{i}$, with warmer colors indicating more rapid turnover. The size of nodes corresponds to their fixation probability $\rho_{i}$.
doi:10.1371/journal.pcbi.1004108.g009
network size $N$ and value $K / u$ (slope $\approx-0.00164$ with $95 \%$ confidence interval ( -0.0023 , -0.001 ) based on the bootstrap method; $R \approx-0.45$ ). This negative relationship persists even if small networks with less than 10 nodes are removed (slope $\approx-0.00156$ with $95 \%$ confidence interval ( $-0.0023,-0.0009$ ); $R \approx-0.43$ ). In summary, while some Twitter ego-networks accelerate the substitution of neutral innovations, the vast majority slow this rate (Figs. 8 and 9).

One possible explanation for the rarity of networks that accelerate idea substitution has to do with the intrinsic relationship between the turnover rates $d_{i}$ and the site-specific fixation probabilities $\rho_{i}$. From Eq. (1), we see that $\rho_{i}$ can be written as the product
$\left(1 / d_{i}\right) \times\left(\sum_{j=1}^{N} e_{i j} \rho_{j}\right)$, where the first factor can be interpreted as the "attention span" of node $i$ and the second can be interpreted as its influence. While these two factors are not strictly independent, we would not necessarily expect a systematic relationship between them in our Twitter network ensemble. In the absence of such a relationship, $\rho_{i}$ is inversely related to $d_{i}$, which implies $K<u$. In other words, the most fertile nodes (in terms of generating new ideas) are also the most fickle (in terms of adopting the ideas of others); thus many new ideas are abandoned as soon as they are generated. This heuristic argument suggests that $K>u$, while possible, might be an uncommon occurrence in networks drawn from statistical or probabilistic ensembles.

## Discussion

The spatial structure of a population affects its evolution in many ways, for example by promoting cooperative behaviors [36, 48-54], genetic variation [55-58], and speciation [59-61]. Asymmetric spatial structure in particular is known to have important consequences for adaptation [23, 27, 28, 35, 62-65] and for genetic diversity [66-68]. Our work shows that asymmetric spatial structure also affects the rate of neutral substitution.

In light of Results 1 and 2, we see that the critical factors driving the changes in molecular clock rate are differential probabilities of turnover $\left(d_{i}\right)$ and net offspring production $\left(b_{i}-d_{i}\right)$ across sites. If both $d_{i}$ and $b_{i}-d_{i}$ differ across sites, the molecular clock rate will in general differ
from that of a well-mixed population (Result 4). If additionally $b_{i}$ is constant across sites, Result 3 guarantees that neutral substitution is slowed relative to the well-mixed case.

Our Result 4 shows that the rate of neutral substitution in a population can come arbitrarily close to $N u$. However, this result depends on the existence of a single "hub" individual seeding the rest of the population as in Fig. 5. A similar but more plausible scenario (especially in sexually reproducing populations) involves a well-mixed hub subpopulation seeding one or more "leaf" subpopulations. In this case, the upper bound on the molecular clock rate is $\left(N / N_{\mathrm{H}}\right) u$, where $N_{\mathrm{H}}$ is the size of the hub subpopulation.

Conditions leading to altered molecular clock rate may occur frequently in natural populations. Asymmetric dispersal may result, for example, from prevailing winds [69, 70], water currents [71, 72], or differences in elevation [73, 74]. Differences in habitat quality may lead to variance in birth and death rates across sites (nonuniform $b_{i}$ and $d_{i}$ ). It is known that such asymmetries in spatial structure have important consequences for adaptation [23, 62-64] and for genetic diversity [66-68]; our work shows that they also have consequences for the rate of neutral genetic change. In particular, the molecular clock is accelerated if there is greater turnover in "upstream" subpopulations.

One important assumption made in our work is that mutation occurs with a constant probability per reproduction. Alternatively, one might suppose that heritable mutations accrue at a constant rate per individual (e.g. due to germline cell divisions). With this alternate assumption, mutations would arise at uniform rates over sites, resulting in a molecular clock rate of $K$ $=u$ for all spatial structures. The applicability of our results thus depends on the mutation process of the population in question, which may in many cases lie between these extremes. In humans, for example, recent evidence suggests that maternal mutations occur with constant probability per reproduction, whereas paternal mutations (which are more frequent) increase in probability with the father's age [75-77]. In general, we expect deviations from $K=u$ to scale with the extent to which mutations in a lineage depend on the number of generations rather than chronological time.

Many somatic cell populations have strongly asymmetric patterns of replacement, with a small number of stem cell pools feeding a much larger number of progenitor and differentiated cells. The rate at which mutations accumulate in these populations has significant implications for the onset of cancer [78-80] and the likelihood of successful cancer therapy [81-84]. It is well-known that this rate is proportional to the rate of stem cell division, since only mutations that arise in stem cells can persist [45, 85-88]. Our work shows this how principle arises from a general analysis of neutral evolution in spatially structured populations. It is important to note, however, that cancer can alter the structure of cell hierarchies; for example, by altering the number and replacement rate of stem cells [41] and/or allowing differentiated cells to revert to stem cells [89]. This restructuring may, in turn, alter the rate of genetic substitution, with further ramifications for cancer progression.

The influence of social network topology on the spread of ideas and behaviors is a question of both theoretical and practical interest [90-97]. The neutral substitution rate $K$ on social networks describes how innovations spread when they are equally likely to be imitated as an existing convention. Our finding that most Twitter ego networks have rates less than those of wellmixed populations contrasts with results from epidemiological models, which generally find that the heterogeneity of real-world social networks accelerates contagion [90, 98, 99]. We note, however, that this finding is sensitive to the assumption that individuals generate new ideas in proportion to the rate of incoming ideas (which itself is proportional to the number of individuals followed). Since the success of an idea varies according to the node at which it arises, the overall rate of substitution depends on the distribution of new ideas among nodes.

For example, if one were to instead assume that each individual generates new ideas at an equal rate, one would find that the network topology has no effect on the rate of substitution.

If spatial structure remains constant over time, then the neutral substitution rate $K$ is in all cases a constant multiple of the mutation rate $u$. In this case, absent other complicating factors such as selection, neutral mutations will accrue at a constant rate that can be inferred from genetic data. However, if the spatial structure changes over time-due, for instance, to changes in climate, tumorogenesis, or social network dynamics [100-103] -the rate of neutral substitutions may change over time as well.

In our framework, the molecular clock rate is assumed to depend only on the rate at which mutations arise and their probability of becoming fixed. This approach assumes that the time to fixation is typically shorter than the expected waiting time $1 /(N u \rho)$ for the next successful mutation. If this is not the case, then substitution rates are also affected by fixation times. These fixation times are themselves affected by spatial structure [22, 30, 58, 104], leading to further ramifications for the molecular clock rate [105].

The starting point of our analysis is that the convention, commonly assumed in evolutionary models, that mutations arise with equal frequency at each site, is not necessarily the most natural choice. If there is a constant probability of mutation per birth, then mutations instead arise in proportion to the rate of turnover at a site. Here we have applied this principle to study the rate of neutral substitution. However, this principle also holds for advantageous and deleterious mutations, as well as those whose effect varies with location. It also applies to frequencydependent selection $[65,106]$. Re-analyzing existing models using this new mutation convention may reveal further surprises about how spatial structure affects evolution.


Fig 10. Illustration of a replacement event. In this case, the occupant of site 3 produced one offspring, which displaced the occupant of site 2 . The occupant of site 4 produced two offspring, one remaining in site 4 and displacing the parent, and the other displacing the occupant of site 5 . Thus the set of replaced positions is $R=\{2,4,5\}$ and the offspring-to-parent map $\alpha$ is given by $\alpha(2)=3, \alpha(4)=4, \alpha(5)=4$. There is no mutation in the example illustrated here, so each offspring inherits the type of the parent. Thus the population transitions from state $\mathbf{s}=(M, R, R, M, R)$ to $\mathbf{s}^{\prime}=(M, R, R, M, M)$.
doi:10.1371/journal.pcbi.1004108.g010

## Methods

## Class of Models

In the class of models we consider, there are $N$ sites indexed $i=1, \ldots, N$, each always occupied by a single individual. At each time-step, a replacement event occurs, in which the occupants of some positions are replaced by the offspring of others. A replacement event is identified by a pair ( $R, \alpha$ ), where $R \subset\{1, \ldots, N\}$ is the set of sites whose occupants are replaced by new offspring, and $\alpha: R \rightarrow\{1, \ldots, N\}$ is a set mapping indicating the parent of each new offspring. (This notation was introduced in Ref. [32].) A sample replacement event is illustrated in Fig. 10.

A model of neutral evolution is specified by a probability distribution over the set of possible replacement events. We call this probability distribution the replacement rule of the model. The probability of a replacement event $(R, \alpha)$ in this distribution will be denoted $p(R, \alpha)$. Neutrality is represented by independence of the probabilities $p(R, \alpha)$ from the state of the evolutionary process.

The only assumption we place on the replacement rule is that it should be possible for at least one site $i$ to contain the eventual ancestor of the population:

Assumption 1. There is an $i \in\{1, \ldots, N\}$, a positive integer $n$ and a finite sequence
$\left\{\left(R_{k}, \alpha_{k}\right)\right\}_{k=1}^{n}$ of replacement events such that

- $p\left(R_{k}, \alpha_{k}\right)>0$ for all $k$, and
- For all individuals $j \in\{1, \ldots, N\}$,

$$
\begin{equation*}
\alpha_{k_{1}} \circ \alpha_{k_{2}} \circ \cdots \circ \alpha_{k_{m}}(j)=i, \tag{7}
\end{equation*}
$$

where $k_{1}, \ldots<k_{m}$ is the maximal subsequence of $1, \ldots, n$ such that the compositions in Eq. (7) are well-defined.

We observe that Eq. (7) traces the ancestors of $j$ backwards in time to $i$. This assumption is equivalent to saying that there is at least one site $i$ such that mutations arising at site $i$ have nonzero fixation probability. It is also equivalent to the statement that the weighted digraph with edge weights $e_{i j}$ is out-connected from at least one vertex. Assumption 1 precludes degenerate cases such as two completely separate subpopulations with no gene flow between them, in which case fixation would be impossible.

