Emergent Neutrality in Adaptive Asexual Evolution

Stephan Schiffels,*,1 Gergely J. Szöllősi,†,1 Ville Mustonen,‡ and Michael Lässig*,2
*Institut für Theoretische Physik, Universität zu Köln, 50937 Köln, Germany, †Centre National de la Recherche Scientifique, UMR 5558–Laboratoire de Biométrie et Biologie Évolutive, Université Claude Bernard, Lyon 1, France, and ‡Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, United Kingdom

ABSTRACT In nonrecombining genomes, genetic linkage can be an important evolutionary force. Linkage generates interference interactions, by which simultaneously occurring mutations affect each other's chance of fixation. Here, we develop a comprehensive model of adaptive evolution in linked genomes, which integrates interference interactions between multiple beneficial and deleterious mutations into a unified framework. By an approximate analytical solution, we predict the fixation rates of these mutations, as well as the probabilities of beneficial and deleterious alleles at fixed genomic sites. We find that interference interactions generate a regime of *emergent neutrality*: all genomic sites with selection coefficients smaller in magnitude than a characteristic threshold have nearly random fixed alleles, and both beneficial and deleterious mutations at these sites have nearly neutral fixation rates. We show that this dynamic limits not only the speed of adaptation, but also a population's degree of adaptation in its current environment. We apply the model to different scenarios: stationary adaptation in a time-dependent environment and approach to equilibrium in a fixed environment. In both cases, the analytical predictions are in good agreement with numerical simulations. Our results suggest that interference can severely compromise biological functions in an adapting population, which sets viability limits on adaptive evolution under linkage.

POPULATIONS adapt to new environments by fixation of beneficial mutations. In linked sequence, simultaneously occurring mutations interfere with each other's evolution and enhance or reduce each other's chance of fixation in the population. We refer to these two cases as positive and negative interference. Several classical studies have shown that interference can substantially reduce the speed of adaptation in large asexual populations (Fisher 1930; Muller 1932; Smith 1971; Felsenstein 1974; Barton 1995; Gerrish and Lenski 1998). Linkage effects are weaker in sexual populations, because they are counteracted by recombination.

Microbial evolution experiments provide a growing amount of data on adaptive evolution under linkage (de Visser *et al.* 1999; Rozen *et al.* 2002; de Visser and Rozen 2006; Desai and Fisher 2007; Perfeito *et al.* 2007; Silander *et al.* 2007; Kao and Sherlock 2008; Barrick *et al.* 2009; Betancourt 2009; Kinnersley *et al.* 2009), and similar data

are available for adaptive evolution in viral systems (Bush *et al.* 1999; Rambaut *et al.* 2008; Neher and Leitner 2010). Modern deep sequencing opens these systems to genomic analysis and poses new questions: How does a continuously changing environment such as the human immune challenge shape the genome of the seasonal influenza virus? How does the fitness of a bacterial population increase in a new environment? To answer such questions, we need to explain how a population and its current fitness values evolve in a time-dependent ecology and fitness landscape and what are the rates of beneficial and deleterious changes observed in this process. Thus, we need to describe the adaptive process in an explicitly genomic context.

In this article, we develop a genomic model of adaptation under linkage, which establishes the conceptual framework to analyze such data. Our model links the *adaptive process*, which changes the frequencies of beneficial and deleterious alleles at polymorphic sites, to the *genome state*, which includes the distribution of beneficial and deleterious alleles at fixed sites. It is the genome state that determines the fitness of a population in its current environment. We show that interference interactions can drastically affect process and state in large asexual populations: Adaptation generates beneficial *driver* mutations, but a substantial fraction of allele changes are *passenger* mutations, whose chance of fixation depends only weakly on their selection coefficient. Thus,

[1]These authors contributed equally to this work.
[2]Corresponding author: Institut für Theoretische Physik, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany. E-mail: mlaessig@uni-koeln.de

a near-neutral dynamic of mutations emerges from sufficiently strong interference interactions. This effect causes a—potentially large—fraction of genomic sites to have nearly random fixed alleles, which do not reflect the direction of selection at these sites. Thus, interference interactions not only reduce the speed of adaptation, but also degrade the genome state and the population's fitness in its current environment. The joint dynamics of driver and passenger mutations have consequences that may appear counterintuitive: deleterious substitutions of a given strength can have a rate increasing with population size, and beneficial substitutions can have a rate decreasing with population size. This behavior is contrary to unlinked sites evolving under genetic drift.

The complexity of genomic linkage and of interference interactions is reflected by the long history of the subject in population genetics literature, which dates back to Fisher and Muller (Fisher 1930; Muller 1932). Their key observation is that in the absence of recombination, two mutations can both reach fixation only if the second mutation occurs in an individual that already carries the first. In other words, mutations occurring in different individuals interfere with one another. Interference inevitably causes a fraction of all mutations to be lost, even if they are beneficial and have already reached substantial frequencies in the population (*i.e.*, have overcome genetic drift). Following a further seminal study, the interference between linked mutations is commonly referred to as the Hill–Robertson effect (Hill and Robertson 1966). This term is also used more broadly to describe the interplay between linkage and selection: interference interactions reduce the fixation probability of beneficial mutations and enhance that of deleterious ones. Hence, they reduce the effect of selection on substitution rates (Felsenstein 1974; Barton 1995), but for which mutations, and by how much? A number of theoretical and experimental studies have addressed these questions and have led to a quite diverse picture of particular interaction effects. These effects, which are sketched in Figure 1, include (i) interference between strongly beneficial mutations (clonal interference and related models) (Gerrish and Lenski 1998; Orr 2000; Rouzine *et al.* 2003; Wilke 2004; Desai *et al.* 2007; Park and Krug 2007, Hallatschek 2011) and between weakly selected mutations (McVean and Charlesworth 2000; Comeron and Kreitman 2002; Comeron *et al.* 2008), (ii) the effects of strongly beneficial mutations on linked neutral mutations (genetic hitchhiking, genetic draft) (Smith and Haigh 1974; Barton 2000; Gillespie 2001; Kim and Stephan 2003; Hermisson and Pennings 2005; Andolfatto 2007), and (iii) the effects of strongly deleterious mutations on linked neutral or weakly selected mutations (background selection) (Charlesworth *et al.* 1993; Charlesworth 1994, 1996; Kim and Stephan 2000; Bachtrog and Gordo 2004; Kaiser and Charlesworth 2009).

Adaptive evolution under linkage contains all of these processes, and the model developed in this article integrates positive and negative interference, background selection, and hitchhiking into a unified treatment of multiple in-
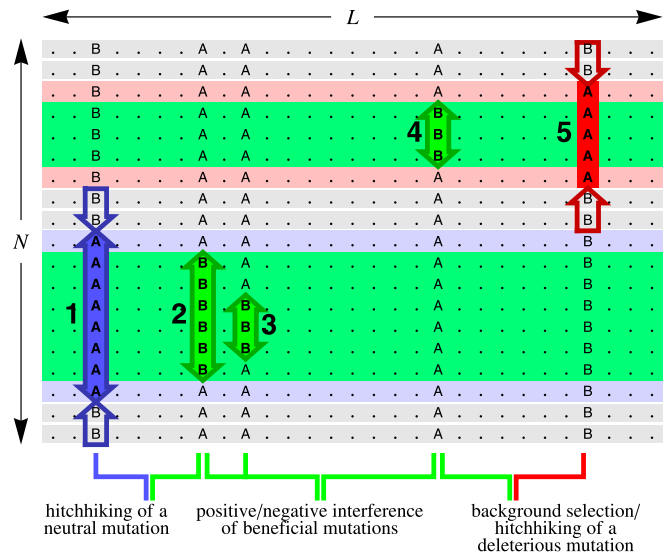

Figure 1 Interference interactions of mutations in linked genomes. Our model describes the evolution of N individuals with two-nucleotide genomes of length L. In a nonrecombining genome, this process is governed by positive and negative interference interactions between beneficial (green), neutral (blue), and deleterious (red) mutations. Shown are five mutations simultaneously present in the population; their expected frequency changes in the absence of genetic linkage are indicated by arrows. The fitness contribution of each mutation additively affects the fitness of all individuals carrying that mutation, which is indicated by the background color of the sequences. Linkage introduces the following interactions: allele 1 may be driven to fixation by allele 2 (hitchhiking of a neutral mutation), alleles 2 and 3 enhance each other's probability of fixation (positive interference between beneficial mutations), alleles 3 and 4 compete for fixation (negative interference between beneficial mutations), allele 4 may be driven to loss by allele 5 (background selection), or allele 5 may be driven to fixation by allele 4 (hitchhiking of a deleterious mutation).

teracting mutations. Specifically, the model describes the adaptive evolution of a finite asexual population, whose individuals have nonrecombining genotypes of finite length. Evolution takes place by mutations, genetic drift, and selection given by a genomic fitness function, which is specified by the distribution of selection coefficients between alleles at individual sequence sites and may be explicitly time dependent. The evolving population is described by its genome state, *i.e.*, by the probabilities of beneficial and deleterious alleles at fixed sites. The genome state influences the rate and distribution of selection coefficients for mutations in individuals: the better the population is adapted, the more sites are fixed at beneficial alleles and the more novel mutations will be deleterious. Thus, the scope of our genomic model goes beyond that of previous studies that analyze the statistics of substitutions given the rate and the effects of mutations as fixed input parameters (Gerrish and Lenski 1998; Desai *et al.* 2007; Park and Krug 2007). In particular, our model naturally includes nonstationary adaptation, *i.e.*, processes in which the distribution of selection effects for mutations becomes itself time dependent.

The key derivation of this article concerns the effects of interference interactions on the evolution of the genome state. We develop an approximate calculus for multiple simultaneous mutations. Specifically, we determine how the fixation probability of a specific target mutation is affected by positive and negative interference of other mutations. Since the target mutation can, in turn, act as interfering mutation, we obtain an approximate, self-consistent summation of interference interactions between all co-occurring mutations. We show that these interactions partition the adaptive dynamics into strongly beneficial driver mutations, which fix without substantial interference, and beneficial or deleterious passenger mutations, which suffer strong positive or negative interference.

