The pace of evolution across fitness valleys

Chaitanya S. Gokhale [a,*], Yoh Iwasa [b], Martin A. Nowak [c], Arne Traulsen [a]

[a] *Max-Planck-Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany*
[b] *Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka 812-8581, Japan*
[c] *Program for Evolutionary Dynamics, Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA*

ABSTRACT

How fast does a population evolve from one fitness peak to another? We study the dynamics of evolving, asexually reproducing populations in which a certain number of mutations jointly confer a fitness advantage. We consider the time until a population has evolved from one fitness peak to another one with a higher fitness. The order of mutations can either be fixed or random. If the order of mutations is fixed, then the population follows a metaphorical ridge, a single path. If the order of mutations is arbitrary, then there are many ways to evolve to the higher fitness state. We address the time required for fixation in such scenarios and study how it is affected by the order of mutations, the population size, the fitness values and the mutation rate.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Evolutionary dynamics is based on natural selection, mutation and genetic drift (Nowak, 2006). It can be illustrated as the dynamics of a population in an abstract, typically high-dimensional fitness landscape. Since individuals with higher fitness produce more offspring, the average density of individuals is highest close to the fitness maxima. Many such features as the stationary population density in the fitness landscape or the mutation rate under which a population can still be concentrated around a fitness maximum have been addressed (Eigen and Schuster, 1977; Eigen et al., 1989; Wilke, 2005; Nowak, 1992). An important question is how a population evolves towards a fitness peak via several intermediate states. If the intermediate states have the same fitness as the initial state, then evolution to higher fitness states is neutral at first and thus poses no significant problems (van Nimwegen and Crutchfield, 2000). If the intermediate states have lower fitness than the initial state, then a fitness valley has to be overcome and it is more difficult to reach the fitness peak (Weinreich et al., 2006; Poelwijk et al., 2007). In this case, population stuck on a local peak cannot escape by natural selection alone, because there is no evolutionary trajectory with successively advantageous mutations. Instead, neutral genetic drift becomes important.

Here, we consider the dynamics of these systems from a different perspective. We address the average time a population needs to transfer from one peak to another one. For small mutation rates and finite populations, we calculate this average time analytically. When mutation rates are high, we can describe the system by a set of differential equations and obtain the relevant times from a numerical integration of the differential equations. In this framework, the relevant question is how fast a population evolves (Traulsen et al., 2007).

In particular, we can address the question whether a population evolves faster from one peak to another via d mutations if

(i) mutations have to occur in a certain order, i.e. only a single evolutionary trajectory is available, or
(ii) the order of the mutations does not matter, i.e. there are $d!$ evolutionary paths.

In the simplest case the intermediate fitness values are identical in both the cases and equal to that of the initial state. Thus the only difference remaining is the number of available paths. When the order of mutations is not fixed then multiple paths are available and the evolutionary dynamics will be faster when compared to a single path. We can then ask the question: Does a population evolve faster on a narrow ridge or a broad valley? This implies that we move away from the simplest case mentioned above and decrease the fitness in the intermediate states of the multiple paths compared to the fitness in the intermediate states of the single path. We show how the pace of evolution depends on the depth of the valley, the number of intermediate states and the size of the population.

In general, evolutionary dynamics depends crucially on the size of the population. In a small population a single mutation will typically reach fixation or extinction before another mutation can

* Corresponding author. Tel.: +49 4522 7630.
  E-mail address: gokhale@evolbio.mpg.de (C.S. Gokhale).

arise. The population moves as a whole step by step on the fitness landscape. For large populations, even for small mutation rates usually multiple types arise at the same time. This results in a non-zero population density in many states at the same time. For intermediate mutation rates, the population can either move stepwise across the fitness landscape or move several steps without getting concentrated in one of the intermediate states. This phenomenon has been termed stochastic tunneling (Iwasa et al., 2004). If the mutation rates are too small, tunneling does not occur because it is unlikely that a second mutation arises before the first one has reached fixation or has gone extinct. If the mutation rates are high, tunneling occurs trivially, because the system can be approximated by differential equations for the densities in the different states. These different scenarios including the limiting cases of stepwise evolution (typical for small populations) and continuous evolution (typical for large populations) can also be observed when the population size is kept constant, but the mutation rates are increased. For computer simulations increasing the mutation rate is more convenient than simulating huge populations for moderate mutation rates.

One important example for an evolutionary process in which the timescales are of crucial importance is the somatic evolution of cancer (Frank, 2007). Cancer progression has been investigated mathematically since the 1950s (Fisher, 1959; Nordling, 1953; Armitage and Doll, 1954). Of special interest are the tumor suppressor genes (Knudson, 1971; Michor et al., 2004). In a normal cell, there are two alleles of the tumor suppressor gene. The mutation in the first allele is neutral if the second wild-type allele can sufficiently perform the function. Inactivation of both the alleles confers a selective advantage to the cell and can lead to cancer progression. This is an example in which the order of mutations does not matter. Many cancers also require certain particular mutations that initiate cancer growth and pave the way for the accumulation of further mutations (Vogelstein and Kinzler, 2004). Recently, it has been shown that after cancer initiation, a large number of different mutations may be involved in cancer progression (Sjöblom et al., 2006; Wood et al., 2007; Jones et al., 2008a, b). So far, it is unclear if the mutations have to occur in a specific order or if there is more variation in the order (Beerenwinkel et al., 2007; Gerstung and Beerenwinkel, 2008).

For simplicity, we consider only very simple fitness landscapes here in which the fitness in all the intermediate states is identical. In natural systems, these fitness values will differ and also the mutation rate may not be constant. In addition, sometimes the order of mutations will matter and sometimes, it will not. Thus, sometimes a particular mutation will be a prerequisite to obtain a new function, but sometimes new mutations do not require any prerequisites. For example, this is the case in the evolution of resistance to β lactam antibiotics studied by Weinreich (2005) and Weinreich et al. (2006). However, here we focus on a very simple model to highlight the general aspects of the dynamics by analytical and numerical considerations.

This paper is organized as follows. We begin with the description of the two ways to order the mutations, the single path and the hypercube. We then derive analytical approximations of the fixation times for small mutation rates and discuss the effect of the different parameters on the fixation times. Next, we address the dynamics for intermediate and high mutation rates. Finally, we explore biological examples which can be modeled using this approach.

2. Model

To model evolutionary dynamics in a haploid population of size N, we use the Moran process (Nowak, 2006; Moran, 1962). In each time step, one individual is selected at random, but proportional to fitness. It produces one offspring, which replaces a randomly chosen individual. In one generation, each individual reproduces on average once. During reproduction, mutations occur with probability μ. We are interested in the time it takes until d mutations reach fixation in the population, starting from a homogeneous population in the initial state without any mutants. Moreover, we aim to explore the dynamical features of this process. We restrict ourselves to two different cases that allow the derivation of some analytical results.