For a specific replacement event $(R, \alpha)$, the sites that are replaced by the offspring of $i$ is the given by the preimage $\alpha^{-1}(i) \subset R$, i.e., the set of indices that map to $i$ under $\alpha$. The number of offspring of site $i$ is equal to $\left|\alpha^{-1}(i)\right|$, the cardinality (size) of this preimage. Taking all possible replacement events into account, the birth rate (expected offspring number) of site $i$ is given by

$$
b_{i}=\mathrm{E}\left[\left|\alpha^{-1}(i)\right|\right]=\sum_{(R, \alpha)} p(R, \alpha)\left|\alpha^{-1}(i)\right| .
$$

The death rate (probability of replacement) of site $i$ is equal to

$$
d_{i}=\operatorname{Pr}[i \in R]=\sum_{\substack{R, \alpha) \\ i \in R}} p(R, \alpha) .
$$

The probability that the offspring of $i$ displaces the occupant of $j$ in a replacement event is

$$
e_{i j}=\operatorname{Pr}[j \in R \text { and } \alpha(j)=i]=\sum_{\substack{(R, \alpha) \\ j \in R \\ \alpha(j)=i}} p(R, \alpha) .
$$

We observe that $b_{i}=\sum_{j=1}^{N} e_{i j}$ and $d_{i}=\sum_{j=1}^{N} e_{j i}$.

## The evolutionary Markov chain

To study the fixation of new mutations, we consider evolution with two genetic types: mutant (M) and resident (R). The type occupying site $i$ in a given state of the evolutionary process is denoted $s_{i} \in\{\mathrm{M}, \mathrm{R}\}$. The overall state of the process can be recorded as a string $\mathbf{s}=\left(s_{1}, \ldots, s_{N}\right)$ of length $N$ with alphabet $\{\mathrm{M}, \mathrm{R}\}$.

We assume that there is no further mutation after an initial mutant appears; thus offspring faithfully inherit the type of the parent. It follows that if the current state is $\mathbf{s}=\left(s_{1}, \ldots, s_{N}\right)$ and the replacement event $(R, \alpha)$ occurs, then the new state $\mathrm{s}^{\prime}=\left(s_{1}^{\prime}, \ldots, s_{N}^{\prime}\right)$ is given by

$$
s_{i}^{\prime}= \begin{cases}s_{i} & i \notin R \\ s_{\alpha(i)} & i \in R\end{cases}
$$

The above assumptions describe a Markov chain on the set of strings of length $N$ with alphabet $\{M, R\}$. We call this the evolutionary Markov chain. It is straightforward to show that, from any initial state, this Markov chain will eventually converge upon one of two absorbing states: $(M, \ldots, M)$ or $(R, \ldots, R)$ [32]. In the former case, we say that the mutant type has gone to fixation; in the latter case we say that the mutant type has disappeared.

## The ancestral Markov chain

It is useful to consider a variation on the evolutionary Markov chain called the ancestral Mar$k o v$ chain, denoted $\mathcal{A}$. Instead of two types, the ancestral Markov chain has $N$ types, labeled 1, $\ldots, N$, which correspond to the $N$ members of a "founding generation" of the population. Evolution proceeds according to the given replacement rule, as described above. The states of the ancestral Markov chain are strings of length $N$ with alphabet $\{1, \ldots, N\}$.

The ancestral Markov chain has a canonical initial state $\mathbf{a}_{0}=(1, \ldots, N)$, in which the type of each individual corresponds to its location. This initial state identifies the locations of each founding $(t=0)$ member of the population. The ancestral Markov chain with initial state $\mathbf{a}_{0}$ in our notation, $\left(\mathcal{A}, \mathbf{a}_{0}\right)$-has the useful feature that at any time $t \geq 0$, the state $\mathbf{a}(t)=\left(a_{1}(t), \ldots\right.$, $\left.a_{N}(t)\right)$ indicates the site occupied by each individual's founding ancestor. In other words, if $a_{j}(t)$ $=i$, then the current occupant of site $j$ is descended from the founder that occupied site $i$.

To relate the evolutionary Markov chains $\mathcal{A}$ and $\mathcal{M}$, consider any set mapping $\gamma:\{1, \ldots, N\}$ $\rightarrow\{\mathrm{M}, \mathrm{R}\}$. We think of $\gamma(i)$ as giving the genetic type (M or R) of each member $i=1, \ldots, N$ of the founding generation. The mapping $\gamma$ induces a mapping $\tilde{\gamma}$ from states of $\mathcal{A}$ to states of $\mathcal{M}$, defined by $\tilde{\gamma}\left(a_{1}, \ldots, a_{N}\right)=\left(\gamma\left(a_{1}\right), \ldots, \gamma\left(a_{N}\right)\right)$. For any state $\mathbf{a}(t)$ of $\left(\mathcal{A}, \mathbf{a}_{0}\right)$, the string $\tilde{\gamma}(\mathbf{a}(t)) \in\{1, \ldots, N\}^{N}$ indicates the genetic type of each individual's ancestor in the founding generation. Since genetic types are inherited faithfully, it follows that $\tilde{\gamma}(\mathbf{a}(t))$ gives the current genetic type of each individual. Thus if $\mathcal{M}$ and $\mathcal{A}$ follow the same replacement rule, we have that for any such mapping $\gamma$ and any string $\mathbf{s} \in\{\mathrm{M}, \mathrm{R}\}^{N}$,

$$
\begin{equation*}
\mathbf{P}_{\left(\mathcal{A}, \mathbf{a}_{0}\right)}[\tilde{\gamma}(\mathbf{a}(t))=\mathbf{s}]=\mathrm{P}_{\left(\mathcal{M}, \tilde{\gamma}\left(\mathbf{a}_{0}\right)\right)}[\mathbf{s}(t)=\mathbf{s}] \tag{8}
\end{equation*}
$$

## Basic results on fixation probabilities

We define the fixation probability from site i , $\rho_{i}$, to be the probability that, from an initial state a mutant in site $i$ and residents in all other sites, the mutant type goes to fixation:

$$
\begin{equation*}
\rho_{i}=\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{M}, \mathbf{m}_{i}\right)}[\mathbf{s}(t)=(\mathrm{M}, \ldots, \mathrm{M})] \tag{9}
\end{equation*}
$$

Above, $\mathbf{m}_{i}$ denotes the initial state consisting of an M in position $i$ and R's elsewhere. The
$\operatorname{ordered} \operatorname{pair}\left(\mathcal{M}, \mathbf{m}_{i}\right)$ refers to the Markov chain $\mathcal{M}$ with initial state $\mathbf{m}_{i} . \mathrm{P}_{\left(\mathcal{M}, \mathbf{m}_{i}\right)}[\mathbf{s}(t)=\mathbf{s}]$ denotes the probability that the state of $\left(\mathcal{M}, \mathbf{m}_{i}\right)$ is $\boldsymbol{s}$ at time $t \geq 0$.

We can use the relationship between the ancestral Markov chain $\mathcal{A}$ and the evolutionary Markov chain $\mathcal{M}$ to obtain an alternate expression for the site-specific fixation probability $\rho_{i}$ :

Theorem 1. $\rho_{i}=\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{A}, \mathbf{a}_{0}\right)}[\mathbf{a}(t)=(i, \ldots, i)]$.
In words, the site-specific fixation probability $\rho_{i}$ equals the probability that founding individual $i$ becomes the eventual ancestor of the whole population.

Proof. For any $i=1, \ldots, N$, define the set mapping $\gamma_{i}:\{1, \ldots, N\} \rightarrow\{\mathrm{M}, \mathrm{R}\}$ by

$$
\gamma_{i}(j)= \begin{cases}\mathrm{M} & \text { if } j=i \\ \mathrm{R} & \text { otherwise }\end{cases}
$$

(Intuitively, this mapping describes the case that individual $i$ of the founding generation is a mutant, and all others in the founding generation are residents.) Note that $\tilde{\gamma}_{i}\left(\mathbf{a}_{0}\right)=\mathbf{m}_{i}$. Combining Eq. (8) with the definition of $\rho_{i}$, we obtain

$$
\begin{aligned}
\rho_{i} & =\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{M}, \mathbf{m}_{i}\right)}[\mathbf{s}(t)=(\mathrm{M}, \ldots, \mathrm{M})] \\
& =\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{A}, \mathbf{a}_{0}\right)}\left[\tilde{\gamma}_{i}(\mathbf{a}(t))=(\mathrm{M}, \ldots, \mathrm{M})\right] \\
& =\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{A}, \mathbf{a}_{0}\right)}[\mathbf{a}(t)=(i, \ldots, i)] .
\end{aligned}
$$

More generally, we can consider the probability of fixation from an arbitrary set of sites. For any set $S \subset\{1, \ldots, N\}$, we let $\rho_{S}$ denote the probability that the mutant type becomes fixed, given the initial state $\mathbf{m}_{S}$ with mutants occupying the sites specified by $S$ and residents occupying all other sites:

$$
\rho_{S}=\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{M}, \mathrm{m}_{S}\right)}[\mathbf{s}(t)=(\mathrm{M}, \ldots, \mathrm{M})]
$$

Site-specific fixation probabilities are additive in the following sense:
Theorem 2. For any set $S \subset\{1, \ldots N\}$ of sites,

$$
\begin{equation*}
\rho_{S}=\sum_{i \in S} \rho_{i} \tag{10}
\end{equation*}
$$

In particular,

$$
\begin{equation*}
\sum_{i=1}^{N} \rho_{i}=1 \tag{11}
\end{equation*}
$$

This result has previously been obtained for a number of specific evolutionary processes on graphs [39, 107, 108]. Intuitively, the probability of fixation from the initial state described by $S$ equals the probability that one of the individuals in a site $i \in S$ becomes the eventual ancestor of the population. Since this cannot be true of more than one site in $S$, the overall probability $\rho_{S}$ is obtained by summing over all $i \in S$ the individual probabilities $\rho_{i}$ that site $i$ contains the eventual ancestor.

Proof. Suppose that we are given a set $S \subset\{1, \ldots, N\}$ of sites initially occupied by mutants. This situation is described by the set mapping $\gamma_{s}:\{1, \ldots, N\} \rightarrow\{\mathrm{M}, \mathrm{R}\}$, given by

$$
\gamma_{S}(i)= \begin{cases}\mathrm{M} & \text { if } i \in S \\ \mathrm{R} & \text { otherwise } .\end{cases}
$$

Invoking the relationship between $\mathcal{A}$ and $\mathcal{M}$-in particular, Eq. (8) and Theorem 1-we have

$$
\begin{aligned}
\rho_{S} & =\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{M}, \tilde{r}_{S}\left(\mathbf{a}_{0}\right)\right)}[\mathbf{s}(t)=(\mathrm{M}, \ldots, \mathrm{M})] \\
& =\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{A}, \mathbf{a}_{0}\right)}\left[\tilde{\gamma}_{S}(\mathbf{a}(t))=(\mathrm{M}, \ldots, \mathrm{M})\right] \\
& =\lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{A}, \mathbf{a}_{0}\right)}[\mathbf{a}(t)=(i, \ldots, i) \text { for some } i \in S] \\
& =\sum_{i \in S} \lim _{t \rightarrow \infty} \mathrm{P}_{\left(\mathcal{A}, \mathbf{a}_{0}\right)}[\mathbf{a}(t)=(i, \ldots, i)] \\
& =\sum_{i \in S} \rho_{i} .
\end{aligned}
$$

This proves Eq. (10). Eq. (2) now follows from letting $S=\{1, \ldots, N\}$, and noting that, in this case,

$$
\rho_{S}=\lim _{t \rightarrow \infty} \mathrm{P}_{(\mathcal{M},(\mathrm{M}, \ldots, \mathrm{M}))}[\mathbf{s}(t)=(\mathrm{M}, \ldots, \mathrm{M})]=1
$$

We now derive Eq. (1), which allows the fixation probabilities $\rho_{i}$ to be calculated from the replacement rates $e_{i j}$.