Our analytical approach differs from the two classes of models analyzed in previous work. The clonal interference calculus (Gerrish and Lenski 1998) focuses on the dynamics of driver mutations, but it does not consider passenger mutations and neglects the effects of multiple co-occurring mutations. On the other hand, the traveling-wave approach assumes an ensemble of many co-occurring mutations, which have the same or a similar selective effect (Rouzine *et al.* 2003; Desai *et al.* 2007). The adaptive processes studied in this article—and arguably those in many real systems—take place in linked genomes with more broadly distributed selection coefficients and differ from both model classes: they are governed by interference interactions between *multiple*, but *few* strongly beneficial substitutions and their effect on weaker selected alleles. We show that this leads to *intermittent* fitness waves, which have large fluctuations and travel faster than fitness waves with deterministic bulk. This scenario and the results of our model are supported by simulations over a wide range of evolutionary parameters, which includes those of typical microbial evolution experiments. Its range of validity and the crossover to other modes of evolution are detailed in the *Discussion*.

We analyze our model and its biological implications for two specific scenarios of adaptive evolution. The first is a stationary adaptive process maintained by an explicitly time-dependent fitness "seascape", in which selection coefficients at individual genomic sites change direction at a constant rate (Mustonen and Lässig 2007, 2008, 2009). Such time dependence of selection describes changing environments, which can be generated by external conditions, migration, or coevolution. An example is the antigen–antibody coevolution of the human influenza virus (Bush *et al.* 1999). Our model predicts the regime of emergent neutral genomic sites, the speed of adaptation, and the population's degree of adaptation in its current environment. The second scenario is the approach to evolutionary equilibrium in a static fitness landscape, starting from a poorly adapted initial state. This case describes, for example, the long-term laboratory evolution of bacterial populations in a constant environment (Barrick *et al.* 2009). The predictions of our model are now time dependent: the regime of emergent neutral sites and the speed of adaptation decrease over time, while the degree of adaptation increases.

This article has two main parts. In the first part, we introduce a minimal genomic model for adaptation under linkage and present its general solution. In the second part, we discuss the application of the model to the scenarios of stationary adaptation and approach to equilibrium. In the *Discussion*, we draw general biological consequences and place our model into a broader context of asexual evolutionary processes.

Minimal Model for Adaptive Genome Evolution

We first introduce our model of genome evolution, as well as two key observables of the evolutionary process: the degree of adaptation, which measures the fitness of a population in its current environment, and the fitness flux, which we use as a measure of the speed of adaptation. We then calculate the fixation probability of beneficial and deleterious mutations and show, in particular, the emergence of neutrality.

Genome state and degree of adaptation

We consider an evolving asexual population of fixed size N, in which each individual has a genome of length L with two possible alleles per site. Our minimal fitness model is additive and fairly standard: each site is assigned a nonnegative selection coefficient f, which equals the fitness difference between its beneficial and its deleterious allele. The site selection coefficients are drawn independently from a normalized distribution $\rho(f)$, which is parameterized by its mean $\bar{f}$ and a shape parameter κ [we use a Weibull distribution, which has a tail of the form $\rho(f) \sim \exp[-(f/\bar{f})^{\kappa}]$; see section 6 of supporting information, File S1 for details]. In section 7.4 of File S1, we introduce an extension of the fitness model with simple fitness interactions (epistasis) between sites, and we show that such interactions do not affect the conclusions of this article.

Point mutations between nucleotides take place with a uniform rate μ. We assume that the population evolves in the low-mutation regime $\mu N \ll 1$, in which one of the two nucleotide alleles is fixed in the population at most sequence sites, and a fraction μN or less of the sites are polymorphic. In this regime, the two-allele genome model adequately describes the evolution of a genome with four nucleotides, because polymorphic sites with more than two nucleotides occur with negligible frequency. The genome state of the population is then characterized by the probability that a fixed site with selection coefficient f carries the beneficial allele, $\lambda_{\mathrm{b}}(f)$, or the probability that it carries the deleterious allele, $\lambda_{\mathrm{d}}(f)$. We use the familiar weak-mutation approximation $\lambda_{\mathrm{b}}(f) + \lambda_{\mathrm{d}}(f) = 1$ (this approximation neglects the probability that a site is polymorphic, which is of order $\mu N \log N$). The selection-dependent degree of adaptation, which is defined as $\alpha(f) = \lambda_{\mathrm{b}}(f) - \lambda_{\mathrm{d}}(f)$, varies between 0 for a randomly fixed genomic site and 1 for a perfectly adapted genomic site, which carries the beneficial allele

with probability 1. For example, a single locus with two alleles and time-independent selection coefficient f has $\alpha(f) = \tanh(Nf)$ at evolutionary equilibrium (Kimura 1962; Mustonen and Lässig 2007).

In a similar way, we define the degree of adaptation in the entire genome as the weighted average over all sites,

$$\alpha = \frac{1}{\bar{f}} \int_0^\infty \rho(f)df\, f[\lambda_b(f) - \lambda_d(f)]. \tag{1}$$

This summary statistic can be written in the form $\alpha = (F - F_0)/(F_{max} - F_0)$, where F is the Malthusian population fitness, F_0 is the fitness of a random genome, and F_{max} is the fitness of a perfectly adapted genome. Hence, α varies between 0 for a random genome and 1 for a perfectly adapted genome, and $1 - \alpha$ is a normalized measure of genetic load (Haldane 1937; Muller 1950). For an adaptive process, the lag of the genomic state behind the current fitness optimum can lead to a substantial reduction of α (Haldane 1957; Smith 1976), which is larger than the reduction due to mutational load (of order μ/f). In the following, we use α to measure the fitness cost of interference.

Mutations and speed of adaptation

A genomic site of selection coefficient f evolves by beneficial mutations with selection coefficient $\sigma = f > 0$ and by deleterious mutations with selection coefficient $\sigma = -f < 0$ (recall that f is, by definition, nonnegative). Hence, the distributions of beneficial and deleterious alleles at fixed sites, $\lambda_{b,d}(f)$, determine the genome-wide rate of mutations with a given selection coefficient occurring in the population,

$$U(\sigma) = \begin{cases} \mu NL\rho(\sigma)\lambda_d(\sigma) & (\sigma>0,\ \text{beneficial mutations}), \\ \mu NL\rho(|\sigma|)\lambda_b(|\sigma|) & (\sigma<0,\ \text{deleterious mutations}). \end{cases} \tag{2}$$

The total rates of beneficial and deleterious mutations are obtained by integration over all positive and negative selection coefficients, $U_b = \int_0^\infty d\sigma U(\sigma)$ and $U_d = \int_{-\infty}^0 d\sigma U(\sigma)$. The distribution $U(\sigma)$ is conceptually different from the distribution $\rho(\sigma)$ of selection coefficients at genomic sites, because it depends on the genome state $\lambda_{b,d}(f)$. Therefore, both the shape of $U(\sigma)$ and the total rates $U_{b,d}$ are in general time dependent, even if $\rho(f)$ is fixed. A similar coupling between the distribution $U(\sigma)$ and the adaptive state of the population has been discussed previously, for example, in the context of Fisher's geometrical model (Martin and Lenormand 2006; Tenaillon *et al.* 2007; Waxman 2007; Rouzine *et al.* 2008). For stationary evolution with an exponential $\rho(f)$, we find the distribution $U(\sigma)$ of beneficial mutations ($\sigma > 0$) is approximately exponential as well, which is in accordance with the form suggested by previous studies (Gillespie 1984; Imhof and Schlotterer 2001; Orr 2003; Rokyta *et al.* 2005; Kassen and Bataillon 2006; Eyre-Walker and Keightley 2007; MacLean and Buckling 2009).

The selection-dependent substitution rate is given by the product of the mutation rate and probability of fixation $G(\sigma)$ of a mutation with selection coefficient σ,

$$V(\sigma) = G(\sigma)U(\sigma). \tag{3}$$

For unlinked sites, the fixation probability is given by Kimura's classical result, $G_0(\sigma) = (1 - \exp[-2\sigma])/(1 - \exp[-2N\sigma])$ (Kimura 1962). Computing this probability for linked sites is at the core of this article: $G(\sigma)$ depends not only on the selection strength σ and population size N, but also on the interference interactions between co-occurring mutations shown in Figure 1.

To measure the speed of adaptation contributed by sites with selection coefficient f, we define the fitness flux $\Phi(f) = f[V(f) - V(-f)]$. The total fitness flux

$$\Phi = \int_0^\infty df\, \Phi(f) = \int_{-\infty}^\infty d\sigma\, \sigma V(\sigma) = V\bar{\sigma}_V \tag{4}$$

is simply the product of the total rate V and the average selection coefficient $\bar{\sigma}_V$ of substitutions (Mustonen and Lässig 2007). If evolution is adaptive, it can be shown that Φ is always positive (Mustonen and Lässig 2010), which reflects an excess of beneficial over deleterious substitutions. In the following, we use Φ to measure the reduction in speed of adaptation due to interference (Gerrish and Lenski 1998).