2.1. Single path

If the mutations can occur only in a particular order, we have a single evolutionary path, see Fig. 1 for an illustration. Individuals in the initial state have fitness $r_0 = 1$ and individuals in the final state have fitness $r_d > 1$. It is instructive to characterize an individual by a string of d sites, which can either be wild-type or mutated. If the order of mutations is fixed, then a particular mutation requires another particular mutation as a prerequisite. For simplicity, we assume that all the $d - 1$ intermediate states have the same fitness $r_j = s < r_d$ ($j = 1, \ldots, d - 1$). For $s < 1$, the joint effect of the set of mutations make up for the loss of fitness caused by the individually deleterious mutations. This can be considered as a very special case of epistasis (Weinreich et al., 2005).

2.2. Hypercube

If the order of mutations does not matter, evolutionary dynamics takes place on a hypercube in d dimensions cf. Fig. 1. Thus, there are 2^d different types of individuals. In the initial state, we have d possible mutations. In the next step, $d - 1$ mutations are available. Consequently, we have $d!$ possible paths to fixation. Again, we assume $r_0 = 1$ and $r_d > 1$. Further, all individuals with some, but not all mutations have fitness $s < r_d$.

If the mutation probability is small, we do not need to make specific assumptions on the mutation process. But when the mutation probability increases, we can no longer be certain that only a single mutation occurs during reproduction. For simplicity, we do not consider the possibility of backward mutations. Although back mutations are often relevant, especially to escape from evolutionary dead ends (DePristo et al., 2007), it is not straightforward to define the speed of evolution in a system with
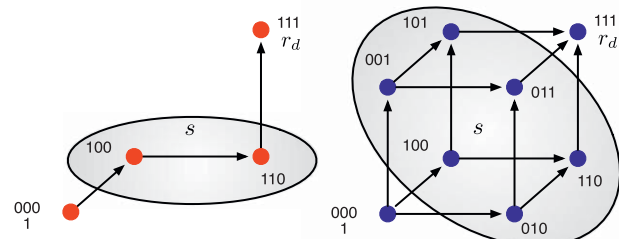

Fig. 1. The order of mutations determines the geometry for evolutionary dynamics, shown here for $d = 3$ sites (e.g. genes, nucleotide sites etc.). If mutations can only occur in a particular order, only a single path is available (left). If the order of mutations is arbitrary, evolutionary dynamics occurs on a hypercube (right). The initial states have fitness 1 and the final states fitness $r \geqslant 1$. All intermediate states are assumed to have the same fitness $s < r$. States are labeled by bit-strings, 0 is an wild-type site and 1 is a mutated site.

backward mutations. This is due to the fact that for sufficiently high mutation rates, fixation in the final state might never occur. Other definitions of the end state of the system become arbitrary to a certain extend. The probability $u_{m \to m+k}$ that the offspring of an individual with m mutations has $m + k$ mutations ($m \leq m + k \leq d$) is

$$u_{m \to m+k} = \binom{d-m}{k} \mu^k (1 - \mu)^{d-m-k}. \tag{1}$$

This equation is valid for the hypercube, where the order of mutations does not matter. Here, $\binom{d-m}{k}$ is the number of different types of mutants with k additional mutations, μ^k is the probability that mutations occur at k sites and $(1 - \mu)^{d-m-k}$ is the probability that no mutation occurs at the remaining $d - m - k$ sites. For the single path, there is only one possibility to arrange the $m + k$ mutations. Thus, for $k > 0$, $u_{m \to m+k}$ is identical to Eq. (1), except that the binomial factor has to be dropped. The probability $u_{m \to m}$ that no mutation occurs follows from normalization, $u_{m \to m} = 1 - \sum_{k=1}^{d-m} u_{m \to m+k}$. Our analytical calculations for small mutation rates as well as the considerations for high mutation rates are independent of the precise form of the mutation rates. However, we need to specify the form of the mutation probabilities to perform our numerical simulations for intermediate and high mutation rates.

3. Small mutation rates

3.1. The pace of evolution for small mutation rates

For small mutation probabilities, double mutations can be neglected. Since mutations occur rarely, we can calculate the average time until d mutations are fixed in the population analytically. Let us first address the evolutionary dynamics when mutants with fitness r_m are already present in a resident population of fitness r_w, but no new mutations occur. This scenario is relevant when mutation rates are sufficiently small. The probability to increase the number of mutants from j to $j + 1$ is

$$T_j^+ = \frac{r_m j}{r_m j + r_w (N - j)} \frac{N - j}{N}. \tag{2}$$

Similarly, the number of mutants decreases from j to $j - 1$ with probability

$$T_j^- = \frac{r_w (N - j)}{r_m j + r_w (N - j)} \frac{j}{N}. \tag{3}$$

The probability that k mutants take over the entire population is given by Nowak (2006), Karlin and Taylor (1975), Ewens (2004) and Crow and Kimura (1970)

$$\phi_k \left(\frac{r_m}{r_w} \right) = \frac{1 + \sum_{i=1}^{k-1} \prod_{j=1}^{i} \frac{T_j^-}{T_j^+}}{1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} \frac{T_j^-}{T_j^+}} = \frac{1 - \left(\frac{r_w}{r_m} \right)^k}{1 - \left(\frac{r_w}{r_m} \right)^N}. \tag{4}$$

If a mutant reaches fixation, the average number of generations this process takes is given by Goel and Richter-Dyn (1974) and Antal and Scheuring (2006)

$$\tau_{\text{fix}} \left(\frac{r_m}{r_w} \right) = \frac{1}{N} \sum_{k=1}^{N-1} \sum_{l=1}^{k} \frac{\phi_l}{T_l^+} \prod_{m=l+1}^{k} \frac{T_m^-}{T_m^+}. \tag{5}$$

For a neutral process with $r_m = r_w$, this reduces to $\tau_{\text{fix}} = N - 1$. For sufficiently large N, this is the maximum conditional fixation time of a mutant. Even for disadvantageous mutants ($r_m < r_w$) the conditional fixation time is smaller than $N - 1$ (Antal and

Scheuring, 2006). Since there are μN mutations per generation, the time between two mutations is $1/\mu N$. Thus, for $\mu \ll N^{-2}$ a mutant reaches fixation before the next one arises and mutations will not occur when a mutant is already present. Thus the population evolves by a process where the mutations occur one after the other, which has been termed periodic selection (Atwood et al., 1951) and theoretically described as the strong-selection weak-mutation regime (Gillespie, 1983, 2004).