Theorem 3. For each $i=1, \ldots, N$,

$$
d_{i} \rho_{i}=\sum_{j=1}^{N} e_{i j} \rho_{j}
$$

Proof. Considering the change that can occur over a single time-step, we have the following recurrence relation:

$$
\rho_{i}=\left(1-d_{i}\right) \rho_{i}+\sum_{(R, \alpha)} p(R, \alpha) \rho_{\alpha^{-1}(i)} .
$$

The first term above represents the case that the occupant of $i$ survives the current time-step and becomes the eventual ancestor of the population, while the second term represents the case that one of $i$ 's offspring from the current time-step is the eventual ancestor. Subtracting ( 1 $\left.-d_{i}\right) \rho_{i}$ from both sides and applying Theorem 2 with $S=\rho_{\alpha^{-1}(i)}$ yields

$$
d_{i} \rho_{i}=\sum_{(R, \alpha)} p(R, \alpha)\left(\sum_{j \in \alpha^{-1}(i)} \rho_{j}\right)
$$

Now interchanging the summations on the right-hand side yields

$$
d_{i} \rho_{i}=\sum_{j=1}^{N} \rho_{j}\left(\sum_{\substack{(R, \alpha) \\ j \in R \\ \alpha(j)=i}} p(R, \alpha)\right)
$$

By definition,

$$
\sum_{\substack{(R, \alpha) \\ j \in R \\ j \in R \\ \alpha(j)=i}} p(R, \alpha)=e_{i j} .
$$

This completes the proof.

## Proof of Result 2

Theorem 4 (Result 2). The fixation probabilities from each site are equal to $1 / N\left(\rho_{i}=1 / N\right.$ for all $i=1, \ldots, N)$ if and only if each site has birth rate equal to death rate $\left(b_{i}=d_{i}\right.$ for all $\left.i=1, \ldots, N\right)$.

Proof. Assume first that the fixation probabilities from each site are all equal to $1 / N$. Substituting $\rho_{i}=1 / N$ for all $i$ into Eq. (1) yields $d_{i}=b_{i}$. This proves the "only if" direction.

Next assume that the birth rate is equal to the death rate at each site $\left(b_{i}=d_{i}\right.$ for all $i=1, \ldots$, $N)$. Eq. (1) can then be rewritten as

$$
\sum_{j=1}^{N} e_{i j} \rho_{i}=\sum_{j=1}^{N} e_{i j} \rho_{j} .
$$

Clearly, $\rho_{i}=1 / N$ for all $i$ satisfies the above equation for all $i$, and also satisfies $\sum_{i=1}^{N} \rho_{i}=1$. Assumption 1 guarantees that the solution to these equations is unique. Therefore $b_{i}=d_{i}$ for all $i$ implies $\rho_{i}=1 / N$ for all $i$, proving the "if" direction.

## Proof of Result 3

Theorem 5 (Result 3). If birth rates are constant over sites, $b_{i}=1 / N$ for all $i=1, \ldots, N$, then $\rho \leq$ $1 / N$ (and consequently $K \leq u$ ) with equality if and only if the death rates are also constant over sites.

Proof. We separate our proof into six steps. First, we show that there is no loss of generality in assuming that $B=1$. Second, we use the method of Lagrange multipliers to find the critical points of the function $\rho=\sum_{i=1}^{N} d_{i} \rho_{i}$ with respect to the variables $\left\{e_{i j}\right\}_{i, j=1}^{N}$ and $\left\{\rho_{i}\right\}_{i=1}^{N}$ and the constraints

$$
\begin{cases}\sum_{j=1}^{N} e_{i j}=\frac{1}{N} & i \in\{1, \ldots, N\}  \tag{12}\\ \left(\sum_{j=1}^{N} e_{j i}\right) \rho_{i}=\sum_{j=1}^{N} e_{i j} \rho_{j} & i \in\{1, \ldots, N\} \\ \sum_{i=1}^{N} \rho_{i}=1\end{cases}
$$

We obtain that the critical points are precisely those for which $d_{i}=1 / N$ for all $i$. In the next three steps, we use partial derivatives of Eqs. (1), (2) and (3) to form a second-order Taylor expansion of $\rho$ in the variables $\left\{e_{i j}\right\}_{i \neq j}$ around the critical points. Finally, we show that the critical points are maxima, completing the proof.

## Step 1: Normalize the expected number of offspring per time-step

We first show that there is no loss of generality in assuming that the total expected number of offspring per time-step, $B \equiv \sum_{i, j} e_{i j}$, is one. We will demonstrate this by showing that, if a
maximum of $\rho$ is achieved by a combination of replacement rates $\left\{e_{i j}^{*}\right\}_{i, j}$, this maximum is also achieved using the normalized rates $\left\{e_{i j^{\prime}}\right\}_{i, j}$ with $e_{i j}^{\prime}=e_{i j}^{*} / B^{*}, B^{*} \equiv \sum_{i, j} e_{i j}^{*}$.

First, we see that the equations

$$
d_{i}^{*} \rho_{i}=\sum_{j=1}^{N} e_{i j}^{*} \rho_{j} \quad \text { and } \quad d_{i}^{\prime} \rho_{i}=\sum_{j=1}^{N} e_{i j}^{\prime} \rho_{j}
$$

are equivalent, differing only by a factor of $B$. (Above, we have introduced the notation $d_{i}^{*}=$ $\sum_{j=1}^{N} e_{i j}^{*}$ and $d_{i}^{\prime}=\sum_{j=1}^{N} e_{i j}^{\prime}$. ) Thus the node-specific fixation probabilities $\left\{\rho_{i}\right\}_{i=1}^{N}$ resulting from $\left\{e_{i j^{\prime}}\right\}_{i, j}$ are the same as those resulting from $\left\{e_{i j}^{*}\right\}_{i, j}$.

Turning now to the overall fixation probabilities $\rho^{*}$ and $\rho^{\prime}$ corresponding to $\left\{e_{i j}^{*}\right\}_{i, j}$ and $\left\{e_{i j^{\prime}}\right\}_{i, j}$ respectively, we see that

$$
\rho^{*}=\frac{1}{B^{*}} \sum_{i=1}^{N} d_{i}^{*} \rho_{i}=\sum_{i=1}^{N} \frac{d_{i}^{*}}{B^{*}} \rho_{i}=\sum_{i=1}^{N} d_{i}^{\prime} \rho_{i}=\rho^{\prime} .
$$

Thus the same maximum is achieved by the normalized replacement rates $\left\{e_{i j^{\prime}}\right\}_{i, j}$, which satisfy $\sum_{i, j} e_{i j}^{\prime}=1$. We conclude that there is no loss of generality in assuming $B=1$. Together with uniform birth rates, this implies the constraint

$$
\begin{equation*}
\sum_{j=1}^{N} e_{i j}=\frac{1}{N} . \tag{13}
\end{equation*}
$$

## Step 2: Determine critical points

We use the method of Lagrange multipliers to find critical points of the function $\rho=\sum_{i=1}^{N} d_{i} \rho_{i}$ with respect to the variables $\left\{e_{i j}\right\}_{i, j=1}^{N}$ and $\left\{\rho_{i}\right\}_{i=1}^{N}$. Eqs. (13), (1), and (2) give the respective constraint equations

$$
\begin{aligned}
g_{i} & \equiv \sum_{j=1}^{N} e_{i j}-\frac{1}{N}=0, \text { for } i \in\{1, \ldots, N\} \\
h_{i} & \equiv \rho_{i}\left(\sum_{j=1}^{N} e_{j i}\right)-\sum_{i=1}^{N} e_{i j} \rho_{j}=0, \text { for } i \in\{1, \ldots, N\} \\
q & \equiv \sum_{i=1}^{N} \rho_{i}-1=0
\end{aligned}
$$

Critical points with respect to the above constraints are solutions to

$$
\nabla \rho=\sum_{i=1}^{N} \lambda_{i} \nabla g_{i}+\sum_{i=1}^{N} \mu_{i} \nabla h_{i}+\sigma \nabla q,
$$

where the $\lambda_{i}{ }^{\prime} s, \mu_{i}$ 's, and $\sigma$ are the Lagrange multipliers. The individual components of the gradient are obtained by taking the partial derivative with respect to $e_{k l}$, yielding

$$
\begin{equation*}
\rho_{l}=\lambda_{k}+\rho_{l} \mu_{l}-\rho_{l} \mu_{k}, \tag{14}
\end{equation*}
$$

for $k, l \in\{1, \ldots, N\}$, and the partial derivative with respect to $\rho_{k}$, yielding

$$
\begin{equation*}
\sum_{j=1}^{N} e_{j k}=\mu_{k} \sum_{j=1}^{N} e_{j k}-\sum_{j=1}^{N} \mu_{j} e_{j k}+\sigma, \tag{15}
\end{equation*}
$$

for $k=1, \ldots, N$. Solving Eq. (14) for $\rho_{l}$ gives

$$
\begin{equation*}
\rho_{l}=\frac{\lambda_{k}}{1-\mu_{l}+\mu_{k}} . \tag{16}
\end{equation*}
$$

Note that in the case $l=k$, Eq. (16) becomes $\rho_{k}=\lambda_{k}$. Therefore, we replace $\lambda_{k}$ with $\rho_{k}$ in Eq. (16):

$$
\begin{equation*}
\rho_{l}=\frac{\rho_{k}}{1-\mu_{l}+\mu_{k}} . \tag{17}
\end{equation*}
$$

Interchanging $k$ and $l$, we obtain

$$
\begin{equation*}
\rho_{k}=\frac{\rho_{l}}{1-\mu_{k}+\mu_{l}} . \tag{18}
\end{equation*}
$$

Combining Eqs. (17) and (18) yields $\mu_{l}=\mu_{k}$ and $\rho_{l}=\rho_{k}$ for all $k, l \in\{1, \ldots, N\}$. That is, all node-specific probabilities $\rho_{i}$ are equal. It follows from Eq. (2) that $\rho_{i}=1 / N$ for all $i=1, \ldots, N$. Furthermore, given that $\mu_{l}=\mu_{k}$, Eq. (15) yields $d_{k}=\sigma$ for all $k=1, \ldots, N$. That is, all death rates $d_{i}$ are equal. Therefore $d_{i}=1 / N$ for $i=1, \ldots, N$ because we have assumed (without loss of generality) that $\sum_{i=1}^{N} d_{i}=1$.