Adaptive dynamics

In an additive fitness model, adaptive evolution can be maintained at a stationary rate if the selection coefficients at individual genomic sites are time dependent: changes in selection open new windows of positive selection and trigger adaptive response by beneficial mutations. Selection changes at a specific genomic site result from changes in a population's environment, as well as from substitutions at other sites coupled by epistasis. In this sense, time-dependent selection is a proxy for epistasis within an additive fitness model (see section 7.4 in File S1 for an explicit epistatic model). Here, we use the minimal dynamic fitness model introduced in Mustonen and Lässig (2007, 2010), in which the direction of selection at each site flips according to an independent Poisson process. A selection flip exchanges the beneficial and the deleterious allele at a given site, whereas the magnitude f of selection remains constant for simplicity. The rate γ of selection flips per site is small, such that a given direction of selection persists for longer than the typical fixation time of a mutation; a sufficient condition is $N\gamma \ll 1$. In this regime, selection flips trigger adaptive substitutions, which occur at a rate γL for independently evolving sites under substantial selection ($Nf \gg 1$). In a linked genome, however, the rate of adaptation can be substantially smaller, because adaptive mutations overlap in time and interfere with each other's evolution. In our model, this effects sets in approximately at $2N\gamma L \sim 1$ (see below).
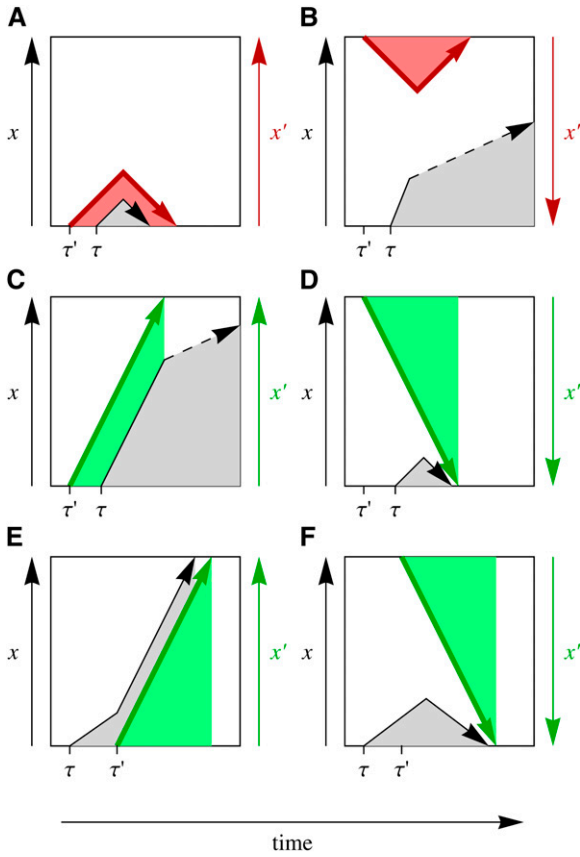
Figure 2 Interference time series diagrams. A target mutation with origination time τ and frequency $x(t)$ (black arrow) is subject to a stronger interfering mutation with origination time τ' and frequency $x'(t)$ (colored arrow). The interactions between this pair of mutations can be classified as follows: (A and B) *Interference by a deleterious background mutation* (red arrow). (A) The target mutation originates on the deleterious allele of the interfering mutation and is driven to loss, and (B) the target mutation originates on the ancestral (beneficial) allele of the interfering mutation and is enhanced in frequency. (C and D) *Interference by a beneficial background mutation* (green arrow). (C) The target mutation originates on the beneficial allele of the interfering mutation and is enhanced in frequency, and (D) the target mutation originates on the ancestral (deleterious) allele of the interfering mutation and is driven to loss. (E and F) *Interference by a beneficial future mutation* (green arrow). (E) The interfering mutation originates on the new allele of the target mutation and drives it to fixation, and (F) the interfering mutation originates on the ancestral background of the target locus and drives the target mutation to loss.

Changes in selection and substitutions together determine the change of the genome state,

$$\frac{d\lambda_{\text{b}}(f)}{dt} = -\frac{d\lambda_{\text{d}}(f)}{dt}$$
$$= \frac{1}{L\rho(f)}[V(f) - V(-f)] + \gamma[\lambda_{\text{d}}(f) - \lambda_{\text{b}}(f)], \quad (5)$$

which, in turn, determines the substitution rates $V(\sigma)$ by Equations 2 and 3. These coupled dynamics admit an iterative solution, once we have computed the fixation probability $G(\sigma)$ (see below). Equation 5 can be expressed as a relation between the selection-dependent fitness flux $\Phi(f)$, the degree of adaptation $\alpha(f)$, and its time derivative,

$$\Phi(f) = L\rho(f)f\left[\frac{1}{2}\frac{d\alpha(f)}{dt} + \gamma\alpha(f)\right]. \quad (6)$$

Hence, a population's fitness in a time-dependent environment increases by fitness flux and decreases by changes of the fitness seascape. A similar relation follows for the total fitness flux Φ (Equation 4) and the genome-averaged degree of adaptation (Equation 1),

$$\Phi = L\bar{f}\left[\frac{1}{2}\frac{d\alpha}{dt} + \gamma\alpha\right]. \quad (7)$$

In the second part of this article, we analyze these dynamics in two specific scenarios.

Stationary adaptation under time-dependent selection: In this case, the genome state becomes static, $d\lambda_{\text{b,d}}/dt = d\alpha/dt = 0$. This state is determined by the fixation probabilities $G(\sigma)$ and the selection flip rate γ,

$$\lambda_{\text{b}}(f) = 1 - \lambda_{\text{d}}(f) = \frac{NG(f) + \gamma/\mu}{NG(f) + NG(-f) + 2\gamma/\mu} \quad (8)$$

(Mustonen and Lässig 2007). The stationary fitness flux becomes proportional to the degree of adaptation,

$$\Phi = \bar{f}\gamma L\alpha \quad \text{(stationary adaptation);} \quad (9)$$

that is, the actual fitness flux Φ is only a fraction α of the maximal flux $\bar{f}\gamma L$ required for perfect adaptation.

Approach to equilibrium under time-independent selection: In a static fitness landscape ($\gamma = 0$), adaptation is the approach to a mutation–selection–drift equilibrium state. In this case, the fitness flux is simply the change in fitness,

$$\Phi = \frac{dF}{dt} = \frac{\bar{f}L}{2}\frac{d\alpha}{dt} \quad \text{(approach to equilibrium);} \quad (10)$$

that is, the population fitness increases with time toward its equilibrium value.

Fixation probability of interacting mutations

The missing piece in the coupled dynamics of substitutions and genomic state (Equations 2, 3, and 5) is the fixation probability $G(\sigma)$. Consider a *target* mutation with origination time τ and selection coefficient σ, which is subject to interference by a mutation with origination time τ' and a selection coefficient σ' larger in magnitude, $|\sigma'| > |\sigma|$; this hierarchical approximation is detailed further below. We classify pairwise interactions between interfering and target mutations by three criteria:

i. Temporal order: The interfering mutation originates either at a time $\tau' < \tau$ (we call this case background interfering mutation) or at a time $\tau' > \tau$ (future interfering mutation).

ii. Direction of selection on the interfering mutation: Deleterious or beneficial.iii. Allele association: The target mutation may occur on the ancestral or the new allele of a background interfering mutation; similarly, a future interfering mutation may occur on the ancestral or the new allele of the target mutation. Figure 2 shows this classification: a target mutation interacts with a deleterious background interfering mutation (Figure 2, A and B), with a beneficial background interfering mutation (Figure 2, C and D), and with a beneficial future interfering mutation (Figure 2, E and F). The case of deleterious future interfering mutations is not shown, because their contribution to the fixation probability is negligible.

As a first step, we evaluate the conditional fixation probability of a target mutation, $G(\sigma, \tau \mid \sigma', \tau')$, for each of the cases of Figure 2, A–F, and for given selection coefficients σ, σ' and origination times τ, τ'. This step is detailed in the *Appendix*. The net contribution of deleterious background (Figure 2, A and B) is found to be small, because the reduction in fixation probability for association with the deleterious allele is offset by an enhancement for association with the beneficial allele. The relationship of this result to previous studies of background selection is discussed in section 1 of File S1. Beneficial background mutations (Figure 2, C and D) retain a net effect on the fixation probability of target mutation. The largest effect turns out to arise from future interfering mutations (Figure 2, E and F).

We now derive an approximate expression for the total fixation probability of a target mutation on the basis of pair interactions with multiple interfering mutations. Clearly, a straightforward "cluster expansion" makes sense only in a regime of *dilute* selective sweeps at sufficiently low rates of beneficial mutations, where the interference interactions of Figure 2, C–F, are infrequent. However, we are primarily interested in adaptive processes under linkage in the *dense-sweep* regime at high rates of beneficial mutations, which generates strongly correlated clusters of fixed mutations nested in each other's background (the crossover between these regimes is further quantified below). We treat the dense-sweep regime by an approximation: each sweep is associated with a unique *driver mutation*, which is the strongest beneficial mutation in its cluster. The driver mutation itself evolves free of interference, but it influences other mutations by interference; that is, we neglect the feedback of weaker beneficial and deleterious mutations on the driver mutation. In this hierarchical approximation, the coherence time of a sweep is set by the fixation time of its driver mutation, $\tau_{\text{fix}}(\sigma) = 2 \log(2N\sigma)/\sigma$ (see section 2 of File S1). The sweep rate becomes equal to the rate of driver mutations, $V_{\text{drive}}(\sigma)$, and is given by the condition that no stronger selective sweep occurs during the interval $\tau_{\text{fix}}(\sigma)$. Hence, we obtain a self-consistent relation,

$$V_{\text{drive}}(\sigma) = p_{\text{drive}}(\sigma)G_0(\sigma)U(\sigma) \qquad (11)$$

with

$$p_{\text{drive}}(\sigma) = \exp\left[-\tau_{\text{fix}}(\sigma) \int_\sigma^\infty V_{\text{drive}}(\zeta)d\zeta \right], \qquad (12)$$

which can be regarded as a partial summation of higher-order interference interactions characteristic of the dense-sweep regime (see the *Appendix* for details). These expressions and the underlying hierarchical approximation are similar to the model of clonal interference by Gerrish and Lenski (1998). This model determines an approximation of the sweep rate, $V_{\text{GL}}(\sigma)$, by requiring that no *negative* interference by a future interfering mutation occurs (see below for a quantitative comparison with our model).