The total time τ until a mutation reaches fixation in a population is the sum of the waiting time until a successful mutant occurs and the fixation time of the mutant $\tau = \tau_{\text{wait}} + \tau_{\text{fix}}$. The waiting time is the inverse of the mutation rate divided by the probability that a particular mutant is successful

$$\tau_{\text{wait}} \left(\frac{r_m}{r_w} \right) = \frac{1}{\mu N} \frac{1}{\phi_1 \left(\frac{r_m}{r_w} \right)}. \tag{6}$$

For $\mu \to 0$, we have $\tau_{\text{wait}} \to \infty$, but τ_{fix} remains approximately constant. Thus, $\tau \approx \tau_{\text{wait}}$ for small mutation rates. In principle, we could calculate τ_{fix} in the presence of mutations. But since our approximation is only valid for small mutation rates, this will be a minor correction.

For $\mu \ll N^{-2}$, the population is homogeneous most of the time. Only occasionally, a mutant arises and reaches fixation or goes to extinction. The total time until d mutations are fixed in the population is the sum of the waiting times for the successful mutants plus the time of the d fixation events. For a single path with initial fitness 1, intermediate fitness s and final fitness r, we find for the total time τ^S

$$\tau^S = \tau_{\text{wait}}(s) + (d - 2)\tau_{\text{wait}}(1) + \tau_{\text{wait}}(r/s) \\ + \tau_{\text{fix}}(s) + (d - 2)\tau_{\text{fix}}(1) + \tau_{\text{fix}}(r/s). \tag{7}$$

For small μ, we have $\tau_{\text{fix}} \ll \tau_{\text{wait}}$ and hence the total time can be approximated by

$$\tau^S = \frac{1}{\mu} \left[\frac{1}{N\phi_1(s)} + d - 2 + \frac{1}{N\phi_1(r/s)} \right]. \tag{8}$$

Consider now a "fitness valley", in which the intermediate states have fitness $s < 1$, but the final state has fitness $r > 1$. To move from the fitness peak in the initial state to the fitness state in the final state, first a disadvantageous mutation has to be fixed in the population. Since $\phi_1(s < 1) \ll \frac{1}{N}$, the waiting time of such an event is very long. The waiting time for the neutral mutations, $\tau_{\text{wait}}(1) = 1/\mu$ and the waiting time for a successful mutation, $\tau_{\text{wait}}(r/s)$ are significantly shorter. Thus, τ^S is dominated by $1/\mu N\phi_1(s)$ for $s < 1$ and sufficiently high N in a fitness valley. Fig. 2 shows a good agreement between exact numerical simulations and our analytical approximation for small mutation rates Eq. (8).

If the order of mutations is arbitrary, evolutionary dynamics occurs on a hypercube. In this case, the whole process will be faster, as we have $d!$ possible paths instead of a single one. Now, the waiting times in the different states depend on the number of mutations that are still available. For the total time τ^H, we obtain,

$$\tau^H = \frac{1}{d} \tau_{\text{wait}}(s) + \sum_{k=1}^{d-2} \frac{1}{d-k} \tau_{\text{wait}}(1) + \tau_{\text{wait}} \left(\frac{r}{s} \right) \\ + \tau_{\text{fix}}(s) + (d - 2)\tau_{\text{fix}}(1) + \tau_{\text{fix}} \left(\frac{r}{s} \right). \tag{9}$$

Note that the time of the fixations alone is identical for the hypercube and the single path. Neglecting these fixation times for small μ (as $\tau_{\text{fix}} \ll \tau_{\text{wait}}$) yields

$$\tau^H = \frac{1}{\mu} \left[\frac{1}{dN\phi_1(s)} + \sum_{k=1}^{d-2} \frac{1}{d-k} + \frac{1}{N\phi_1(\frac{r}{s})} \right]. \tag{10}$$
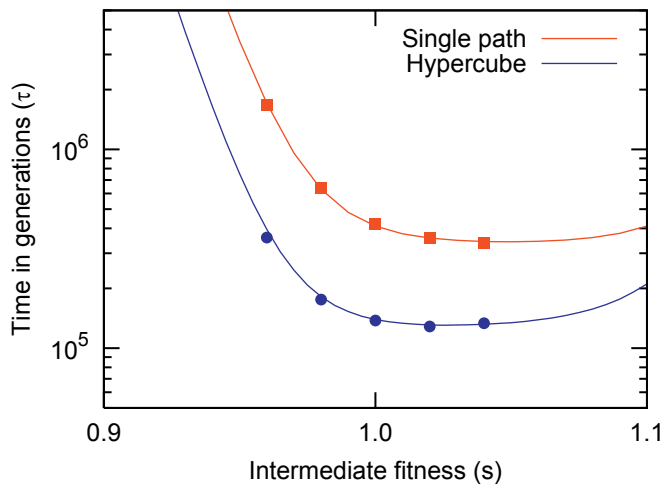
Fig. 2. Fixation time for a single path (squares) and a hypercube (circles) with small mutation rates ($\mu \ll N^{-2}$) for different intermediate fitness values. Evolutionary dynamics is always faster in the hypercube. Solid lines show the analytical approximation for small mutation rates, Eqs. (8) and (10). Numerical simulations shown by symbols agree well with the analytical approximation (population size $N = 100$, mutation rate $\mu = 10^{-5}$, $d = 5$, $r = 1.1$, simulations averaged over a 1000 realizations).

Since $1/(dN\phi_1(s)) < 1/(N\phi_1(s))$ and $\sum_{k=1}^{d-2} 1/(d-k) < \sum_{k=1}^{d-2} 1 = d - 2$ it is obvious that $\tau^H < \tau^S$, i.e. evolutionary dynamics is faster if the order of mutations is arbitrary. For fitness valleys with $s < 1$ and a large population size, τ^H is dominated by $1/d\mu N\phi_1(s)$. As d more mutations are available, this is always faster than the corresponding equation for a single path, see Fig. 2.

3.2. Thresholds of the waiting times

Next, we derive expressions for some interesting thresholds of the waiting times in the limit of small mutation rates. Since evolutionary dynamics is always faster if many paths are available, we now compare a fitness valley in which many paths are available to a single path in which the order of mutations is important, but fitness does not decrease in the course of evolution. The basic question we address here is, whether it is faster to cross a broad valley or a narrow ridge in fitness space. In other words, we compare $\tau^S(s=1)$ to $\tau^H(s<1)$. Since we consider only small mutation rates μ, we neglect the fixation times τ_{fix} here, although they will not be identical in the two scenarios. For $s = 1$, the single path is neutral. We decrease s in the hypercube until we have identical waiting times. This yields an implicit expression for s

$$d - 1 + \frac{1}{N}\frac{1 - \frac{1}{r^N}}{1 - \frac{1}{r}} = \frac{1}{dN}\frac{1 - \frac{1}{s^N}}{1 - \frac{1}{s}} + \sum_{k=1}^{d-2}\frac{1}{d-k} + \frac{1}{N}\frac{1 - \left(\frac{s}{r}\right)^N}{1 - \frac{s}{r}}. \quad (11)$$

From this equation, we can numerically determine s for any given N. For large N, Eq. (11) simplifies to

$$d(d-1) - \sum_{k=1}^{d-2}\frac{d}{d-k} = \frac{1}{N}\frac{1 - \frac{1}{s^N}}{1 - \frac{1}{s}} \approx \frac{e^{N(1-s)-1}}{N(1-s)}, \quad (12)$$

where we used $(1 - x/N)^{-N} \to e^x$ for large N. Thus, the quantity $N(1 - s)$ becomes constant for large N, see Fig. 3. Thus, we can now say how broad and deep a fitness valley has to be to lead to the same cumulative waiting time as a single neutral path.