In summary, we have shown that the overall fixation probability $\rho$ has a critical point whenever all node-specific fixation probabilities and death rates are constant $\left(d_{i}=\rho_{i}=1 / N\right.$ for all $i=$ $1, \ldots, N)$. Consequently, $\rho=1 / N$ at all critical points. It still remains to prove that this critical value of $\rho$ is a maximum.

## Step 3: Taylor series expansion

To prove that the overall fixation probability has an absolute maximum of $1 / N$, we pick a critical point $\left\{e_{i j}^{*}\right\}_{\substack{i, j \\ i \neq j}}$. Then, viewing $\rho$ as a function of the independent variables $\left\{e_{i j}\right\}_{\substack{i, j \\ i \neq j}}$, we form a second-order Taylor expansion of $\rho$ around the critical point. (Since we are operating under the constraint $\sum_{j=1}^{N} e_{i j}=1 / N$ for all $i$, this set of variables suffices to determine the value of $\rho$.) We will show that the second-order term of this Taylor expansion is always less than or equal to zero.

The second order multivariate Taylor series expansion for $\rho$ about the critical point $\left\{e_{i j}\right\}_{\substack{i, j \\ i \neq j}}$ can be written as

$$
\begin{equation*}
\rho=\frac{1}{N}+\sum_{\substack{k, l \\ k \neq l}} \frac{\partial \rho}{\partial e_{k l}} \Delta e_{k l}+\frac{1}{2} \sum_{\substack{k, l, m, n \\ m \neq n \\ k \neq l}} \frac{\partial^{2} \rho}{\partial e_{k l} \partial e_{m n}} \Delta e_{k l} \Delta e_{m n}+\mathcal{O}\left(|\Delta \vec{e}|^{3}\right), \tag{19}
\end{equation*}
$$

where all derivatives are taken at $\left\{e_{i j}^{*}\right\}_{i, j}^{i \neq j}$, and $\Delta e_{k l}=e_{k l}-e_{k l}^{*}$. More simply, we write this expansion as

$$
\rho=\frac{1}{N}+\rho^{(1)}+\rho^{(2)}+\mathcal{O}\left(|\Delta \vec{e}|^{3}\right)
$$

with $\rho^{(1)}$ and $\rho^{(2)}$ representing the first- and second-order terms, respectively, in Eq. (19). The first-order term $\rho^{(1)}$ is zero since $\left\{e_{i j}^{*}\right\}_{i, j}$ ifj , is a critical point. Our goal will be to show that the sec-ond-order term, $\rho^{(2)}$ is always negative or zero.

We can find an alternative expression for $\rho^{(2)}$ using the definition of $\rho$ in Eq. (3) as follows. We introduce the notation $\Delta d_{i}=d_{i}-d_{i}^{*}=d_{i}-1 / N$. With this notation, Eq. (3) becomes

$$
\begin{align*}
\rho & =\frac{1}{N} \sum_{i=1}^{N} \rho_{i}+\sum_{i=1}^{N} \Delta d_{i} \rho_{i} \\
& =\frac{1}{N}+\sum_{i=1}^{N} \Delta d_{i} \rho_{i} . \tag{20}
\end{align*}
$$

We substitute the first-order Taylor series expansion for $\rho_{i}$,

$$
\rho_{i}=\frac{1}{N}+\rho_{i}^{(1)}+\mathcal{O}\left(|\Delta \vec{e}|^{2}\right)
$$

into Eq. (20), noting that $\Delta d_{i}=\mathcal{O}(|\Delta \vec{e}|)$ and $\sum_{i=1}^{N} \Delta d_{i}=0$. This yields

$$
\begin{aligned}
\rho & =\frac{1}{N}+\sum_{i=1}^{N} \Delta d_{i}\left(\frac{1}{N}+\rho_{i}^{(1)}+\mathcal{O}\left(|\Delta \vec{e}|^{2}\right)\right) \\
& =\frac{1}{N}+\sum_{i=1}^{N} \Delta d_{i} \rho_{i}^{(1)}+\mathcal{O}\left(|\Delta \vec{e}|^{3}\right) .
\end{aligned}
$$

Thus, the second-order term of the overall fixation probability can be written as

$$
\begin{equation*}
\rho^{(2)}=\sum_{i=1}^{N} \Delta d_{i} \rho_{i}^{(1)} . \tag{21}
\end{equation*}
$$

## Step 4: Obtain recursion for first-order term of the site-specific fixation probability

We now investigate the properties of $\rho_{i}^{(1)}$. We begin by rewriting Eq. (1) as

$$
d_{i} \rho_{i}=e_{i i} \rho_{i}+\sum_{j \neq i} e_{i j} \rho_{j}
$$

Replacing $e_{i i}$ with $\frac{1}{N}-\sum_{j \neq i} e_{i j}$ and simplifying, we obtain

$$
\Delta d_{i} \rho_{i}=\sum_{j \neq i} e_{i j}\left(\rho_{j}-\rho_{i}\right) .
$$

Next we take the partial derivative of both sides with respect to $e_{k l}$, where $k \neq l$,

$$
\rho_{i} \frac{\partial\left(\Delta d_{i}\right)}{\partial e_{k l}}+\Delta d_{i} \frac{\partial \rho_{i}}{\partial e_{k l}}=\sum_{\substack{i, j \\ j \neq i}} \frac{\partial e_{i j}}{\partial e_{k l}}\left(\rho_{j}-\rho_{i}\right)+\sum_{\substack{i, j \\ j \neq i}} e_{i j}\left(\frac{\partial \rho_{j}}{\partial e_{k l}}-\frac{\partial \rho_{i}}{\partial e_{k l}}\right),
$$

and evaluate at the critical point ( $\rho_{i}^{*}=\frac{1}{N}$ and $\Delta d_{i}=0$ ):

$$
\frac{1}{N} \frac{\partial\left(\Delta d_{i}\right)}{\partial e_{k l}}=\sum_{j \neq i} e_{i j}^{*}\left(\frac{\partial \rho_{j}}{\partial e_{k l}}-\frac{\partial \rho_{i}}{\partial e_{k l}}\right) .
$$

Multiplying both sides by $\Delta e_{k l}$ and then summing over all $k, l \in\{1, \ldots, N\}$ for $k \neq l$ yields

$$
\begin{equation*}
\frac{1}{N} \sum_{\substack{k, l \\ k \neq l}} \frac{\partial\left(\Delta d_{i}\right)}{\partial e_{k l}} \Delta e_{k l}=\sum_{\substack{j, k, l \\ k \neq l \\ j \neq i}} e_{i j}^{*} \frac{\partial \rho_{j}}{\partial e_{k l}} \Delta e_{k l}-\sum_{\substack{j, k, l \\ k \neq l \\ j \neq i}} e_{i j}^{*} \frac{\partial \rho_{i}}{\partial e_{k l}} \Delta e_{k l} . \tag{22}
\end{equation*}
$$

We observe that $\rho_{i}^{(1)}=\sum_{k^{k l} \neq l} \frac{\partial \rho_{i}}{\partial e_{k l}} \Delta e_{k l}$ is the first order term of the Taylor series expansion of $\rho_{i}$ about the critical point, and $\sum_{k^{k} \neq l} \frac{\partial\left(\Delta d_{i}\right)}{\partial e_{k l}} \Delta e_{k l}=\Delta d_{i}$, since $\Delta d_{i}$ is a linear function of the $e_{k l}$. Thus Eq. (22) becomes

$$
\frac{1}{N} \Delta d_{i}=\sum_{j \neq i} e_{i j}^{*} \rho_{j}^{(1)}-\sum_{j \neq i} e_{i j}^{*} \rho_{i}^{(1)}
$$

The restriction $j \neq i$ in the sums on the right-hand side can be removed, since the entire righthand side is zero when $j=i$. Rearranging further, we have

$$
N \rho_{i}^{(1)} \sum_{j=1}^{N} e_{i j}^{*}=-\Delta d_{i}+N \sum_{j=1}^{N} e_{i j}^{*} \rho_{j}^{(1)} .
$$

Since the birth rate is uniform over sites $\left(b_{i}=\sum_{j=1}^{N} e_{i j}^{*}=\frac{1}{N}\right.$ for all $i$ ), the above equation simplifies to

$$
\begin{equation*}
\rho_{i}^{(1)}=-\Delta d_{i}+N \sum_{j=1}^{N} e_{i j}^{*} \rho_{j}^{(1)} \tag{23}
\end{equation*}
$$

This equation provides a recurrence relation for the first-order term of the Taylor expansion of $\rho_{i}$ about the critical point.