Consistent with the hierarchical approximation, we can interpret the diagrams of Figure 2, C–E, as *effective* pair interactions between a target mutation and a selective sweep, which is represented by its driver mutation. The target mutation can strongly interact with two such sweeps, the last sweep before its origination (with parameters $\sigma' > \sigma$ and $\tau' > \tau$) and the first sweep after its origination (with parameters $\tau'' > \tau$ and $\sigma'' > \sigma$). These two sweeps affect the fixation probability $G(\sigma)$ in a combined way: the target mutation can be fixed only if it appears on the background of the last background sweep and if it is itself the background of the first future sweep. The resulting conditional fixation probability of the target mutation, $G(\sigma, \tau \mid \sigma', \tau', \sigma'', \tau'')$, is a straightforward extension of the form obtained for a single interfering mutation. The full fixation probability $G(\sigma)$ is then obtained by integration over the selection coefficients σ', σ'' with weights $V_{\text{drive}}(\sigma')$ and $V_{\text{drive}}(\sigma'')$ and over the waiting times $\tau - \tau'$, $\tau'' - \tau$. This calculation and the result for $G(\sigma)$, Equation A6, are given in the *Appendix*.

The fixation probability can be expressed as the sum of driver and passenger contributions,

$$G(\sigma) = p_{\text{drive}}(\sigma)G_0(\sigma) + [1 - p_{\text{drive}}(\sigma)]G_{\text{pass}}(\sigma). \qquad (13)$$

A passenger mutation fixes predominantly by interference from other stronger mutations, which results in a fixation probability G_{pass} (see Equation A8). The arguments leading to Equations 11 and 12 must be modified, if the distribution of selection coefficients $p(f)$ falls off much faster than exponentially (Fogle *et al.* 2008). In that case, we can still describe the fixation probability of a target mutation as the result of interference interaction with the closest past and future sweep, but these sweeps may contain several driver mutations of comparable strength (see *Discussion*). The system of Equations 2, 3, 8, 11, 12, and 13 can be solved numerically using a straightforward iterative algorithm, which is detailed in section 4 of File S1.

Emergent neutrality

For mutations of sufficiently weak effect, the fixation probability takes the particularly simple form

$$G(\sigma) \simeq G_{\text{pass}}(\sigma) \simeq G_0\left(\frac{\sigma}{2N\tilde{\sigma}}\right) \quad (\text{for } -\tilde{\sigma} < \sigma < \tilde{\sigma}), \qquad (14)$$

where the neutrality threshold $\tilde{\sigma}$ is given by the total sweep rate $V_{\mathrm{drive}} = \int_0^\infty V_{\mathrm{drive}}(\zeta)d\zeta$,

$$\tilde{\sigma} = \frac{1}{2N} + V_{\mathrm{drive}} \qquad (15)$$

(see the *Appendix*). These are the central equations of this article. They show how neutrality emerges for strong adaptive evolution under linkage. Specifically, the relation for $\tilde{\sigma}$ delineates two dynamical modes: the *dilute sweep mode* ($V_{\mathrm{drive}} \lesssim 1/2N$), where the neutrality threshold is set by genetic drift to the Kimura value $\tilde{\sigma} \simeq 1/2N$ (Kimura 1962), and the *dense sweep mode* ($V_{\mathrm{drive}} \lesssim 1/2N$), where interference effects generate a broader neutrality regime with $\tilde{\sigma} \simeq V_{\mathrm{drive}}$. The transition between these modes marks the onset of clonal interference as defined in previous work (Wilke 2004; Park and Krug 2007). For stationary adaptation in a time-dependent fitness seascape, the upper bound $V_{\mathrm{drive}} \simeq \gamma L$ produces the estimate $2N\gamma L > 1$ for the crossover from dilute to dense sweeps. However, Equations 14 and 15 remain valid for nonstationary adaptation, where the neutrality threshold $\tilde{\sigma}$ becomes time dependent (see below).

In summary, interference interactions in the dense-sweep mode produce the following selection classes of mutations and genomic sites.

Emergent neutrality regime: Mutations with selection coefficients $-\tilde{\sigma}<\sigma<\tilde{\sigma}$ fix predominantly as passenger mutations. Their near-neutral fixation probability (Equation 14) is the joint effect of positive and negative interference. Compared to unlinked mutations, $G(\sigma)$ is reduced for beneficial mutations and enhanced for deleterious mutations. Accordingly, sites with selection coefficients $f<\tilde{\sigma}$ have near-random probabilities of their alleles.

Adaptive regime: Mutations with effects $\sigma>\tilde{\sigma}$ have a fixation rate significantly above the neutral rate and, hence, account for most of the fitness flux. Moderately beneficial mutations ($\sigma\gtrsim\tilde{\sigma}$) still fix predominantly as passengers, whereas strongly beneficial mutations ($\sigma\gg\tilde{\sigma}$) are predominantly drivers. Hence, the fixation rate increases to values of order $G_0(\sigma) \simeq 2\sigma$, which are characteristic of unlinked mutations. Accordingly, sites with $f>\tilde{\sigma}$ evolve toward a high degree of adaptation.

Deleterious passenger regime: Mutations with $\sigma<-\tilde{\sigma}$ can fix by positive interference, *i.e.*, by hitchhiking in selective sweeps. This effect drastically enhances the fixation rate in comparison to the unlinked case. It follows the heuristic approximation $G(\sigma) \approx G_{\mathrm{pass}}(\sigma) \sim \exp(-|\sigma|/\tilde{\sigma})$, which extends the linear reduction in the effective strength of selection (Equation 14) obtained in the emergent neutrality regime.

Applications to Adaptive Scenarios

Here we analyze our results for two specific adaptive scenarios: stationary adaption in a fitness seascape and

approach to equilibrium in a static fitness landscape. Detailed comparisons of our analytical results with numerical simulations show that our approach is valid in both cases. In particular, we always find a regime of emergent neutrality with a threshold $\tilde{\sigma}$, which is time dependent for nonstationary processes.

Stationary adaptation in a fitness seascape

Stationary adaptation in our minimal fitness seascape is characterized by ongoing selection flips, which occur with rate γ per site and generate an excess of beneficial over deleterious substitutions, with rates $V(\sigma) > V(-\sigma)$ (see Equation 5). Figure 3A shows the selection-dependent fixation probability $G(\sigma)$ in a linked genome undergoing stationary adaptive evolution. The emergent neutrality regime ($|\sigma|<\tilde{\sigma}$), the adaptive regime ($\sigma>\tilde{\sigma}$), and the deleterious passenger regime ($\sigma<-\tilde{\sigma}$) are marked by color shading. The self-consistent solution of our model (Figure 3A, red line) is in good quantitative agreement with simulation results for a population of linked sequences (Figure 3A, open circles; simulation details are given in section 6 of File S1). Data and model show large deviations from single-site theory (Figure 3A, long-dashed blue line), which demonstrate strong interference effects in the dense-sweep regime. Figure 3A also shows an effective single-site probability with a globally reduced efficacy of selection, $G_0(\sigma/2N\tilde{\sigma})$ (short-dashed blue line). As discussed above, a global reduction in selection efficacy fails to capture the adaptive regime, where mutations have a fixation probability approaching the single-site value 2σ.

The crossover between adaptive and emergent-neutrality regimes implies a nonmonotonic dependence of the substitution rate $V(\sigma)$ on the population size: in sufficiently small populations sizes (where $\tilde{\sigma}<\sigma$), beneficial mutations of strength σ are likely to be driver mutations. Hence, $V(\sigma)$ is an increasing function of N with the asymptotic behavior $V(\sigma) \simeq \mu N\sigma$ familiar for unlinked sites. In larger populations (where $\tilde{\sigma}>\sigma$), the same mutations are likely to be passenger mutations and $V(\sigma)$ decreases with N toward the neutral rate $V(\sigma) \simeq \mu$. The maximal substitution rate is expected to be observed in populations where $\tilde{\sigma}$ is similar to σ. By the same argument, deleterious mutations have a minimum in their substitution rate in populations where $\tilde{\sigma}$ is similar to $|\sigma|$.

Furthermore, it is instructive to compare our results for stationary adaptation with the classical clonal interference model (Gerrish and Lenski 1998) (see section 5 of File S1 for details). This model focuses exclusively on negative interference between strongly beneficial mutations, the case shown in Figure 2F. The resulting approximation for the fixation probability, $G_{\mathrm{GL}}(\sigma) = V_{\mathrm{GL}}(\sigma)/U(\sigma)$, is shown in Figure 3A (brown line). It captures two salient features of the stationary adaptive process: the behavior of strongly beneficial driver mutations and the drastic reduction of the total substitution rate caused by interference. However, the full spectrum of $G(\sigma)$ requires taking into account positive and
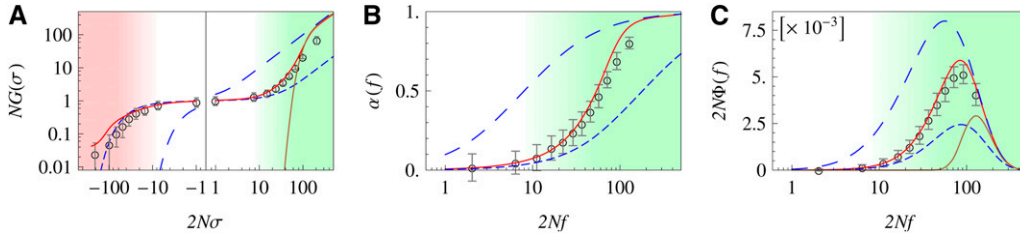
Figure 3 Selection regimes of stationary adaptation. (A) Selection-dependent fixation probability of mutations $G(\sigma)$, scaled by the population size N. Analytic model solution (red line) and simulation results (circles) show three regimes of selection: (i) effective neutrality regime (white background), where $G(\sigma)$ takes values similar to the fixation probability of independent sites with reduced selection, $G_0(\sigma/2N\tilde{\sigma})$ (short-dashed blue line); (ii) adaptive regime (green), where $G(\sigma)$ crosses over to the fixation probability for unlinked sites with full selection, $G_0(\sigma)$ (long-dashed blue line) [the strong-selection part of this crossover is captured by the Gerrish–Lenski model, $G_{GL}(\sigma)$ (brown line, see section 5 of File S1); and (iii) strongly deleterious passenger regime (red), where $G(\sigma)$ is exponentially suppressed, but drastically larger than for unlinked sites (long-dashed blue line) due to hitchhiking in selective sweeps. (B and C) Selection-dependent degree of adaptation $\alpha(f)$ and fitness flux $\Phi(f)$, scaled by N. Analytical model solution (red line) and simulation results (circles) show two regimes of selection: (i) effective neutrality regime (white background), where $\alpha(f)$ and $\Phi(f)$ take values similar to those of unlinked sites with reduced selection (short-dashed blue line), and (ii) adaptive regime (green), where $\alpha(f)$ and $\Phi(\sigma)$ cross over to values of unlinked sites with full selection (short-dashed blue lines). The strong-selection part of the crossover for Φ is captured by the Gerrish–Lenski model, $\Phi_{GL}(f)$ (brown line, see section 5 of File S1). System parameters are $N = 2000$, $L = 1000$, $2N\mu = 0.025$, $2N\gamma = 0.1$, and $2N\bar{f} = 50$, and simulation time is 2×10^6 generations.