Next, we address the effect of the intermediate fitness s, which has an important influence on the cumulative waiting
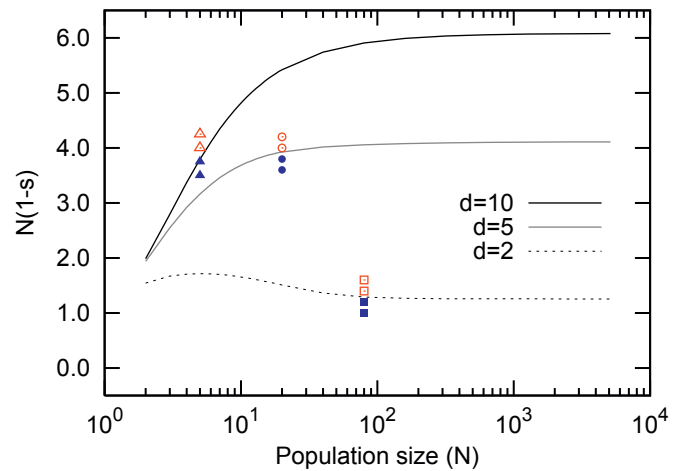

Fig. 3. The figure shows the threshold values for which evolution on a hypercube with fitness $s < 1$ in the intermediate states proceeds as fast as on a single neutral path with $s = 1$. Full lines show $N(1 - s)$ based on the numerical solution of Eq. (11). Above the lines, evolution proceeds faster on the hypercube. Below them, the neutral single path is faster. For $N \to \infty$, the lines converge to a constant, see Eq. (12). The symbols show the results from numerical simulations for $N = 5$ and $d = 10$ (triangles), $N = 20$ and $d = 5$ (circles) as well as $N = 80$ and $d = 2$ (squares). Symbols are open when the single path is faster and filled when the hypercube is faster ($\mu = 10^{-5}$, $r = 1.1$, simulations averaged over a 1000 realizations).

time τ. Fig. 2 shows how the waiting time decreases with increasing fitness in the intermediate states s. If s comes very close to the fitness in the final state, the waiting time increases. This increase is seen both in the single path and the hypercube. An increase in the intermediate state fitness will not always lead to a reduction in waiting times. Instead, the fixation times reach a minimum when the fitness growth is constant between any two consecutive states (Weinreich and Chao, 2005; Traulsen et al., 2007). For the hypercube, the fastest trajectory will be steeper than on a single path: at first, many mutations are available and a big fitness increase is not necessary. Later, fewer mutations are available and thus, the fitness should increase faster. The precise form of the trajectory will in this case depend on the number of mutations d and the population size N. We note that a similar reasoning can be applied to construct a fitness landscape that allows to cross a fitness valley fastest. The fastest trajectory has the same form regardless if a fitness peak is approached ($r > 1$) or a fitness minimum is approached ($r < 1$). Thus, the fastest way to cross a fitness valley is to descend to the minimum with exponentially decreasing fitness and to increase from the minimum again with exponentially increasing fitness.

Now, we turn to the effect of the intermediate fitness s on the individual waiting times. Eqs. (8) and (10) both consist of three terms each. The first term denotes the time required to leave the initial state. The second term is the time spent in moving through all the intermediate states. This second term is independent of s, because the transitions are neutral. The last term denotes the time required to reach the ultimate state from the penultimate state. For small values of s, the probability to fixate the disadvantageous mutation is very small. Thus, the total time is dominated by the first term. When s is increased to a threshold value s_1, then the time for leaving the first state is identical to the waiting time in the intermediate states. For the hypercube, s_1 is given by $(1/d)\tau_{\text{wait}}(s_1) = \sum_{k=1}^{d-2}(1/(d-k))\tau_{\text{wait}}(1)$, which reduces to

$$\frac{1 - \frac{1}{s_1^N}}{1 - \frac{1}{s_1}} = dN\sum_{k=1}^{d-2}\frac{1}{d-k}. \quad (13)$$

This equation can be solved numerically for specific values of N and d. For the single path, the right hand side of this equation has to be replaced by $N(d-2)$. For $s > s_1$, the time to cross the intermediate states is larger than the waiting time in the first state. On the hypercube, we can define a second threshold for which the waiting time in the first state is the same as the time required to reach the final state from the penultimate state. This arises because the effective mutation rate in state 0 is d times larger than the effective mutation rate in state $d-1$. The threshold s_2 is given by $(1/d)\tau_{\text{wait}}(s_2) = \tau_{\text{wait}}(r/s_2)$ or

$$\frac{1}{d}\frac{1 - \frac{1}{s_2^N}}{1 - \frac{1}{s_2^1}} = \frac{1 - \left(\frac{s_2}{r}\right)^N}{1 - \frac{s_2}{r}}. \tag{14}$$

Again, s_2 has to be determined numerically. For a single path, the factor d^{-1} in Eq. (14) has to be dropped. Thus, the threshold s_2 occurs for $s > 1$ and is simply given by $s_2 = \sqrt{r}$.

The fixation time is also strongly influenced by the number of mutations d. A larger d increases the length of the path and usually also the fixation times. For the single path, this increase results only from the increase in the time required to cross the intermediate states, because the time for leaving the initial state and the time to reach the final state from the penultimate state are independent of d. The time required to reach the ultimate state from the penultimate state is also independent of d for the hypercube, but the time required to leave the initial state decreases with increasing d. This is because as d increases, there are more states available in the first error class and thus the effective rate of mutation out of the initial state increases. As for the single path, the time to cross the intermediate states increases with d in the hypercube. For the hypercube, this interplay can lead to a non-monotonic dependence of the fixation time on d. For example, for $N = 100$ and $s = 0.95$, the fixation time τ^H decreases with d for $d < 31$, but it increases with d for $d > 31$. In contrast, the fixation time always increases monotonically with d for the single path.

Increasing the fitness of the final state r increases the advantage of the final state over the intermediate states. This will result in the decrease in the time required for the population to make the last move. Increasing r has no effect on the time required to cross the intermediate states or the time required to move away from the initial state. As a result, those two times remain constant even as r increases, both in the single path and the hypercube.