## Step 5: Show second-order term of the overall fixation probability is nonpositive

We now combine Eq. (23) with Eq. (21) to show that $\rho^{(2)} \leq 0$, with equality if and only if $\Delta d_{i}=0$ for all $i$, according to the following argument. We first multiply both sides of Eq. (23) by $\rho_{i}^{(1)}$ and sum over $i=1, \ldots, N$ :

$$
\sum_{i=1}^{N}\left(\rho_{i}^{(1)}\right)^{2}=-\sum_{i=1}^{N} \Delta d_{i} \rho_{i}^{(1)}+N \sum_{i, j} e_{i j}^{*} \rho_{i}^{(1)} \rho_{j}^{(1)}
$$

By Eq. (21) we can substitute $\rho^{(2)}$ for $\sum_{i=1}^{N} \Delta d_{i} \rho_{i}^{(1)}$. Then, solving for $\rho^{(2)}$, we obtain

$$
\begin{equation*}
\rho^{(2)}=-\sum_{i=1}^{N}\left(\rho_{i}^{(1)}\right)^{2}+N \sum_{i, j} e_{i j}^{*} \rho_{i}^{(1)} \rho_{j}^{(1)} \tag{24}
\end{equation*}
$$

We now make use of the fact that the product $\rho_{i}^{(1)} \rho_{j}^{(1)}$ can be written as a difference of squares:

$$
\rho_{i}^{(1)} \rho_{j}^{(1)}=-\frac{1}{2}\left[\left(\rho_{i}^{(1)}-\rho_{j}^{(1)}\right)^{2}-\left(\rho_{i}^{(1)}\right)^{2}-\left(\rho_{j}^{(1)}\right)^{2}\right] .
$$

Substituting this identity into Eq. (24) yields

$$
\begin{aligned}
\rho^{(2)}=- & \sum_{i=1}^{N}\left(\rho_{i}^{(1)}\right)^{2}-\frac{N}{2} \sum_{i, j} e_{i j}^{*}\left(\rho_{i}^{(1)}-\rho_{j}^{(1)}\right)^{2}+\frac{N}{2} \sum_{i, j} e_{i j}^{*}\left(\rho_{i}^{(1)}\right)^{2} \\
& +\frac{N}{2} \sum_{i, j} e_{i j}^{*}\left(\rho_{j}^{(1)}\right)^{2} \\
=- & \sum_{i=1}^{N}\left(\rho_{i}^{(1)}\right)^{2}-\frac{N}{2} \sum_{i, j} e_{i j}^{*}\left(\rho_{i}^{(1)}-\rho_{j}^{(1)}\right)^{2}+\frac{N}{2} \sum_{i=1}^{N}\left(\rho_{i}^{(1)}\right)^{2} \sum_{j=1}^{N} e_{i j}^{*} \\
& +\frac{N}{2} \sum_{j=1}^{N}\left(\rho_{j}^{(1)}\right)^{2} \sum_{i=1}^{N} e_{i j}^{*} .
\end{aligned}
$$

Uniform birth rates implies that $\sum_{j=1}^{N} e_{i j}^{*}=1 / N$ for each $i$. Furthermore, since $\left\{e_{i j}^{*}\right\}_{i, j}$ is a critical point, death rates are uniform as well; thus $\sum_{i=1}^{N} e_{i j}^{*}=1 / N$ for each $j$. Making these substitutions yields

$$
\begin{align*}
\rho^{(2)} & =-\sum_{i=1}^{N}\left(\rho_{i}^{(1)}\right)^{2}-\frac{N}{2} \sum_{i, j} e_{i j}^{*}\left(\rho_{i}^{(1)}-\rho_{j}^{(1)}\right)^{2}+\frac{1}{2} \sum_{i=1}^{N}\left(\rho_{i}^{(1)}\right)^{2}+\frac{1}{2} \sum_{j=1}^{N}\left(\rho_{j}^{(1)}\right)^{2}  \tag{25}\\
& =-\frac{N}{2} \sum_{i, j} e_{i j}^{*}\left(\rho_{i}^{(1)}-\rho_{j}^{(1)}\right)^{2} .
\end{align*}
$$

This shows that $\rho^{(2)} \leq 0$.
We now consider the case that $\rho^{(2)}=0$. By Eq. (25), this occurs if and only if $\rho_{i}^{(1)}=\rho_{j}^{(1)}$ for any pair $i, j$ with $e_{i j}^{*} \neq 0$. It follows that each term $e_{i j}^{*} \rho_{j}^{(1)}$ in the sum on the right-hand side of Eq. (23) can be replaced by $e_{i j}^{*} \rho_{i}^{(1)}$. Making these substitutions and using $\sum_{j=1}^{N} e_{i j}^{*}=1 / N$, we find that $\Delta d_{i}=0$. Thus $\rho^{(2)}=0$ implies that $\Delta d_{i}=0$ for each $i$.

Conversely, if $\Delta d_{i}=0$ for each $i$, then Eq. (21) implies that $\rho^{(2)}=0$. Overall, we have shown that $\rho^{(2)} \leq 0$ with equality if and only if $\Delta d_{i}=0$ for each $i$.

## Step 6: Show that the critical points are maxima

We now adopt a vector notation, in which $\mathbf{e}=\left\{e_{i j}\right\}_{\substack{i, j \\ i \neq j}}$ is an arbitrary combination of replacement probabilities satisfying the conditions of the theorem, and $\mathbf{e}^{*}=\left\{e_{i j}^{*}\right\}_{\substack{i, j \\ i \neq j}}$ is an arbitrary critical point of $\rho$.

By Taylor's theorem, we can rewrite the second-order expansion (19) of $\rho$ around $\mathbf{e}^{*}$ as

$$
\begin{equation*}
\rho(\mathbf{e})=\frac{1}{N}+\frac{1}{2}\left(\mathbf{e}-\mathbf{e}^{*}\right)^{T} H_{\mathrm{e}^{*}}\left(\mathbf{e}-\mathbf{e}^{*}\right)+\left(\mathbf{e}-\mathbf{e}^{*}\right)^{T} R_{\mathrm{e}, \mathbf{e}^{*}}\left(\mathbf{e}-\mathbf{e}^{*}\right), \tag{26}
\end{equation*}
$$

where $H_{\mathbf{e}^{*}}$ is the Hessian matrix of $\rho$ at $\mathbf{e}^{*}$ and $\left\{R_{\mathbf{e} \cdot \mathbf{e}^{*}}\right\}$ is a family of symmetric matrices with the property that $R_{\mathbf{e} \cdot \mathbf{e}^{*}}$ converges to the zero matrix as $\mathbf{e} \rightarrow \mathbf{e}^{*}$.

Since $H_{\mathrm{e}^{*}}$ and $R_{e ; \mathrm{e}^{*}}$ are symmetric, all eigenvalues of these matrices are real, and the eigenspaces corresponding to these eigenvalues are orthogonal. We showed in Step 5 that $\rho^{(2)} \leq 0$, which implies that $H_{\mathbf{e}^{*}}$ is negative semidefinite. We also showed that $\rho^{(2)}=0$ if and only if $\Delta d_{i}=$ 0 for all $i$; thus the null space of $H_{\mathbf{e}^{*}}$ is the vector space $D$ consisting of all displacement vectors for which $\Delta d_{i}=0$ for all $i$. It follows that $H_{\mathbf{e}^{*}}$ is negative definite as an operator on the orthogonal complement $D^{\perp}$ of $D$.

We now fix a particular critical point $\mathbf{e}_{0}^{*}$, and consider a neighborhood $N\left(\mathbf{e}_{0}^{*} ; \varepsilon\right)$ of the form

$$
N\left(\mathbf{e}_{0}^{*} ; \varepsilon\right)=\left\{\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w} \mid \mathbf{v} \in D,\|\mathbf{v}\|<\varepsilon, \mathbf{w} \in D^{\perp},\|\mathbf{w}\|<\varepsilon\right\} .
$$

For each point $\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w} \in N\left(\mathbf{e}_{0}^{*} ; \varepsilon\right)$, we observe that $\mathbf{e}_{0}^{*}+\mathbf{v}$ is also a critical point of $\rho$, since the addition of $\mathbf{v} \in D$ leaves the value of $d_{i}$ unchanged for each $i$. Now, applying the Taylor expansion (26) of $\rho\left(\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w}\right)$ about $\mathbf{e}_{0}^{*}+\mathbf{v}$, yields

$$
\begin{equation*}
\rho\left(\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w}\right)=\frac{1}{N}+\frac{1}{2} \mathbf{w}^{T}\left(H_{\mathrm{e}_{0}^{*}+\mathbf{v}}\right) \mathbf{w}+\mathbf{w}^{T}\left(R_{\mathrm{e}_{0}^{*}+\mathbf{v}+\mathbf{w} ; \mathbf{e}_{0}^{*}+\mathbf{v}}\right) \mathbf{w} . \tag{27}
\end{equation*}
$$

Let $\lambda_{H}(\mathbf{v})$ denote the largest nonzero eigenvalue of $H_{\mathrm{e}_{0}^{*}+\mathbf{v}}$, and let $\lambda_{R}(\mathbf{v}, \mathbf{w})$ denote the largest eigenvalue of $R_{\mathrm{e}_{0}^{*}+\mathbf{v}+\mathbf{w} ; \mathrm{e}_{0}^{*}+\mathrm{v}}$. Since $H_{\mathrm{e}_{0}^{*}+\mathrm{v}}$ is negative semidefinite, $\lambda_{H}(\mathbf{v})<0$. Further, since $H_{\mathrm{e}_{0}^{*}+\mathrm{v}}$ is negative definite as an operator on $D^{\perp}$ and since $\mathbf{w} \in D^{\perp}$, it follows that

$$
\begin{equation*}
\mathbf{w}^{T} H_{\mathrm{e}_{0}^{*}+\mathbf{v}} \mathbf{w} \leq \lambda_{H}(\mathbf{v})\|\mathbf{w}\|^{2} . \tag{28}
\end{equation*}
$$

We also have that, for each $\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w} \in N\left(\mathbf{e}_{0}^{*} ; \varepsilon\right)$,

$$
\begin{equation*}
\mathbf{w}^{T} R_{\mathrm{e}_{0}^{*}+\mathbf{v}+\mathbf{w} ; \mathrm{e}_{0}^{*}+\mathrm{v}} \mathbf{w} \leq \lambda_{R}(\mathbf{v}, \mathbf{w})\|\mathbf{w}\|^{2} \tag{29}
\end{equation*}
$$

Furthermore, since $R_{\mathrm{e}_{0}^{*}+\mathbf{v}+\mathbf{w} ; \mathrm{e}_{0}^{*}+\mathbf{v}}$ converges to the zero matrix as $\mathbf{w} \rightarrow \mathbf{0}$, we have $\lim _{\mathbf{w} \rightarrow \mathbf{0}} \lambda_{R}(\mathbf{v}$, $\mathbf{w})=0$ for each fixed $\mathbf{v}$. It follows that, for sufficiently small $\varepsilon>0$, all points $\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w} \in$ $N\left(\mathbf{e}_{0}^{*} ; \varepsilon\right)$ satisfy

$$
\begin{equation*}
\frac{1}{2} \lambda_{H}(\mathbf{v})+\lambda_{R}(\mathbf{v}, \mathbf{w})<0 . \tag{30}
\end{equation*}
$$

Combining Eqs. (27)-(30), we have that for all $\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w} \in N\left(\mathbf{e}_{0}^{*} ; \varepsilon\right)$ with $\varepsilon$ sufficiently small,

$$
\begin{equation*}
\rho\left(\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w}\right) \leq 1 / N+\left(\frac{1}{2} \lambda_{H}(\mathbf{v})+\lambda_{R}(\mathbf{v}, \mathbf{w})\right)\|\mathbf{w}\|^{2} \leq 1 / N . \tag{31}
\end{equation*}
$$

Equality is achieved in Eq. (31) if and only if $\mathbf{w}=\mathbf{0}$, in which case $\mathbf{e}_{0}^{*}+\mathbf{v}+\mathbf{w}$ is another critical point. Thus $\rho$ has a local maximum of $1 / N$ at every critical point. It follows that these points are global maxima as well. We showed in Step 2 that the critical points of $\rho$ are precisely those for which $d_{i}=1 / N$ for all $i$. This completes the proof.