negative interference. Furthermore, mutation-based models with rate and effect of beneficial mutations as input parameters cannot predict the degree of adaptation, as discussed in section 7.3 of File S1.

An important feature of the adaptive dynamics under linkage is the relative weight of driver and passenger mutations in selective sweeps. The fixation probability is highest for strongly beneficial mutations ($\sigma \gg \tilde{\sigma}$), which are predominantly driver mutations. Nevertheless, the majority of observed substitutions can be moderately adaptive or deleterious passenger mutations (in the process of Figure 3, for example, ~60% of all substitutions are passengers, 20% of which are deleterious). The distribution of mutation rates, $U(\sigma)$, and the distribution of fixation rates, $V(\sigma) = V_{drive}(\sigma) + V_{pass}(\sigma)$, are shown in Figure S3 in File S1.

In Figure 3, B and C, we plot the selection-dependent degree of adaptation $\alpha(f)$ and the fitness flux $\Phi(f)$ at stationarity, which are proportional to each other according to Equation 6. Simulation results are again in good agreement with our self-consistent theory, but they are not captured by single-site theory, single-site theory with globally reduced selection efficacy, or Gerrish–Lenski theory. The functions $\alpha(f)$ and $\Phi(f)$ display the emergent neutrality regime ($f < \tilde{\sigma}$) and the adaptive selection regime ($f > \tilde{\sigma}$) for genomic sites, which are again marked by color shading. Using Equations 8 and 14, we can obtain approximate expressions for both regimes. Consistent with near-neutral substitution rates, sites in the emergent neutrality regime have a low degree of adaptation and fitness flux:

$$\alpha(f) = \frac{\Phi(f)}{f\gamma L\rho(f)} \simeq \frac{1}{(1 + \gamma/\mu)} \frac{f}{2\tilde{\sigma}}. \quad (16)$$

Hence, fixed sites in this regime have nearly random alleles: they cannot carry genetic information. Two processes contribute to this degradation: negative interference slows down the adaptive response to changes in selection, and hitchhiking in selective sweeps increases the rate of deleterious substitutions. By contrast, sites in the adaptive regime

($f > \tilde{\sigma}$) have a high degree of adaptation and generate most of the fitness flux. Sites under moderate selection ($f \gtrsim \tilde{\sigma}$) are still partially degraded by interference, and the negative component of fitness flux (*i.e.*, the contribution from deleterious substitutions) is peaked in this regime (see Figure S4 in File S1). Strongly selected sites ($f \gg \tilde{\sigma}$) are approximately independent of interference. Hence, their degree of adaptation and fitness flux increase to values characteristic of unlinked sites,

$$\alpha(f) = \frac{\Phi(f)}{f\gamma L\rho(f)} \simeq \frac{f}{f + \gamma/(\mu N)}. \quad (17)$$

In addition to the selection-dependent quantities discussed so far, our theory also predicts how genome-wide characteristics of the adaptive process depend on its input parameters. The adaptively evolving genome is parameterized by the mutation rate μ, by the effective population size N, and by three parameters specific to our genomic model: average strength $\bar{f}$ and flip rate γ of selection coefficients, and genome length L. As an example, Figure 4 shows the dependence of the average degree of adaptation α on γ and on L, with all other parameters kept fixed (recall that according to Equation 9, this also determines the behavior of the total fitness flux, $\Phi = \alpha\bar{f}\gamma L$). The genome-wide rate of selection flips, γL, describes the rate at which new opportunities for adaptive substitutions arise at genomic sites. With increasing supply of opportunities for adaption, interference interactions become stronger. This leads to an increase in the neutrality threshold $\tilde{\sigma}$, a decrease in the degree of adaptation α, and a sublinear increase of the fitness flux Φ. All of these effects are quantitatively reproduced by the self-consistent solution of our model. As shown in Figure 4, low values of the degree of adaptation α are observed over large regions of the evolutionary parameters γ and L. This indicates that a substantial part of the genome can be degraded to a nearly random state, implying that interference effects can compromise biological functions (see *Discussion*).
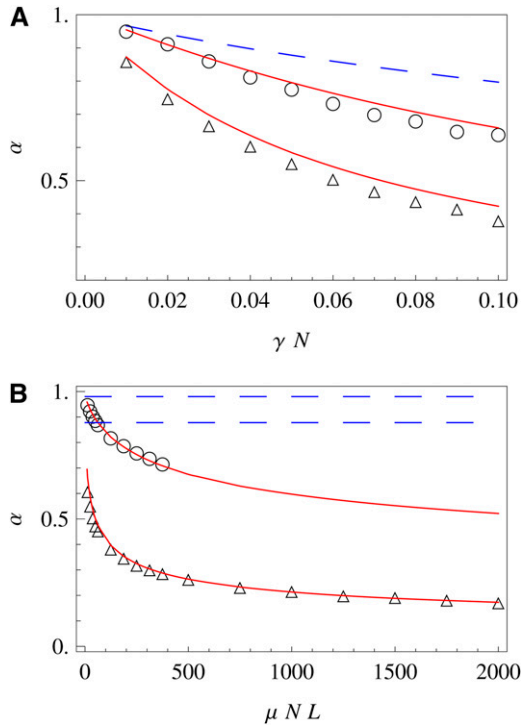
Figure 4 Degree of adaptation at stationarity. Results from our model (red lines), simulation results (circles), and values for independent sites (dashed blue lines) of the degree of adaptation are plotted (A) against the total selection flip rate γL for two different values of the genome length, $L = 200$ (circles) and $L = 2000$ (diamonds), and (B) against the total genomic mutation rate μNL for two different values of the selection flip rate, $2N\gamma = 0.01$ (circles) and $2N\gamma = 0.1$ (triangles). Note that for the smaller value of γ, the time to reach stationarity is very long, which limits the numerical results to smaller values of the system size. Other system parameters are $N = 4000$, $2N\mu = 0.025$, and $2N\bar{f} = 50$, and simulation time is 8×10^5 generations.

Approach to equilibrium in a fitness landscape

A particular nonstationary adaptive process is the approach to evolutionary equilibrium in a static fitness landscape, starting from a poorly adapted initial genome state. Here, we analyze this mode in our minimal additive fitness model. We choose an initial state at time $t = 0$ from the family of stationary states presented in the previous section (which is characterized by a value $\gamma_{initial} > 0$). We study the evolution of this state for $t > 0$ toward equilibrium under time-independent selection ($\gamma = 0$; see section 4 of File S1 for details of the numerical protocol). Unlike for stationary adaptation, the observed dynamics now depend on the initial state. Figure 5 shows the selection-dependent degree of adaptation $\alpha(f)$ of this process at three consecutive times. The self-consistent solution of our model is again in good agreement with simulation data. There is still a clear grading of genomic sites into an emergent neutrality regime and an adaptive regime, which is again marked by color shading. The neutrality threshold $\tilde{\sigma}(t)$ is now a decreasing function of time. Figure 5 also shows the time derivative of the degree of adaptation, which equals half the adaptive substitution
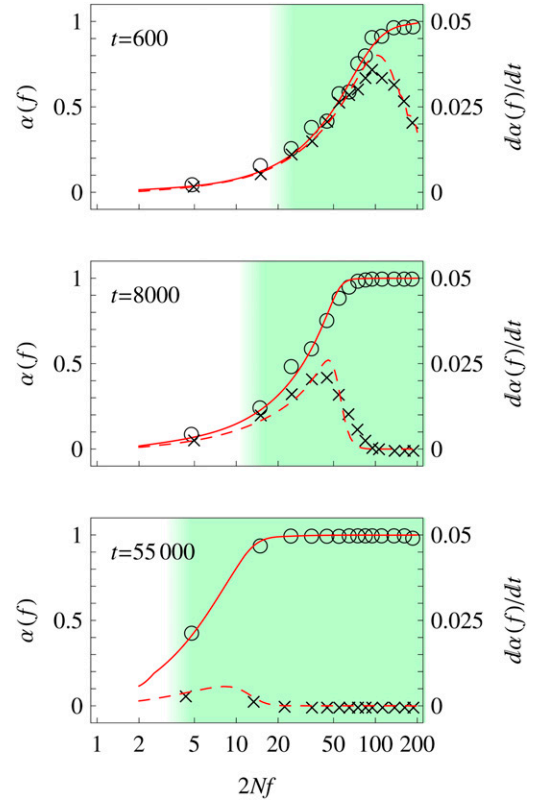