4. Intermediate mutation rates

The analytical approach is only valid as long as the mutation rate is small, $\mu \ll N^{-2}$. For higher mutation rates, the population does not have to consist of at most two different types at any time. Instead, d mutations can be fixed in the population without sequentially fixing one after the other. This process has been termed stochastic tunneling and is of great importance in the context of cancer initiation (Iwasa et al., 2004; Michor et al., 2004; Nowak et al., 2004; Beerenwinkel et al., 2007). Tunneling across fitness valleys is more likely than tunneling across a flat fitness landscape (see Fig. 4). Even for $d = 2$, the evolutionary dynamics is characterized by a doubly stochastic process, which makes analytical approaches tedious (Iwasa et al., 2004). As discussed above, for $\mu N^2 \ll 1$ the population usually contains at most two different types. In this case, the probability of stochastic tunneling will be very small. On the other hand, for $\mu N > 1$, at least one mutant is produced per generation. Thus, the probability of stochastic tunneling approaches 1. For $N^{-2} < \mu < N^{-1}$, the
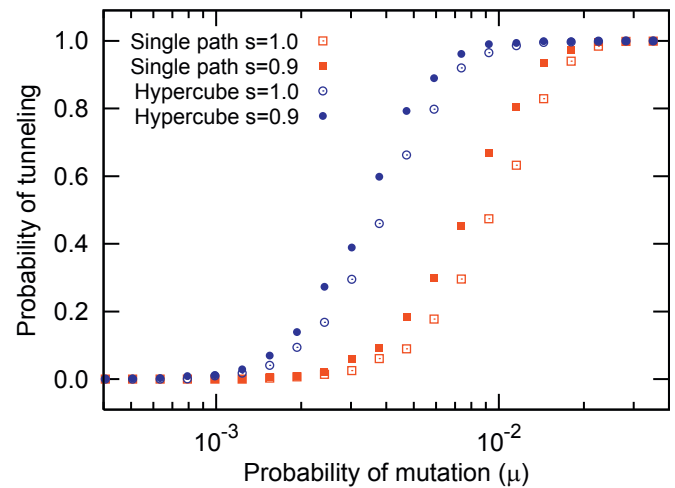

Fig. 4. The probability of tunneling across the hypercube (circles) is larger than in the single path (squares) due to the higher effective mutation rate. The tunneling across the valley denoted by the filled symbols ($s = 0.9$) is always larger than the probability of tunneling across a flat fitness landscape denoted by open symbols ($s = 1.0$). This arises from the fact that the stepwise accumulation of mutations would involve the fixation of a disadvantageous mutation in the first step. All symbols show the probabilities that the population tunnels at least across one state for the single path or at least one error class for the hypercube ($N = 100$, $d = 5$, $r = 1.1$, averaged over a 1000 realizations).
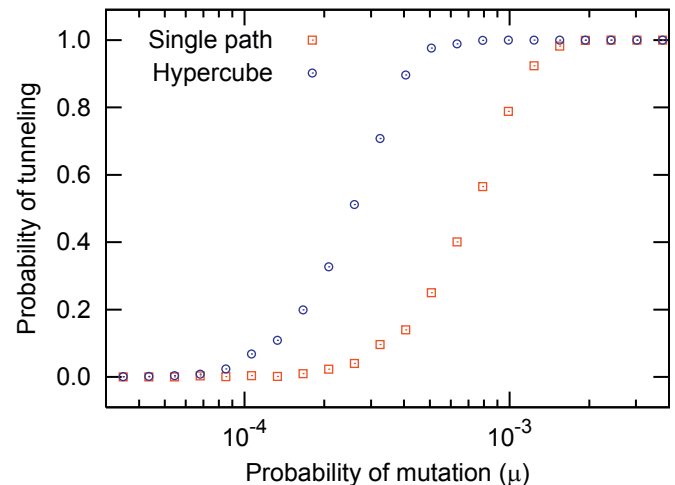

Fig. 5. The probability of tunneling across a neutral hypercube (circles) is always higher than the probability of tunneling across a neutral single path (squares). Here, the probability that the system tunnels across at least one state or one error class is shown for $d = 5$ and $N = 1000$. As expected, for $N\mu > 1$ the probability of tunneling approaches 1. In contrast to our conservative estimate that tunneling can be neglected only as long as the mutation rate is below N^{-2}, even for mutation rates as large as $100N^{-2}$ the probability of tunneling remains close to zero ($s = 1.0$, $r = 1.1$, averaged over a 1000 realizations).

mutations are sometimes fixed sequentially and sometimes via stochastic tunneling. Fig. 5 shows how the tunneling probability increases from 0 to 1 in this interval.

For intermediate mutation rates, it is likely that the population contains more than two different types. The types with beneficial mutations will compete for fixation. This process is termed clonal interference (Crow and Kimura, 1970; Fisher, 1930; Muller, 1932; Gerrish and Lenski, 1998; Park and Krug, 2007). Clonal interference has been considered to slow down adaptation, but recently it has been shown that it can have a positive influence on a rugged fitness landscape (Gerrish and Lenski, 1998; Wilke, 2004; Jain and Krug, 2007).

The states in a single path can be characterized by the number of mutations. In the hypercube, the states are characterized by the types of mutations that have already occurred. Thus, there are many different types that have undergone a specific number of mutations. However, all types that have already accumulated k mutations can be pooled into the error class k. The number of different types in the error class k is given by $\binom{d}{k} = d!/k!(d-k)!$.

In a single path, a population is said to tunnel across a state if it passes through a state without ever reaching fixation in that state. Analogously, in a hypercube a population said to tunnel across an error class if it passes through that error class without ever reaching fixation in it. Within the error classes, tunneling can occur across individual states, but also across several states at once. This means that the whole population passes only across that particular state and not across any other, without ever reaching fixation in that particular state. Tunneling across an error class can also occur in a second way: the population can use all of the available states in the error class, but the total number of individuals in the error class never reaches N. Thus, the probability of tunneling via the individual states is always lower than the probability of tunneling across the error classes. Fig. 6 shows the relation between the different probabilities of tunneling in the hypercube with respect to the rate of mutation μ for the special case $d = 2$.

Due to higher effective rates of mutation, the probability of tunneling across a hypercube is expected to be greater than or equal to the probability of tunneling across a single path. However, numerical simulations reveal that for $d = 2$ the probability of tunneling in a single path is higher than in the hypercube. This is a special case: for $d = 2$ in a hypercube, the number of states into which the initial state can mutate into is 2. The effective rate of mutations is thus twice as much as in the single path. The number of states which can be mutated into next is one, both in the single path and the hypercube. Thus the rate at which the individuals are pushed into the first state is higher in hypercube than in the single path while the rate of individuals being pushed out is the same. Thus there is a higher probability of reaching fixation in the first error class in a hypercube (see Fig. 6). We only observe this effect for $d = 2$, for $d > 2$, the probability of

tunneling is higher in a hypercube than in a single path, as expected (see Fig. 4).