## Result 4

We now turn to the full range of possible values for $K$ with no constraints on birth and death rates. Consider the example spatial structure illustrated in Fig. 5, which consists of a hub with outgoing edges to $n$ leaves, so that the population size is $N=n+1$. There is an edge of weight $a$, $0 \leq a<1 /(N-1)$, from the hub to each leaf. The hub also has a self-loop of weight $1-(N-1) a$, so that $B=1$ births are expected per time-step. The death rates are $d_{\mathrm{H}}=1-(N-1) a$ for the hub and $d_{\mathrm{L}}=a$ for each leaf. Solving Eqs. (1)-(2) we obtain the site-specific fixation probabilities $\rho_{\mathrm{L}}$ $=0$ for each leaf and $\rho_{\mathrm{H}}=1$ for the hub. By Eq. (3), the overall fixation probability is equal to the rate of turnover at the hub:

$$
\rho=d_{\mathrm{H}}=1-(N-1) a .
$$

Any value $0 \leq \rho<1$ can be obtained by an appropriate choice of $a$, with $0<a \leq 1 /(N-1)$, specifically, $a=(1-\rho) /(N-1)$. This proves:

Theorem 6 (Result 4). For arbitrary spatial population structure (no constraints on $e_{i j}$ ) the fixation probability can take any value $0 \leq \rho<1$, and consequently, the molecular clock can take any rate $0 \leq K<N u$.
We observe that if $a=1 / N$ then the death rate is $1 / N$ at each site, and therefore $\rho=1 / N$ by Result 1. If $a=1 /(N-1)$ then there is no turnover at the hub and thus $\rho=0$. At the other extreme, as $a$ approach zero, $\rho$ comes arbitrarily close to unity.

## Exact expressions for $\rho$ with $N \leq 3$

To complement the above arguments, which apply to general population size $N$, we here derive exact expressions for the overall fixation probability $\rho$ in the cases $N=2$ and $N=3$. Our results for node-specific fixation probabilities coincide with those found previous for a different model of evolution in a subdivided population [40].

## Population size $\mathbf{N}=\mathbf{2}$

We solve Eqs. (1)-(2) to obtain expressions for $\rho_{1}$ and $\rho_{2}$ :

$$
\begin{align*}
& \rho_{1}=\frac{e_{12}}{e_{12}+e_{21}}  \tag{32}\\
& \rho_{2}=\frac{e_{21}}{e_{12}+e_{21}} . \tag{33}
\end{align*}
$$

We first consider uniform birth rates ( $b_{i}=\sum_{j=1}^{2} e_{i j}=\frac{1}{2}$ for $i=1,2$ ). Substituting Eqs. (32) and (33) into Eq. (3) ( $B=1$ in this case), gives an expression for the overall fixation probability:

$$
\rho=\frac{1}{2}-\frac{\left(e_{12}-e_{21}\right)^{2}}{e_{12}+e_{21}}
$$

Thus, in accordance with Result $3, \rho$ is less than or equal to $\frac{1}{2}$ and consequently $K \leq u$. Equality occurs if and only if $e_{12}=e_{21}$, which corresponds to equal node-specific fixation probabilities ( $\rho_{1}=\rho_{2}=1 / N$ from Eqs. (32) and (33)) and equal death rates ( $d_{1}=d_{2}=1 / N$ ).

When the condition of uniform birth rates is relaxed, we find, by substituting Eqs. (32) and (33) into Eq. (3), the following expression for the overall fixation probability:

$$
\rho=\frac{1}{2}-\frac{\left(e_{12}-e_{21}\right)\left(\frac{B}{2}-d_{1}\right)}{B\left(e_{12}+e_{21}\right)} .
$$

We discover that $\rho=\frac{1}{2}$, and consequently $K=u$, if (i) death rates are equal, $d_{1}=d_{2}=\frac{B}{2}$ (Result 1), or (ii) node-specific fixation probabilities are equal, $\rho_{1}=\rho_{2}=\frac{1}{2}$ (Result 2). When death rates are unequal and node-specific fixation probabilities are unequal, there are cases in which $\rho<\frac{1}{2}(K<u)$ and cases in which $\rho>\frac{1}{2}(K>u)$. In particular, suppose that $d_{1}>\frac{B}{2}$. Then $\rho$ is less than $\frac{1}{2}$ if $e_{12}<e_{21}$ and greater than $\frac{1}{2}$ if $e_{12}>e_{21}$.

## Population size $\boldsymbol{N}=\mathbf{3}$

We solve Eqs. (1) and (2) to obtain expressions for $\rho_{1}, \rho_{2}$ and $\rho_{3}$ :

$$
\begin{align*}
& \rho_{1}=\frac{e_{12} e_{13}+e_{12} e_{23}+e_{13} e_{32}}{e_{12}\left(e_{13}+e_{23}+e_{31}\right)+e_{13}\left(e_{21}+e_{32}\right)+\left(e_{21}+e_{31}\right)\left(e_{23}+e_{32}\right)},  \tag{34}\\
& \rho_{2}=\frac{e_{21} e_{13}+e_{21} e_{23}+e_{23} e_{31}}{e_{12}\left(e_{13}+e_{23}+e_{31}\right)+e_{13}\left(e_{21}+e_{32}\right)+\left(e_{21}+e_{31}\right)\left(e_{23}+e_{32}\right)},  \tag{35}\\
& \rho_{3}=\frac{e_{12} e_{31}+e_{21} e_{32}+e_{31} e_{32}}{e_{12}\left(e_{13}+e_{23}+e_{31}\right)+e_{13}\left(e_{21}+e_{32}\right)+\left(e_{21}+e_{31}\right)\left(e_{23}+e_{32}\right)} . \tag{36}
\end{align*}
$$

Substituting Eqs. (34)-(36) into Eq. (3) gives an expression for the overall fixation probability $\rho=\frac{\text { num }}{\text { denom }}$ where

$$
\begin{aligned}
& \text { num }=d_{1}\left(e_{12} e_{13}+e_{12} e_{23}\right.\left.+e_{13} e_{32}\right)+d_{2}\left(e_{21} e_{13}+e_{21} e_{23}+e_{23} e_{31}\right) \\
&+d_{3}\left(e_{12} e_{31}+e_{21} e_{32}+e_{31} e_{32}\right) \\
& \text { denom }=B\left(e_{12}\left(e_{13}+e_{23}+e_{31}\right)+e_{13}\left(e_{21}+e_{32}\right)+\left(e_{21}+e_{31}\right)\left(e_{23}+e_{32}\right)\right) .
\end{aligned}
$$

We factor to obtain an expression for num $(\Delta \rho)$ in terms of $\Delta d_{i}=d_{i}-\frac{B}{3}$ for $i \in\{1,2,3\}$,

$$
\begin{equation*}
\operatorname{num}(\Delta \rho)=\left(e_{12} \tilde{\Delta} d_{1}+e_{21} \tilde{\Delta} d_{2}\right)\left(d_{3}-b_{3}\right)-e_{32} \tilde{\Delta} d_{1}\left(d_{1}-b_{1}\right)-e_{31} \tilde{\Delta} d_{2}\left(d_{2}-b_{2}\right) \tag{37}
\end{equation*}
$$

From Eq. (37) we see that $\rho=\frac{1}{3}$, and consequently $K=u$, in the case of equal death rates $\left(d_{1}=d_{2}=d_{3}=\frac{B}{3}\right)$ and in the case of equivalent birth and death rate at each site ( $b_{i}=d_{i}$ for $i \in$ $\{1,2,3\})$. This conforms to Result 1 and 2. When death rates are nonuniform and node-specific probabilities are also nonuniform, we find cases in which $K>u$ and $K<u$. For example, if $\tilde{\Delta} d_{1}=0, \Delta d_{2}>0, d_{2}>b_{2}$, and $d_{3}<b_{3}$ then Eq. (37) yields num $(\Delta \rho)<0$ and consequently, $\rho<\frac{1}{3}$ and $K<u$. On the other hand, if $\Delta d_{1}=0, b_{2}>d_{2}>\frac{B}{3}$, and $d_{3}>b_{3}$ then Eq. (37) yields num $(\Delta \rho)>0$ and therefore, $\rho>\frac{1}{3}$ and $K>u$.

## Upstream-downstream populations

We now turn to the upstream-downstream model introduced in the Results and in Fig. 6.

Theorem 7. In the upstream-downstream model, $\rho>1 / N$, and consequently $K>u$, if and only if $d_{\uparrow}>d_{\downarrow}$.

Proof. From Eq. (3) we obtain the following expression for the overall fixation probability

$$
\begin{align*}
\rho & =\frac{1}{B}\left(N_{\uparrow} d_{\uparrow} \rho_{\uparrow}+N_{\downarrow} d_{\downarrow} \rho_{\downarrow}\right) \\
& =\frac{1}{N}+\frac{1}{B}\left[N_{\uparrow} d_{\uparrow}\left(\rho_{\uparrow}-\frac{1}{N}\right)+N_{\downarrow} d_{\downarrow}\left(\rho_{\downarrow}-\frac{1}{N}\right)\right] . \tag{38}
\end{align*}
$$

Since fixation probabilities sum to one, and since the total population size is $N=N_{\uparrow}+N_{\downarrow}$, we have

$$
\begin{equation*}
N_{\downarrow}\left(\rho_{\downarrow}-\frac{1}{N}\right)=-N_{\uparrow}\left(\rho_{\uparrow}-\frac{1}{N}\right) \tag{39}
\end{equation*}
$$

COMPUTATIONAL BIOLOGY

Substituting in Eq. (38) yields

$$
\begin{equation*}
\rho=\frac{1}{N}+\left(d_{\uparrow}-d_{\downarrow}\right) \frac{N_{\uparrow}}{B}\left(\rho_{\uparrow}-\frac{1}{N}\right) . \tag{40}
\end{equation*}
$$

It follows from Eq. (5) and from $e_{\rightarrow}>e_{\leftarrow}$ that $\rho_{\uparrow}>\rho_{\downarrow}$. Moreover, Eq. (39) implies that $\rho_{\uparrow}-1 / N$ and $\rho_{\downarrow}-1 / N$ have opposite signs, and since $\rho_{\uparrow}>\rho_{\downarrow}$, it follows that $\rho_{\uparrow}-1 / N$ must be positive. Thus the second term on the right-hand side of Eq. (40) has the sign of $d_{\uparrow}-d_{\downarrow}$. We therefore conclude that $\rho>1 / N$, and consequently the molecular clock is accelerated relative to the well-mixed case $(K>u)$, if and only if $d_{\uparrow}>d_{\downarrow}$.