Figure 5 Selection regimes for approach to equilibrium. The population evolves from a poorly adapted initial state to a high-fitness equilibrium state. The degree of adaptation (circles and solid line) and its time derivative (which is related to the adaptive rate per site, crosses and dashed line) are shown for three consecutive times, $t = 600$, 8000, and $55,000$ generations. Theory lines are obtained by numerically solving Equation 6. The emergent neutrality regime ($\sigma < \tilde{\sigma}$) and the adaptive regime ($\sigma > \tilde{\sigma}$) are shown by color shading; the neutrality threshold $\tilde{\sigma}$ decreases with time. Parameters are $N = 1000$, $L = 1000$, $2N\bar{f} = 50$, and $2N\mu = 0.025$. The initial state is a stationary state with $2N\gamma = 0.1$.

rate per site: $d\alpha(f)/dt = [V(f) - V(-f)]/2\rho(f)$ by Equations 4 and 6. Data and model solution show that the adaptive process is nonuniform: at a given time t, adaptation is peaked at sites of effect $f \sim \tilde{\sigma}(t)$, while sites with stronger selection have already adapted at earlier times and sites with weaker selection are delayed by interference. Thus, our model predicts a nonmonotonic behavior of the adaptive rate $d\alpha(f)/dt$ on time: for sites with a given selection coefficient f, this rate has a maximum at some intermediate time when $\tilde{\sigma}(t) = f$, after interference effects have weakened and before these sites have reached equilibrium. This result mirrors the maximum of the substitution rate $V(\sigma)$ at some intermediate population size for stationary adaptation. As before, a substantial fraction of substitutions are passengers in selective sweeps.

Figure 6A shows the evolution of the genome-averaged degree of adaptation, α, and of the mean population fitness, F, which are linearly related by Equation 1. The fitness flux $\Phi = dF/dt$ and the total substitution rate V are plotted in Figure 6B. According to Equation 4, these quantities are
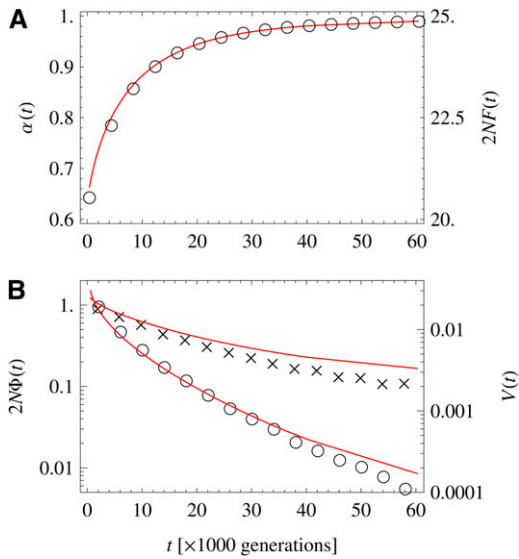
Figure 6 Degree and speed of adaptation for approach to equilibrium. (A) The degree of adaptation evolves from a poorly adapted initial state toward equilibrium at high fitness. Simulation results are shown as circles, and theory predictions as red lines. (B) The fitness flux (circles) and the substitution rate (crosses) as a function of time, together with theory (red lines). The substitution rate decays slower than the fitness flux, due to a decrease in $\tilde{\sigma}$. Theory predictions are obtained by numerically solving Equation 6. Parameters are the same as in Figure 5.

linked in a time-dependent way, $\Phi(t) = V(t)\tilde{\sigma}(t) \approx V(t)\tilde{\sigma}(t)$. We observe that fitness increases monotonically with time. Its rate of increase Φ rapidly slows down as the system comes closer to evolutionary equilibrium, whereas the total substitution rate V shows a slower approach to equilibrium. A qualitatively similar time dependence of fitness and substitution rate has been reported in a long-term bacterial evolution experiment by Barrick *et al.* (2009).

Discussion

Interference can dominate genetic drift

Interference interactions in the dense-sweep regime may be complicated in their details, but their net effect is simple: genomic sites with selection coefficients σ smaller than a threshold $\tilde{\sigma}$ have nearly random fixed alleles, and mutations at these sites fix with near-neutral rates. The neutrality threshold $\tilde{\sigma}$ is given by the total rate of selective sweeps, V_{drive}, as shown in Equation 15. Emergent neutral mutations, as well as more deleterious changes, fix as passengers in selective sweeps. That is, both classes of mutations are subject to interference, not genetic drift, as a dominant stochastic force. The resulting fixation rates $V(\sigma)$ depend only weakly on the effective population size N. Mutations with larger beneficial effect ($\sigma > \tilde{\sigma}$) suffer gradually weaker interference interactions. Hence, their fixation rates show a drastic increase toward the Haldane–Kimura value $V(\sigma) = 2\mu N\sigma$ set by genetic drift.

At a qualitative level, these results tell the story of the Hill–Robertson effect: genetic linkage reduces the efficacy of selection. Quantitatively, they demonstrate that emergent

neutrality is not equivalent to a simple reduction in effective population size. The fixation rate of emergent neutral and deleterious passenger mutations can heuristically be interpreted as a linear reduction in effective population size by a factor $2N\tilde{\sigma}$, but this approximation breaks down for mutations with larger beneficial effect; see Equation 14 and Figure 3A. In other words, we cannot absorb the effects of interference into a single modified strength of genetic drift. Of course, both interference and genetic drift are stochastic processes that randomize alleles of genomic sites. However, they have fundamentally different characteristics: genetic drift is a diffusion process causing *independent* changes in allele frequencies in each generation, whereas interference generates *coherent* changes over time intervals given by the inverse selection coefficient of the driver mutation.

Fluctuations and intermittency of the adaptive process

Given genetic drift and interference as stochastic driving forces, how stochastic are the resulting adaptive substitution dynamics? This question has been addressed in several recent studies, which treat the adaptive process as a traveling fitness wave (Rouzine *et al.* 2003, 2008; Desai *et al.* 2007; Hallatschek 2011). If all mutations are assumed to have the same effect, these models are solvable. One finds a traveling wave with a deterministic bulk of stationary shape (given by a mutation–selection flux state) and a stochastic tip. The variance of this wave determines its speed (*i.e.*, the fitness flux) by Fisher's fundamental theorem. Given a stationary bulk of the wave, the fitness flux has only small fluctuations around its mean value. However, the recent solvable model of Hallatschek (2011) contains large fluctuations in population size, which may be related to fluctuations in fitness flux.

The adaptive process studied in this article shows a drastically different behavior. In our model, fitness effects at genomic sites follow a distribution $\rho(f)$ with shape parameter κ. For the case of exponential $\rho(f)$ (given by $\kappa = 1$), a snapshot of the population's fitness distribution at a given point in time is shown in Figure 7A. This distribution has large shape fluctuations throughout its bulk, not just at the tip. It shows that the adaptive process is dominated by *few* co-occurring beneficial mutations of large effect, whereas a stationary wave is maintained by many mutations of smaller effect. As a consequence, the fitness flux becomes *intermittent*: on small timescales, it has large fluctuations around its mean value, as shown in Figure 7A. Movies of the intermittent fitness wave are available; see Figure 8. Importantly, this strong stochasticity accelerates evolution: at given rate U_{b} and mean effect $\sigma_{\text{b}} = (1/U_{\text{b}}) \int_0^\infty \sigma U(\sigma)d\sigma$ of beneficial mutations, our model produces a much higher mean fitness flux than the traveling-wave solution. The reason is simple: a distribution of selection coefficients generates dynamics dominated by strong driver mutations, whose effect is substantially larger than the mean (Park *et al.* 2010).

How relevant is this mode of intermittent adaptation for actual populations? We expect our model to be applicable to microbial laboratory populations, which often fall into the
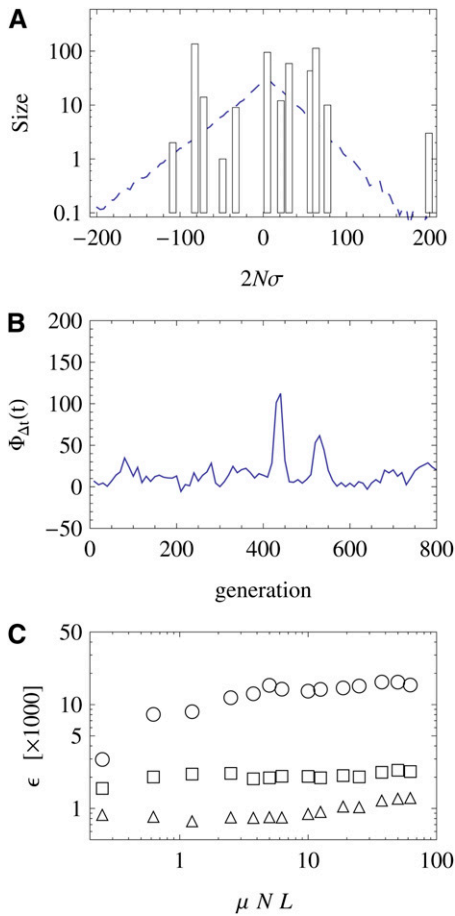
Figure 7 Stochasticity of the adaptive process. The speed of adaptation and the shape of the distribution of fitnesses in the population are governed by large fluctuations. (A) Snapshot of the fitness distribution in the population, centered around the mean fitness. The shape of this distribution is very different from the average shape, shown as a dashed line. The dynamics are governed by few fitness classes with a large number of individuals (note the logarithmic axis). The evolution of this distribution is shown in three movies; see Figure 8. (B) Time series of the cumulative fitness flux $\Phi_{\Delta t}(t)$, which is defined as the net selective effect of all allele frequency changes $\Delta x_i (i = 1, \dots, L)$ within a short time interval Δt (Mustonen and Lässig 2010): $\Phi_{\Delta t}(t) = \sum_{i=1}^{L} (\Delta x_i) \partial F(x_1(t), \dots, x_L(t), t)/\partial x_i$ (we use $\Delta t = 20$ generations). This flux is intermittent; *i.e.*, the traveling fitness wave has short-term boosts in its speed. (C) Stochasticity of the fitness flux. The ratio of variance and mean fitness flux $\Phi_{\Delta t}(t)$ over a population's history is plotted as a function of the total mutation rate μNL for selection shape parameters $\kappa = \frac{1}{2}$ (circles), $\kappa = 1$ (squares), and $\kappa = 2$ (triangles). For a given total mutation rate, the stochasticity is highest for $\kappa = \frac{1}{2}$ and decreases with increasing κ. However, the stochasticity remains approximately constant for increasing system size. Other simulation parameters are $N = 500$ (a and b), $N = 1000$ (c), $L = 500$, $2N\gamma = 0.1$, $2N\gamma = 0.025$, $\kappa = 1$, and $2N\bar{f} = 50$.

range of evolutionary parameters considered by this study. For example, the population- and genome-wide mutation rate in an *Escherichia coli* population of size $N = 10^5$ is $\mu NL = 250$ (Drake *et al.* 1998). Our simulations cover system sizes up to $\mu NL = 2000$; large populations are simulated by scaling up μ while keeping μNL constant (for details, see section 6 of File S1).