5. High mutation rates

For $\mu N > 1$, the stochastic features of the dynamics become less important. In this case, the system can be described by a set of $d + 1$ deterministic differential equations for the fraction $x_k(t)$ of the population that has k mutations (Jain and Krug, 2007). Obviously, we have $\sum_{k=0}^{d} x_k(t) = 1$. Transitions out of state 0 occur with probability $T_{0\to} = (1 - (x_0/\phi)u_{0\to0})x_0$, where $\phi = x_0 + (1 - x_0 - x_d)s + x_d r$ is the average fitness of the population. This includes all the reproductive events except for the one where a type 0 is produced. Transitions into state 0 occur with probability $T_{\to0} = (x_0/\phi)u_{0\to0}(1 - x_0)$. Thus, the fraction of individuals in the initial state follows the differential equation:

$$\dot{x}_0(t) = \frac{1}{N}\left[\frac{x_0}{\phi}u_{0\to0}(1 - x_0) - \left(1 - \frac{x_0}{\phi}u_{0\to0}\right)x_0\right]. \tag{15}$$

The probability that an offspring is of type k is given by $\lambda_k = \sum_{j=0}^{k}(x_j r_j/\phi)u_{j\to k}$. The difference between the hypercube and the single path only occurs in the quantity $u_{j\to k}$, which is given above for both cases. The sum in λ_k is over all individuals with k or less mutations and r_j is the fitness of individuals with j mutations. This leads to the differential equation for the fraction of individuals with k mutations

$$\dot{x}_k(t) = \frac{1}{N}[\lambda_k(1 - x_k) - (1 - \lambda_k)x_k], \tag{16}$$

where $k = 0, \ldots, d$. Of course, the special case $k = 0$ recovers Eq. (15). This set of $d + 1$ differential equations describes how the system moves from state $k = 0$ to state $k = d$. In general, only a numerical solution of this system of equations is feasible. While this allows us to infer details of the dynamics, our main interest is the time required for fixation of d number of mutations. Thus, we solve the differential equation numerically using a standard Runge–Kutta algorithm (Press et al., 2007). To find an equivalent to the fixation time in a stochastic simulation, we average between fixation ($x_d = 1$) and the time when there are on average less than 1 individuals outside the final state ($x_d = 1 - 1/N$). Thus, the fixation time is the time when the solution of the differential equation crosses $x_d = 1 - 1/2N$.

Fig. 7 shows an overview of the fixation times, covering the full range of mutation rates. For small mutation rates, we have sequential fixation of mutations and the time can be well approximated by Eqs. (8) and (10). For high mutation rates, the numerical solution of Eq. (16) leads to a good approximation for the fixation times.

6. Discussion

We have determined the average time during which a population moves from a certain initial state to a final state of higher fitness. The initial and the final states are separated by a fixed number of mutations d. The mutations jointly confer a fitness advantage to the final mutant which can be represented by a peak in the fitness landscape. If the intermediate mutations need to occur in a specific order for the evolution of the final mutant then it corresponds to the single path. Otherwise, evolution occurs on a hypercube and there are $d!$ ways of reaching the final state.

We have explored the simplest system in which the fitness in all intermediate states is the same. As expected, the fixation times on a hypercube are shorter than on a single path, due to the
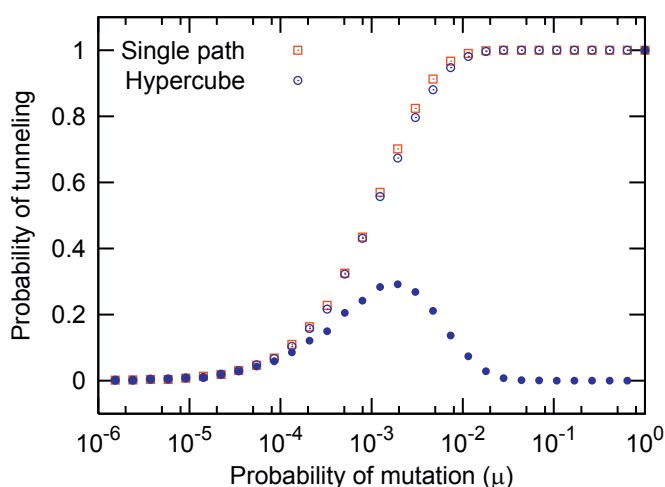

Fig. 6. For $d = 2$, the probability of tunneling across the intermediate state is slightly higher in the single path (squares) than in the hypercube (open circles), shown for $N = 100$ here. This is because the effective mutation rate into the intermediate state is twice as big in the hypercube, leading to a higher probability of fixation. Filled circles show the probability to tunnel across individual states of the hypercube. For $\mu N > 1$, the system always tunnels. In the hypercube, both states are used for this. As expected, we need $\mu \ll N^{-2}$ for the tunneling probability to vanish ($s = 1.0$, $r = 1.1$, averaged over a 1000 realizations).
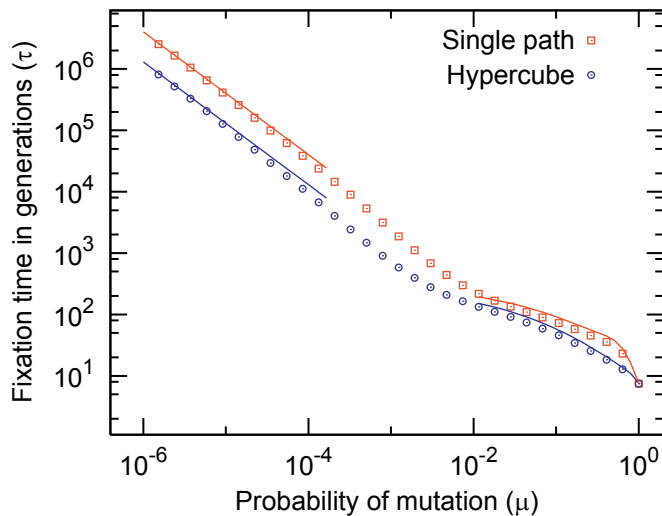
Fig. 7. The fixation times decrease with increasing mutation rate. Fixation always occurs faster on the hypercube (circles) than in the single path (squares). For small mutation rates, mutations fixate sequentially and the fixation time can be well approximated by Eqs. (8) and (10). Here, the fixation times decrease as μ^{-1}. For high mutation rates, the system can be approximated by a set of deterministic differential equations and the simulation results for the fixation times can be approximated based on the numerical solution of Eq. (16). In this case, fixation times decrease in general slower than μ^{-1} with increasing mutation rate (population size $N = 1000$, $d = 5$, $s = 1$, $r = 1.1$, averages over 1000 realizations).

Table 1

The time required for fixation of d mutations in units of 10^{10} generations for a mutation rate of $\mu = 10^{-10}$ based on Eqs. (8) and (10).