## Acknowledgments

We are grateful to Joshua B. Plotkin and David McCandlish for helpful comments.

## Author Contributions

Conceived and designed the experiments: BA CS YD MAN. Performed the experiments: BA CS YD RCM CP. Analyzed the data: BA CS YD RCM CP. Wrote the paper: BA CS YD RCM CP MAN.

## References

1. Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. Evolving genes and proteins 97: 97-166. doi: 10.1016/B978-1-4832-2734-4.50017-6
2. Kimura M (1968) Evolutionary rate at the molecular level. Nature 217: 624-626. doi: 10.1038/ 217624a0 PMID: 5637732
3. Ayala FJ (1999) Molecular clock mirages. BioEssays 21:71-75. doi: 10.1002/(SICI)1521-1878 (199901)21:1\%3C71::AID-BIES9\%3E3.3.CO;2-2 PMID: 10070256
4. Bromham L, Penny D (2003) The modern molecular clock. Nature Reviews Genetics 4: 216-224. doi: 10.1038/nrg1020 PMID: 12610526
5. Pollak E (1982) The rate of mutant substitution in populations with overlapping generations. Genetical Research 40: 89-94. doi: 10.1017/S0016672300018930 PMID: 7141224
6. Balloux F, Lehmann $L$ (2012) Substitution rates at neutral genes depend on population size under fluctuating demography and overlapping generations. Evolution 66: 605-611. doi: 10.1111/j.15585646.2011.01458.x PMID: 22276552
7. Lehmann $L$ (2014) Stochastic demography and the neutral substitution rate in class-structured populations. Genetics 197: 351-360. doi: 10.1534/genetics.114.163345 PMID: 24594520
8. Ohta T, Kimura M (1971) On the constancy of the evolutionary rate of cistrons. Journal of Molecular Evolution 1: 18-25. doi: 10.1007/BF01659391 PMID: 4377445
9. Kimura M, Ohta T (1974) On some principles governing molecular evolution. Proceedings of the National Academy of Sciences 71: 2848-2852. doi: 10.1073/pnas.71.7.2848
10. Langley CH , Fitch WM (1974) An examination of the constancy of the rate of molecular evolution. Journal of Molecular Evolution 3: 161-177. doi: 10.1007/BF01797451 PMID: 4368400
11. Kimura M (1984) The Neutral Theory of Molecular Evolution. Cambridge University Press.
12. Ayala FJ (1986) On the virtues and pitfalls of the molecular evolutionary clock. Journal of Heredity 77: 226-235. PMID: 3020121
13. Li WH, Tanimura M, Sharp PM (1987) An evaluation of the molecular clock hypothesis using mammalian dna sequences. Journal of Molecular Evolution 25: 330-342. doi: 10.1007/BF02603118 PMID: 3118047
14. Kumar S (2005) Molecular clocks: four decades of evolution. Nature Reviews Genetics 6: 654-662. doi: 10.1038/nrg1659 PMID: 16136655
15. Maruyama T (1970) On the fixation probability of mutant genes in a subdivided population. Genetical Research 15: 221-225. doi: 10.1017/S0016672300001543 PMID: 5480754
16. Lande $R$ (1979) Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. Evolution 33: 234-251. doi: 10.2307/2407380

COMPUTATIONAL BIOLOGY
17. Slatkin $M$ (1981) Fixation probabilities and fixation times in a subdivided population. Evolution: 477488. doi: 10.2307/2408196
18. Nagylaki T (1982) Geographical invariance in population genetics. Journal of Theoretical Biology 99: 159-172. doi: 10.1016/0022-5193(82)90396-4 PMID: 7169795
19. Tachida H, lizuka M (1991) Fixation probability in spatially changing environments. Genetical Research 58: 243-251. doi: 10.1017/S0016672300029992 PMID: 1802806
20. Barton N (1993) The probability of fixation of a favoured allele in a subdivided population. Genetical Research 62: 149-149. doi: 10.1017/S0016672300031748
21. Roze D, Rousset $F$ (2003) Selection and drift in subdivided populations: a straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. Genetics 165: 2153-2166. PMID: 14704194
22. Whitlock MC (2003) Fixation probability and time in subdivided populations. Genetics 164: 767-779. PMID: 12807795
23. Lieberman E, Hauert C, Nowak MA (2005) Evolutionary dynamics on graphs. Nature 433: 312-316. doi: 10.1038/nature03204 PMID: 15662424
24. Broom M, Rychtář J (2008) 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 Science 464: 2609-2627. doi: 10.1098/rspa.2008.0058
25. Patwa Z, Wahl LM (2008) The fixation probability of beneficial mutations. Journal of The Royal Society Interface 5: 1279-1289. doi: 10.1098/rsif.2008.0248
26. Masuda N, Ohtsuki H (2009) Evolutionary dynamics and fixation probabilities in directed networks. New Journal of Physics 11:033012. doi: 10.1088/1367-2630/11/3/033012
27. Houchmandzadeh B, Vallade $M$ (2011) The fixation probability of a beneficial mutation in a geographically structured population. New Journal of Physics 13: 073020. doi: 10.1088/1367-2630/13/7/ 073020
28. Shakarian P, Roos P, Johnson A (2012) A review of evolutionary graph theory with applications to game theory. Biosystems 107: 66-80. doi: 10.1016/j.biosystems.2011.09.006 PMID: 22020107
29. Voorhees $B$ (2013) Birth-death fixation probabilities for structured populations. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 469: 20120248. doi: 10.1098/ rspa.2012.0248
30. Constable GWA, McKane AJ (2014) Population genetics on islands connected by an arbitrary network: An analytic approach. Journal of Theoretical Biology 358: 149-165. doi: 10.1016/j.jtbi.2014.05. 033 PMID: 24882790
31. Monk T, Green P, Paulin M (2014) Martingales and fixation probabilities of evolutionary graphs. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 470: 20130730. doi: 10.1098/rspa. 2013.0730
32. Allen B, Tarnita CE (2014) Measures of success in a class of evolutionary models with fixed population size and structure. Journal of Mathematical Biology 68: 109-143. doi: 10.1007/s00285-012-0622-x PMID: 23179131
33. Liggett TM (1985) Particle Systems. Springer.
34. Nowak MA, Bonhoeffer S, May RM (1994) More spatial games. International Journal of Bifurcation and Chaos 4:33-56. doi: 10.1142/S0218127494000046
35. Antal T, Redner S, Sood V (2006) Evolutionary dynamics on degree-heterogeneous graphs. Physical Review Letters 96: 188104. doi: 10.1103/PhysRevLett.96.188104 PMID: 16712402
36. Nowak MA, May RM (1992) Evolutionary games and spatial chaos. Nature 359: 826-829. doi: 10 1038/359826a0
37. Taylor P, Lillicrap T, Cownden D (2011) Inclusive fitness analysis on mathematical groups. Evolution 65: 849-859. doi: 10.1111/j.1558-5646.2010.01162.x PMID: 21044061
38. Wakano JY, Ohtsuki H, Kobayashi Y (2012) A mathematical description of the inclusive fitness theory. Theoretical Population Biology.
39. Maciejewski W (2014) Reproductive value in graph-structured populations. Journal of Theoretical Biology 340: 285-293. doi: 10.1016/j.jtbi.2013.09.032 PMID: 24096097
40. Lundy IJ, Possingham HP (1998) Fixation probability of an allele in a subdivided population with asymmetric migration. Genetical Research 71: 237-245. doi: 10.1017/S0016672398003139
41. Kozar S, Morrissey E, Nicholson AM, van der Heijden M, Zecchini HI, et al. (2013) Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. Cell Stem Cell 13: 626-633. doi: 10.1016/j.stem.2013.08.001 PMID: 24035355

COMPUTATIONAL BIOLOGY
42. Lopez-Garcia C, Klein AM, Simons BD, Winton DJ (2010) Intestinal stem cell replacement follows a pattern of neutral drift. Science 330: 822-825. doi: 10.1126/science. 1196236 PMID: 20929733
43. Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, et al. (2010) Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Igr5 stem cells. Cell 143: 134144. doi: 10.1016/j.cell.2010.09.016 PMID: 20887898
44. Vermeulen L, Snippert HJ (2014) Stem cell dynamics in homeostasis and cancer of the intestine. Nature Reviews Cancer 14: 468-480. doi: 10.1038/nrc3744 PMID: 24920463
45. Vermeulen L, Morrissey E, van der Heijden M, Nicholson AM, Sottoriva A, et al. (2013) Defining stem cell dynamics in models of intestinal tumor initiation. Science 342: 995-998. doi: 10.1126/science. 1243148 PMID: 24264992
46. Barker N, van Oudenaarden A, Clevers H (2012) Identifying the stem cell of the intestinal crypt: strategies and pitfalls. Cell Stem Cell 11: 452-460. doi: 10.1016/j.stem.2012.09.009 PMID: 23040474
47. Leskovec $J$ (2011). Stanford large network dataset collection. http://snap.stanford.edu/data/index. html URL http://snap.stanford.edu/data/index.html
48. Durrett R, Levin S (1994) The importance of being discrete (and spatial). Theoretical Population Biology 46: 363-394. doi: 10.1006/tpbi.1994.1032
49. Nakamaru M, Matsuda H, Iwasa Y (1997) The evolution of cooperation in a lattice-structured population. Journal of Theoretical Biology 184: 65-81. doi: 10.1006/jtbi.1996.0243 PMID: 9039401
50. van Baalen M, Rand DA (1998) The unit of selection in viscous populations and the evolution of altruism. Journal of Theoretical Biology 193: 631-648. doi: 10.1006/jtbi.1998.0730 PMID: 9750181
51. Ohtsuki H, Hauert C, Lieberman E, Nowak MA (2006) A simple rule for the evolution of cooperation on graphs and social networks. Nature 441:502-505. doi: 10.1038/nature04605 PMID: 16724065
52. Nowak MA, Tarnita CE, Antal T (2010) Evolutionary dynamics in structured populations. Philosophical Transactions of the Royal Society B: Biological Sciences 365: 19-30. doi: 10.1098/rstb.2009.0215
53. Allen B, Nowak MA (2014) Games on graphs. EMS Surveys in Mathematical Sciences 1:113-151. doi: 10.4171/EMSS/3
54. Débarre F, Hauert C, Doebeli M (2014) Social evolution in structured populations. Nature Communications 5: 4409. doi: 10.1038/ncomms4409 PMID: 24598979
55. Felsenstein $J$ (1976) The theoretical population genetics of variable selection and migration. Annual Review of Genetics 10: 253-280. doi: 10.1146/annurev.ge.10.120176.001345 PMID: 797310
56. Hedrick PW, Ginevan ME, Ewing EP (1976) Genetic polymorphism in heterogeneous environments. Annual Review of Ecology and Systematics 7: 1-32. doi: 10.1146/annurev.es.07.110176.000245
57. Slatkin M (1987) Gene flow and the geographic structure of natural populations. Science 236: 787792. doi: 10.1126/science. 3576198 PMID: 3576198
58. Cherry JL, Wakeley J (2003) A diffusion approximation for selection and drift in a subdivided population. Genetics 163: 421-428. PMID: 12586727
59. MacArthur RH, Wilson EO (1967) The Theory of Island Biogeography. Princeton, NJ: Princeton University Press.
60. Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. Nature 421:259-264 doi: 10.1038/nature01274 PMID: 12529641
61. De Aguiar M, Baranger M, Baptestini E, Kaufman L, Bar-Yam Y (2009) Global patterns of speciation and diversity. Nature 460: 384-387. doi: 10.1038/nature08168 PMID: 19606148
62. Kawecki TJ, Holt RD (2002) Evolutionary consequences of asymmetric dispersal rates. The American Naturalist 160: 333-347. doi: 10.1086/341519 PMID: 18707443
63. Lenormand $T$ (2002) Gene flow and the limits to natural selection. Trends in Ecology \& Evolution 17: 183-189. doi: 10.1016/S0169-5347(02)02497-7
64. Garant D, Forde SE, Hendry AP (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. Functional Ecology 21: 434-443. doi: 10.1111/j.1365-2435.2006.01228.x
65. Maciejewski W, Fu F, Hauert C (2014) Evolutionary game dynamics in populations with heterogenous structures. PLoS Computational Biology 10: e1003567. doi: 10.1371/journal.pcbi. 1003567 PMID: 24762474
66. Nagylaki T (1978) Clines with asymmetric migration. Genetics 88: 813-827. PMID: 17248820
67. Vuilleumier S, Wilcox C, Cairns BJ, Possingham HP (2007) How patch configuration affects the impact of disturbances on metapopulation persistence. Theoretical Population Biology 72: 77-85. doi: 10.1016/j.tpb.2006.11.001 PMID: 17275866