To further test the range of applicability of our model, we evaluate the stochasticity of the fitness flux for different
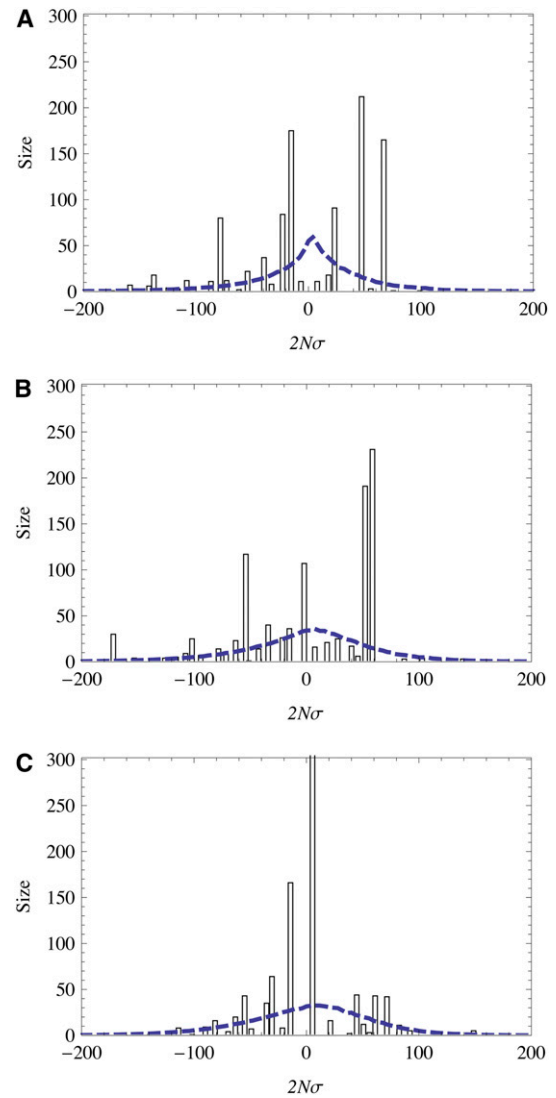

Figure 8 Intermittent fitness waves. The three videos show the distribution of genotype fitness values in a population, as it changes over time in a dense-sweep adaptive process. The mean fitness is kept to 0 by normalization. The fitness distribution is strongly stochastic: at any specific time, its shape is very different from the long-time average shape (plotted as a dashed line). There are recurrent selective sweeps: a new high-fitness genotype appears in the right-hand tail of the distribution, gains frequency, and moves quickly toward the center. Simulations are shown for three different distributions $\rho(f)$ of genomic selection coefficients, characterized by the shape parameter κ. (a) $\kappa = 0.5$ (stretched exponential tail). This case has the strongest stochasticity in the shape of the fitness wave, because exceptionally strong sweeps are generated with significant probability. (b) $\kappa = 1$ (exponential tail). (c) $\kappa = 2$ (Gaussian tail). This case shows a smoother evolution of the fitness wave, which results from average-strength sweeps generated at a more regular pace. Other simulation parameters are $N = 1000$, $L = 1000$, $2N\mu = 0.025$, $2N\gamma = 0.1$, and $2Nf = 50$.
(A) Movie 1: http://www.genetics.org/content/suppl/2011/09/16/genetics.111.132027.DC1/Movie1.mov
(B) Movie 2: http://www.genetics.org/content/suppl/2011/09/16/genetics.111.132027.DC1/Movie2.mov
(C) Movie 3: http://www.genetics.org/content/suppl/2011/09/16/genetics.111.132027.DC1/Movie3.mov

evolutionary parameters. We define the cumulative fitness flux $\Phi_{\Delta t}(t)$ as the selective effect of allele frequency changes over a short time interval Δt (we use $\Delta t = 20$ generations), and we evaluate the ratio ε of variance and mean of $\Phi_{\Delta t}(t)$ over a population's history (Mustonen and Lässig 2010). In Figure 7C, we plot this ratio as a function of the genome length L for fitness effect distributions $\rho(f)$ of different shapes, which have a stretched exponential ($\kappa = \frac{1}{2}$), an exponential ($\kappa = 1$), or a Gaussian ($\kappa = 2$) tail for large values of f. Movies of the fitness wave are available for all three cases; see Figure 8. As expected, we observe a decrease in stochasticity with increasing shape parameter κ. This is consistent with a crossover to fitness waves with deterministic bulk shape in the limit of a sharp distribution ($\kappa \rightarrow \infty$). However, we do not see any evidence of a crossover to deterministic fitness waves with increasing genome length L: the stochasticity ratio ε stays roughly constant and the fitness wave retains strong shape fluctuations even for the largest values of L. At the same time, our model predictions of the fixation probability $G(\sigma)$ overestimate the simulation results at large L, in particular for strongly beneficial driver mutations. This may indicate a crossover to a new mode of adaptive evolution: selective sweeps are driven cooperatively by multiple beneficial mutations, but the adaptive dynamics remain intermittent. This regime is not yet covered by any analytical scheme, but we expect that our interaction calculus can be extended to more complex multidriver sweeps.

Passenger mutations and the inference of selection

Our model predicts an important consequence of interference interactions: a substantial fraction of the genomic substitutions observed in laboratory or field data of dense-sweep processes are not driver mutations, but moderately beneficial or deleterious passenger mutations fixed by hitchhiking. This fraction increases with increasing population size N or genome length L. Disentangling driver and passenger mutations in the dense-sweep regime poses a challenge for the inference of selection from such data. This problem arises not only for asexual populations, but also for linked genome segments of recombining genomes (Fay 2011). The general rationale of this article applies to recombining populations as well, if we replace the total genome length by the linkage correlation length. Adaptation under linkage confounds the standard population-genetic inference of selection based on the statistics of polymorphisms and substitutions, because it reduces the statistical differences between weakly and strongly selected genomic changes. To develop selection inference methods applicable under strong interference, we have to extend our picture of the genome state to polymorphism spectra. This will be the subject of a future study.

Viability limits of evolution under linkage

To summarize, we have shown that interference can strongly reduce the degree of adaptation of an evolving population and hence its viability. This result is likely to be valid beyond the specifics of our model: in any ongoing adaptive process driven by time-dependent selection, a large reduction in the speed of adaptation due to interference is inextricably linked to a large fitness cost compared to unlinked sites. However, genome states with a large fraction of effectively randomized sites are not a plausible scenario for microbial evolution, where a substantial degree of adaptation is maintained and dysfunctional parts are expected to be pruned from the genome. This suggests that natural populations limit genome degradation by interference. Microbial experiments and genomic analysis may help us to better understand how.

Literature Cited

Andolfatto, P., 2007 Hitchhiking effects of recurrent beneficial amino acid substitutions in the Drosophila melanogaster genome. Genome Res. 17(12): 1755–1762.

Bachtrog, D., and I. Gordo, 2004 Adaptive evolution of asexual populations under Muller's ratchet. Evolution 58(7): 1403–1413.

Barrick, J. E., D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh *et al.*, 2009 Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature 461(7268): 1243–1247.

Barton, N. H., 1995 Linkage and the limits to natural selection. Genetics 140: 821–841.

Barton, N. H., 2000 Genetic hitchhiking. Philos. Trans. R. Soc. Lond. B Biol. Sci. 355(1403): 1553–1562.

Betancourt, A. J., 2009 Genomewide patterns of substitution in adaptively evolving populations of the RNA bacteriophage ms2. Genetics 181: 1535–1544.

Bush, R. M., C. A. Bender, K. Subbarao, N. J. Cox, and W. M. Fitch, 1999 Predicting the evolution of human influenza A. Science 286(5446): 1921–1925.

Charlesworth, B., 1994 The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. 63(3): 213–227.

Charlesworth, B., 1996 Background selection and patterns of genetic diversity in Drosophila melanogaster. Genet. Res. 68(2): 131–149.

Charlesworth, B., M. T. Morgan, and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics 134: 1289–1303.

Comeron, J. M., and M. Kreitman, 2002 Population, evolutionary and genomic consequences of interference selection. Genetics 161: 389–410.

Comeron, J. M., A. Williford, and R. M. Kliman, 2008 The Hill-Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. Heredity 100(1): 19–31.

de Visser, J., C. W. Zeyl, P. J. Gerrish, J. L. Blanchard, and R. E. Lenski, 1999 Diminishing returns from mutation supply rate in asexual populations. Science 283(5400): 404–406.

de Visser, J. A. G. M., and D. E. Rozen, 2006   Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. Genetics 172: 2093–2100.