N	$d = 3$		$d = 10$	
	Single path	Hypercube	Single path	Hypercube
10^2	2.10999	0.943325	9.10999	2.03896
10^4	2.0011	0.834433	9.0011	1.93007
10^6	2.00001	0.833344	9.00001	1.92898

The intermediate mutations are neutral, $s = 1$. For small mutation rates, the fixation times scale linearly with μ^{-1}. For $N \to \infty$, the fixation time on the single path approaches $\mu^{-1}(d - 1)$ and the fixation time on the hypercube approaches $\mu^{-1} \sum_{k=0}^{d-2}(d - k)$. However, the mutation rates have to decrease with increasing N to make the approximation for the fixation times valid (initial fitness 1.0 and final fitness $r = 1.1$).

presence of multiple paths available in a hypercube. This observation leads to the question: for which parameters does the hypercube show shorter fixation times than the single path, even with an added disadvantage? The fitness in the intermediate states was then set to lower values than the ones in the single path. Up to a certain threshold value of the fitness of the intermediate states, the hypercube shows shorter fixation times than in the single path. The value of the threshold depends on the population size, total number of required mutations and the fitness in the final state.

The fixation times for large populations largely depend on the fitness function and are qualitatively independent of the order of mutations. Let us first focus on a flat landscape: when the intermediate states have a fitness equal to the fitness of the initial wild-type, then for small mutation rates large populations have shorter fixation times than small populations. This is because the neutral rate of evolution does not depend on the population size. But the waiting time for fixation of the last mutation becomes shorter with larger population size. For intermediate mutation rates, tunneling starts earlier in larger populations. This leads to a marked decrease in the fixation time with larger population size. For high mutation rates, the time to fixation is no longer dominated by the time for the first mutant to reach the final state, but by the time until all individuals are in that state. Due to this, for high mutation rates the time required for fixation can be shorter in smaller population as compared to larger populations. Next, we focus on fitness valley: if the fitness landscape consists of a valley with reduced fitness of the intermediate states, small populations have an advantage for small mutation rates, as they can easily leave the initial state and enter the valley. But for high mutation rates, large populations reach fixation faster, because they can explore states within and beyond the fitness valley more easily.

Our numerical simulations reveal that tunneling can be neglected even when the mutation rate exceeds N^{-2}, at least by one order of magnitude. Thus, Eqs. (8) and (10) provide good estimates for the fixation times even in relatively large populations.

Concrete values for fixation times are collected in Table 1. They reveal that even in long-term studies of experimental evolution, it is difficult to observe the consecutive fixation of neutral mutants (Cooper et al., 2003). Consecutive fixation of advantageous mutants, however, is significantly faster. For example, while Table 1 reveals a fixation time of $\sim 10^{11}$ generations on a single path for $d = 10$, $s = 1$ and $N = 10^6$, an optimal choice of the intermediate fitness values (Traulsen et al., 2007) would lead to a fixation time of $\sim 10^7$ generations.

While we have focused on the simplest possible system which allows analytical approximations, experimental studies reveal of course a much higher complexity. Weinreich et al. (2006) studied experimentally the point mutations in the β-lactamase gene of bacteria. β lactam antibiotics are commonly used, but the bacteria can develop resistance to the drugs. Five point mutations in a particular allele of the β-lactamase gene increases the resistance of the bacteria to cefotaxime by a factor of $\sim 100,000$. Theoretically the mutations leading from the wild-type allele to the resistant allele can occur in $5! = 120$ ways. These can be represented by a hypercube of $d = 5$. But in only 18 of the 120 trajectories, the intermediate mutations are either neutral with respect to the initial state or beneficial. Weinreich and colleagues have shown that these have the highest probability of realization. For all beneficial intermediates the fastest way to reach the final state would be when the relative fitness increase between any two consecutive mutations is constant (Traulsen et al., 2007), but usually in nature several different mutations are available and the population first evolves to states that provide the highest selective advantage.

In another experimental study the sequence space of the 5s rRNA of a marine bacterium, *Vibrio proteolyticus* was explored (Lee et al., 1997). The sequences from *Vibrio proteolyticus* and *Vibrio alginolyticus* differ in only four positions. All the possible intermediates were constructed by the authors and the fitness of each was calculated (Chao and McBroom, 1985). Two of the valid intermediates have a fitness lower than the initial wild-type. We have shown how such fitness valleys can be crossed by exploring the phenomenon of tunneling or multiple mutations (for high mutation rates). Thus, the population does not need to move in a Wrightian fashion (the whole population moving as a whole across the valley).

The theory discussed herein deals with basic evolutionary concepts which are important to the kind of biological examples described above. More complex properties of the experimental studies like more general cases of epistasis and compensatory mutations can easily be incorporated, but there is a huge number of possibilities. Even if we are only interested in the ordering of fitness values, we can have up to $2^d!$ distinct epistatic patterns.

Thus, one should rather focus on concrete systems instead. For example, one could simulate the dynamics in a system with experimentally derived fitness values and mutation rates. Not all the paths of a hypercube might be accessible for selection, but still some of them might prove to be significant depending upon the particular values of the parameters, such as fitness values and population size. Our goal here was to characterize the simplest features of the dynamics of a population crossing a fitness valley. This approach can be helpful when more realistic scenarios are addressed.

Acknowledgments

References

Antal, T., Scheuring, I., 2006. Fixation of strategies for an evolutionary game in finite populations. Bull. Math. Biol. 68, 1923–1944.

Armitage, P., Doll, R., 1954. The age distribution of cancer and a multi-stage theory of carcinogenesis. Br. J. Cancer 8, 1–12.

Atwood, K.C., Schneider, K.L., Ryan, F.J., 1951. Periodic selection in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 37, 146–155.

Beerenwinkel, N., Antal, T., Dingli, D., Traulsen, A., Kinzler, K.W., Velculescu, V.E., Vogelstein, B., Nowak, M.A., 2007. Genetic progression and the waiting time to cancer. PLoS Comput. Biol. 3, e225.

Chao, L., McBroom, S.M., 1985. Evolution of transposable elements: an IS10 insertion increases fitness in Escherichia coli. Mol. Biol. Evol. 2, 359–369.

Cooper, T., Rozen, D., Lenski, R.E., 2003. Parallel changes in gene expression after 20,000 generations of evolution in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1072–1077.

Crow, J.F., Kimura, M., 1970. An Introduction to Population Genetics Theory. Harper & Row, New York.

DePristo, M.A., Hartl, D.L., Weinreich, D.M., 2007. Mutational reversions during adaptive protein evolution. Mol. Biol. Evol. 24, 1608–1610.

Eigen, M., McCaskill, J., Schuster, P., 1989. The molecular quasi-species. Adv. Chem. Phys. 75, 149–263.