COMPUTATIONAL BIOLOGY
68. Morrissey MB, de Kerckhove DT (2009) The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. The American Naturalist 174:875-889. doi: 10.1086/648311 PMID: 19860541
69. Tackenberg O, Poschlod P, Bonn S (2003) Assessment of wind dispersal potential in plant species. Ecological Monographs 73: 191-205. doi: 10.1890/0012-9615(2003)073\%5B0191:AOWDPI\%5D2. 0.CO;2
70. Muñoz J, Felicísimo ÁM, Cabezas F, Burgaz AR, Martínez I (2004) Wind as a long-distance dispersal vehicle in the southern hemisphere. Science 304: 1144-1147. doi: 10.1126/science. 1095210 PMID: 15155945
71. Collins CJ, Fraser CI, Ashcroft A, Waters JM (2010) Asymmetric dispersal of southern bull-kelp (durvillaea antarctica) adults in coastal new zealand: testing an oceanographic hypothesis. Molecular Ecology 19: 4572-4580. doi: 10.1111/j.1365-294X.2010.04842.x PMID: 20875065
72. Pringle JM, Blakeslee AM, Byers JE, Roman J (2011) Asymmetric dispersal allows an upstream region to control population structure throughout a speciesí range. Proceedings of the National Academy of Sciences 108: 15288-15293. doi: 10.1073/pnas. 1100473108 PMID: 21876126
73. Angert AL (2009) The niche, limits to species' distributions, and spatiotemporal variation in demography across the elevation ranges of two monkeyflowers. Proceedings of the National Academy of Sciences 106: 19693-19698. doi: 10.1073/pnas. 0901652106 PMID: 19805178
74. O'keefe K, Ramakrishnan U, Van Tuinen M, Hadly EA (2009) Source-sink dynamics structure a common montane mammal. Molecular Ecology 18: 4775-4789. doi: 10.1111/j.1365-294X.2009.04382.x PMID: 19863718
75. Sun JX, Helgason A, Masson G, Ebenesersdóttir SS, Li H, et al. (2012) A direct characterization of human mutation based on microsatellites. Nature genetics 44:1161-1165. doi: 10.1038/ng.2398 PMID: 22922873
76. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, et al. (2012) Rate of de novo mutations and the importance of father/'s age to disease risk. Nature 488: 471-475. doi: 10.1038/nature11396 PMID: 22914163
77. Campbell CD, Eichler EE (2013) Properties and rates of germline mutations in humans. Trends in $\mathrm{Ge}-$ netics 29: 575-584. doi: 10.1016/j.tig.2013.04.005 PMID: 23684843
78. Knudson AG (2001) Two genetic hits (more or less) to cancer. Nature Reviews Cancer 1: 157-162. doi: 10.1038/35101031 PMID: 11905807
79. Attolini CSO, Michor F (2009) Evolutionary theory of cancer. Annals of the New York Academy of Sciences 1168: 23-51. doi: 10.1111/j.1749-6632.2009.04880.x PMID: 19566702
80. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, et al. (2013) Cancer genome landscapes. Science 339: 1546-1558. doi: 10.1126/science. 1235122 PMID: 23539594
81. Goldie JH, Coldman AJ (1998) Drug resistance in cancer: mechanisms and models. Cambridge University Press.
82. Komarova NL, Wodarz D (2005) Drug resistance in cancer: principles of emergence and prevention. Proceedings of the National Academy of Sciences 102: 9714-9719. doi: 10.1073/pnas. 0501870102 PMID: 15980154
83. Michor F, Nowak MA, Iwasa Y (2006) Evolution of resistance to cancer therapy. Current Pharmaceutical Design 12: 261-271. doi: 10.2174/138161206775201956 PMID: 16454743
84. Bozic I, Reiter JG, Allen B, Antal T, Chatterjee K, et al. (2013) Evolutionary dynamics of cancer in response to targeted combination therapy. Elife 2. doi: 10.7554/eLife. 00747 PMID: 23805382
85. Nowak MA, Michor F, Iwasa Y (2003) The linear process of somatic evolution. Proceedings of the National Academy of Sciences 100: 14966-14969. doi: 10.1073/pnas.2535419100 PMID: 14657359
86. Dingli D, Traulsen A, Pacheco JM (2007) Stochastic dynamics of hematopoietic tumor stem cells. Cell Cycle 6: 461-466. doi: 10.4161/cc.6.4.3853 PMID: 17329969
87. Bozic I, Nowak MA (2013) Unwanted evolution. Science 342: 938-939. doi: 10.1126/science. 1247887 PMID: 24264980
88. Snippert HJ, Schepers AG, van Es JH, Simons BD, Clevers H (2014) Biased competition between Igr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. EMBO Rep 15: 6269. doi: 10.1002/embr. 201337799 PMID: 24355609
89. Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, et al. (2011) Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell 146: 633-644. doi: 10.1016/j.cell.2011. 08.030 PMID: 21854987
90. May RM, Lloyd AL (2001) Infection dynamics on scale-free networks. Physical Review E 64: 066112. doi: 10.1103/PhysRevE.64.066112

COMPUTATIONAL BIOLOGY
91. Watts DJ (2002) A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences 99: 5766-5771. doi: 10.1073/pnas. 082090499 PMID: 16578874
92. Christakis NA, Fowler JH (2007) The spread of obesity in a large social network over 32 years. New England Journal of Medicine 357: 370-379. doi: 10.1056/NEJMsa066082 PMID: 17652652
93. Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences 105: 1118-1123. doi: 10.1073/pnas. 0706851105 PMID: 18216267
94. Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. Reviews of Modern Physics 81:591. doi: 10.1103/RevModPhys.81.591
95. Centola D (2010) The spread of behavior in an online social network experiment. Science 329: 11941197. doi: 10.1126/science. 1185231 PMID: 20813952
96. Hill AL, Rand DG, Nowak MA, Christakis NA (2010) Infectious disease modeling of social contagion in networks. PLoS Computational Biology 6: e1000968. doi: 10.1371/journal.pcbi. 1000968 PMID: 21079667
97. Bakshy E, Rosenn I, Marlow C, Adamic L (2012) The role of social networks in information diffusion. In: Proceedings of the 21st international conference on World Wide Web. ACM, pp. 519-528. doi: 10. 1145/2187836.2187907
98. Moreno Y, Pastor-Satorras R, Vespignani A (2002) Epidemic outbreaks in complex heterogeneous networks. The European Physical Journal B-Condensed Matter and Complex Systems 26: 521-529. doi: 10.1007/s10051-002-8996-y
99. Nekovee M, Moreno Y, Bianconi G, Marsili M (2007) Theory of rumour spreading in complex social networks. Physica A: Statistical Mechanics and its Applications 374: 457-470. doi: 10.1016/j.physa. 2006.07.017
100. Skyrms B, Pemantle R (2009) A dynamic model of social network formation. In: Adaptive Networks, Springer. pp. 231-251.
101. Rosvall M, Bergstrom CT (2010) Mapping change in large networks. PloS ONE 5: e8694. doi: 10. 1371/journal.pone. 0008694 PMID: 20111700
102. Rand DG, Arbesman S, Christakis NA (2011) Dynamic social networks promote cooperation in experiments with humans. Proceedings of the National Academy of Sciences 108: 19193-19198. doi: 10. 1073/pnas. 1108243108 PMID: 22084103
103. Wardil L, Hauert C (2014) Origin and structure of dynamic cooperative networks. Scientific Reports 4. doi: 10.1038/srep05725 PMID: 25030202
104. Baxter GJ, Blythe RA, McKane AJ (2008) Fixation and consensus times on a network: A unified approach. Physical Review Letters 101:258701. doi: 10.1103/PhysRevLett.101.258701 PMID: 19113759
105. Frean M, Rainey PB, Traulsen A (2013) The effect of population structure on the rate of evolution. Proceedings of the Royal Society B: Biological Sciences 280. doi: 10.1098/rspb.2013.0211 PMID: 23677339
106. Tarnita CE, Taylor PD (2014) Measures of relative fitness of social behaviors in finite structured population models. The American Naturalist 184:477-488. doi: 10.1086/677924 PMID: 25226183
107. Broom M, Hadjichrysanthou C, Rychtár J, Stadler B (2010) Two results on evolutionary processes on general non-directed graphs. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 466: 2795-2798. doi: 10.1098/rspa.2010.0067
108. Shakarian P, Roos P, Moores $G$ (2013) A novel analytical method for evolutionary graph theory problems. Biosystems 111:136-144. doi: 10.1016/j.biosystems.2013.01.006 PMID: 23353025


[^0]:    doi:10.1371/journal.pcbi.1004108.g008