Eigen, M., Schuster, P., 1977. The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. Die Naturwiss. 64, 541–565.

Ewens, W.J., 2004. Mathematical Population Genetics. Springer, New York.

Fisher, J., 1959. Multiple-mutation theory of carcinogenesis. Nature 181, 651–652.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Frank, S.A., 2007. Evolutionary Dynamics of Cancer. Princeton University Press, Princeton, NJ.

Gerrish, P.J., Lenski, R.E., 1998. The fate of competing beneficial mutations in a asexual population. Genetica 102/103, 127–144.

Gerstung, M., Beerenwinkel, N., 2008. Waiting time models of cancer progression. Mathematical Population Studies, to appear. arXiv:0807.3638.

Gillespie, J.H., 1983. Some properties of finite populations experiencing strong selection and weak mutation. Am Nat. 121 (5), 691–708.

Gillespie, J.H., 2004. Population Genetics: A Concise Guide, second ed. The Johns Hopkins University Press.

Goel, N., Richter-Dyn, N., 1974. Stochastic Models in Biology. Academic Press, New York.

Iwasa, Y., Michor, F., Nowak, M.A., 2004. Stochastic tunnels in evolutionary dynamics. Genetics 166, 1571–1579.

Jain, K., Krug, J., 2007. Deterministic and stochastic regimes of asexual evolution on rugged fitness landscapes. Genetics 175, 1275–1288.

Jones, S., Chen, W.-D., Parmigiani, G., Diehl, F., Beerenwinkel, N., Antal, T., Traulsen, A., Nowak, M.A., Siegel, C., Velculescu, V.E., Kinzler, K.W., Vogelstein, B., Willis, J., Markowitz, S.D., 2008a. Comparative lesion sequencing provides insights into tumor evolution. Proc. Natl. Acad. Sci. U.S.A. 105 (11), 4283–4288.

Jones, S., Zhang, X., Parsons, D.W., Lin, J.C.-H., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Kamiyama, H., Jimeno, A., Hong, S.-M., Fu, B., Lin, M.-T., Calhoun, E.S., Kamiyama, M., Walter, K., Nikolskaya, T., Nikolsky, Y., Hartigan, J., Smith, D.R., Hidalgo, M., Leach, S.D., Klein, A.P., Jaffee, E.M., Goggins, M., Maitra, A., Iacobuzio-Donahue, C., Eshleman, J.R., Kern, S.E., Hruban, R.H., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V.E., Kinzler, K.W., 2008b. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science 321 (5897), 1801–1806.

Karlin, S., Taylor, H.M.A., 1975. A First Course in Stochastic Processes, second ed. Academic Press, London.

Knudson, A.G., 1971. Mutation and cancer: statistical study of retinoblastoma. Proc. Natl. Acad. Sci. U.S.A. 68, 820–823.

Lee, T.H., DSouza, L.M., Fox, G.E., 1997. Equally parsimonious pathways through an RNA sequence space are not equally likely. J. Mol. Evol. 45, 278–284.

Michor, F., Iwasa, Y., Nowak, M.A., 2004. Dynamics of cancer progression. Nat. Rev. Cancer 4, 197–205.

Moran, P.A.P., 1962. The Statistical Processes of Evolutionary Theory. Clarendon Press, Oxford.

Muller, H.J., 1932. Some genetic aspects of sex. Am. Nat. 66 (703), 118–138.

Nordling, C., 1953. A new theory on cancer inducing mechanism. Br. J. Cancer 7, 68–72.

Nowak, M.A., 1992. What is a quasispecies? Trends in Ecology and Evolution 7, 118–121.

Nowak, M.A., 2006. Evolutionary Dynamics. Harvard University Press, Cambridge, MA.

Nowak, M.A., Michor, F., Komarova, N.L., Iwasa, Y., 2004. Evolutionary dynamics of tumor suppressor gene inactivation. Proc. Natl. Acad. Sci. U.S.A. 101, 10635–10638.

Park, S.-C., Krug, J., 2007. Clonal interference in large populations. Proc. Natl. Acad. Sci. U.S.A. 104 (46), 18135–18140.

Poelwijk, F.J., Kiviet, D.J., Weinreich, D.M., Tans, S.J., 2007. Empirical fitness landscales reveal accessible evolutionary paths. Nature 445, 383–386.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press, New York.

Sjöblom, T., Jones, S., Wood, L., Parsons, D., Lin, J., Barber, T., Mandelker, D., Leary, R., Ptak, J., Silliman, N., Szabo, S., Buckhaults, P., Farrell, C., Meeh, P., Markowitz, S., Willis, J., Dawson, D., Willson, J., Gazdar, A., Hartigan, J., Wu, L., Liu, C., Parmigiani, G., Park, B., Bachman, K., Papadopoulos, N., Vogelstein, B., Kinzler, K., Velculescu, V., 2006. The consensus coding sequences of human breast and colorectal cancers. Science 314, 268–274.

Traulsen, A., Iwasa, Y., Nowak, M.A., 2007. The fastest evolutionary trajectory. J. Theor. Biol. 249 (3), 617–623.

van Nimwegen, E., Crutchfield, J., 2000. Metastable evolutionary dynamics: crossing fitness barriers or escaping via neutral paths? Bull. Math. Biol. 62, 799–848.

Vogelstein, B., Kinzler, K., 2004. Cancer genes and the pathways they control. Nat. Med. 10, 789–799.

Weinreich, D., 2005. The rank ordering of genotypic fitness values predicts genetic constraint on natural selection on landscapes lacking sign epistasis. Genetics 171, 1397–1405.

Weinreich, D., Delaney, N., DePristo, M., Hartl, D., 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 312, 111–114.

Weinreich, D.M., Chao, L., 2005. Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution 59, 1175–1182.

Weinreich, D.M., Watson, R., Chao, L., 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. Evolution 56 (6), 1165–1174.

Wilke, C., 2004. The speed of adaptation in large asexual populations. Genetics 167, 2045–2053.

Wilke, C., 2005. Quasispecies theory in the context of population genetics. BMC Evol. Biol. 5, 44.

Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J., Silliman, N., Szabo, S., Dezso, Z., Ustyankky, V., Nikolskaya, T., Nikolsky, Y., Karchin, R., Wilson, P.A., Kaminker, J.S., Zhang, Z., Croshaw, R., Willis, J., Dawson, D., Shipitsin, M., Willson, J.K.V., Sukumar, S., Polyak, K., Park, B.H., Pethiyagoda, C.L., Pant, P.V.K., Ballinger, D.G., Sparks, A.B., Hartigan, J., Smith, D.R., Suh, E., Papadopoulos, N., Buckhaults, P., Markowitz, S.D., Parmigiani, G., Kinzler, K.W., Velculescu, V.E., Vogelstein, B., 2007. The genomic landscapes of human breast and colorectal cancers. Science 318 (5853), 1108–1113.