Desai, M. M., and D. S. Fisher, 2007   Beneficial mutation selection balance and the effect of linkage on positive selection. Genetics 176: 1759–1798.

Desai, M. M., D. S. Fisher, and A. W. Murray, 2007   The speed of evolution and maintenance of variation in asexual populations. Curr. Biol. 17(5): 385–394.

Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow, 1998   Rates of spontaneous mutation. Genetics 148: 1667–1686.

Eyre-Walker, A., and P. D. Keightley, 2007   The distribution of fitness effects of new mutations. Nat. Rev. Genet. 8(8): 610–618.

Fay, J. C., 2011   Weighing the evidence for adaptation at the molecular level. Trends Genet 27: 343–349.

Felsenstein, J., 1974   The evolutionary advantage of recombination. Genetics 78: 737–756.

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

Fogle, C. A., J. L. Nagle, and M. M. Desai, 2008   Clonal interference, multiple mutations and adaptation in large asexual populations. Genetics 180: 2163–2173.

Gerrish, P. J., and R. E. Lenski, 1998   The fate of competing beneficial mutations in an asexual population. Genetica 102–103 (1–6): 127–144.

Gillespie, J., 1984   Molecular evolution over the mutational landscape. Evolution 38(5): 1116–1129.

Gillespie, J. H., 2001   Is the population size of a species relevant to its evolution? Evolution 55(11): 2161–2169.

Haldane, J., 1937   The effect of variation of fitness. Am. Nat. 71 (735): 337–349.

Haldane, J., 1957   The cost of natural selection. J. Genet. 55: 511–524.

Hallatschek, O., 2011   From the cover: the noisy edge of traveling waves. Proc. Natl. Acad. Sci. USA 108(5): 1783–1787.

Hermisson, J., and P. S. Pennings, 2005   Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169: 2335–2352.

Hill, W., and A. Robertson, 1966   The effect of linkage on limits to artificial selection. Genet. Res. 8: 269–294.

Imhof, M., and C. Schlotterer, 2001   Fitness effects of advantageous mutations in evolving Escherichia coli populations. Proc. Natl. Acad. Sci. USA 98(3): 1113–1117.

Kaiser, V. B., and B. Charlesworth, 2009   The effects of deleterious mutations on evolution in non-recombining genomes. Trends Genet. 25(1): 9–12.

Kao, K., and G. Sherlock, 2008   Molecular characterization of clonal interference during adaptive evolution in asexual populations of Saccharomyces cerevisiae. Nat. Genet. 40(12): 1499–1504.

Kassen, R., and T. Bataillon, 2006   Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. Nat. Genet. 38(4): 484–488.

Kim, Y., and W. Stephan, 2000   Joint effects of genetic hitchhiking and background selection on neutral variation. Genetics 155: 1415–1427.

Kim, Y., and W. Stephan, 2003   Selective sweeps in the presence of interference among partially linked loci. Genetics 164: 389–398.

Kimura, M., 1962   On the probability of fixation of mutant genes in a population. Genetics 47: 713–719.

Kinnersley, M. A., W. E. Holben, and F. Rosenzweig, 2009   E unibus plurum: genomic analysis of an experimentally evolved polymorphism in Escherichia coli. PLoS Genet. 5(11): e1000713.

MacLean, R. C., and A. Buckling, 2009   The distribution of fitness effects of beneficial mutations in Pseudomonas aeruginosa. PLoS Genet. 5(3): e1000406.

Martin, G., and T. Lenormand, 2006   A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60(5): 893–907.

McVean, G., and B. Charlesworth, 2000   The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. Genetics 155: 929–944.

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

Muller, H., 1950   Our load of mutations. Am. J. Hum. Genet. 2(2): 111–176.

Mustonen, V., and M. Lässig, 2007   Adaptations to fluctuating selection in Drosophila. Proc. Natl. Acad. Sci. USA 104(7): 2277–2282.

Mustonen, V., and M. Lässig, 2008   Molecular evolution under fitness fluctuations. Phys. Rev. Lett. 100(10): 108101.

Mustonen, V., and M. Lässig, 2009   From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. Trends Genet. 25(3): 111–119.

Mustonen, V., and M. Lässig, 2010   Fitness flux and ubiquity of adaptive evolution. Proc. Natl. Acad. Sci. USA 107(9): 4248–4253.

Neher, R. A., and T. Leitner, 2010   Recombination rate and selection strength in HIV intra-patient evolution. PLoS Comput. Biol. 6(1): e1000660.

Orr, H. A., 2000   The rate of adaptation in asexuals. Genetics 155: 961–968.

Orr, H. A., 2003   The distribution of fitness effects among beneficial mutations. Genetics 163: 1519–1526.

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

Park, S.-C., D. Simon, and J. Krug, 2010   The speed of evolution in large asexual populations. J. Stat. Phys. 138(1–3): 381–410.

Perfeito, L., L. Fernandes, C. Mota, and I. Gordo, 2007   Adaptive mutations in bacteria: high rate and small effects. Science 317 (5839): 813–816.

Rambaut, A., O. G. Pybus, M. I. Nelson, C. Viboud, J. K. Taubenberger *et al.*, 2008   The genomic and epidemiological dynamics of human influenza A virus. Nature 453(7195): 615–619.

Rokyta, D. R., P. Joyce, S. B. Caudle, and H. A. Wichman, 2005   An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. Nat. Genet. 37(4): 441–444.

Rouzine, I. M., J. Wakeley, and J. Coffin, 2003   The solitary wave of asexual evolution. Proc. Natl. Acad. Sci. USA 100(2): 587–592.

Rouzine, I. M., E. Brunet, and C. O. Wilke, 2008   The traveling-wave approach to asexual evolution: Muller's ratchet and speed of adaptation. Theor. Popul. Biol. 73(1): 24–46.

Rozen, D. E., J. A. G. M. de Visser, and P. J. Gerrish, 2002   Fitness effects of fixed beneficial mutations in microbial populations. Curr. Biol. 12(12): 1040–1045.

Silander, O. K., O. Tenaillon, and L. Chao, 2007   Understanding the evolutionary fate of finite populations: the dynamics of mutational effects. PLoS Biol. 5(4): e94.

Smith, J. M., 1971   What use is sex? J. Theor. Biol. 30(2): 319–335.

Smith, J. M., 1976   What determines the rate of evolution? Am. Nat. 110(973): 331–338.

Smith, J. M., and J. Haigh, 1974   The hitch-hiking effect of a favourable gene. Genet. Res. 23(1): 23–35.

Tenaillon, O., O. K. Silander, J.-P. Uzan, and L. Chao, 2007   Quantifying organismal complexity using a population genetic approach. PLoS ONE 2(2): e217.

Waxman, D., 2007   Mean curvature *vs.* normality: a comparison of two approximations of Fisher's geometrical model. Theor. Popul. Biol. 71(1): 30–36.

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

Communicating editor: H. G. Spencer

Appendix

Fixation Probability Under Interference Interactions

Here we compute the conditional fixation probability $G(\sigma, \tau \mid \sigma', \tau')$, of a target mutation with selection coefficient σ and origination time τ, which is subject to an interfering mutation with selection coefficient σ' and origination time τ', for the different cases of Figure 2: interference by deleterious background mutations ($\tau' < \tau$ and $\sigma' < -|\sigma|$, Figure 2, A and B), beneficial background mutations ($\tau' < \tau$ and $\sigma' > |\sigma|$, Figure 2, C and D), and beneficial future mutations ($\tau' > \tau$ and $\sigma' > |\sigma|$, Figure 2, E and F). We neglect the effects of interference mutations weaker than the target mutations ($-|\sigma| < \sigma' < |\sigma|$), which is consistent with the hierarchy approximation. We also neglect the effects of deleterious future mutations, which are small (if a deleterious mutation arises after the target mutation, this deleterious mutation cannot prevent fixation of the target mutation). Without interference, a target mutation of frequency x_0 has a fixation probability $G_0(x_0, \sigma) = (1 - e^{-2N\sigma x_0})/(1 - e^{-2N\sigma})$, which is determined by selection and genetic drift. In contrast, we treat beneficial interfering mutations as destined for fixation; that is, we assume they have overcome genetic drift.

Interference by deleterious background mutations

The diagrams in Figure 2, A and B, describe background selection caused by strongly deleterious alleles originating before the target mutation. Case a occurs with probability $x'Q(x', \sigma)$, where x' is the frequency of the interfering mutation at time τ and $Q(x', \sigma)$ is the probability distribution of this frequency. This case results in likely loss of the target mutation, because the interfering mutation has a selection coefficient stronger in magnitude than the target mutation. Case b occurs with probability $(1 - x')Q(x', \sigma)$, and its dominant effect is to boost the initial frequency x_0 of the target mutation by a factor $1/(1 - x')$ (see section 1 of File S1 for a numerical validation). The resulting conditional fixation probability of the target mutation is

$$G(\sigma, \tau|\sigma', \tau') = \int_0^1 dx' Q(x'; \sigma)(1 - x')G_0\left(\frac{1}{N(1 - x')}, \sigma\right). \tag{A1}$$

Because the interfering mutation is deleterious, its probability distribution $Q(x'; \sigma)$ is dominated by very small frequencies $x' \ll 1$. For $\sigma \geq 0$ and $x' \ll 1$, the fixation probability $G_0(x_0, \sigma)$ is in good approximation linear in its first argument. In that case, the factor $(1 - x')$ cancels out and we recover the unlinked fixation probability $G(\sigma, \tau|\sigma', \tau') = G_0(1/N, \sigma)$. This argument does not apply to deleterious target mutations, but they can be neglected because $G_0(x_0, \sigma)$ is exponentially small for $\sigma < 0$, even including a boost in the initial frequency.

Interference by beneficial background mutations

The diagrams of Figure 2, C and D, describe positive and negative interference by a selective sweep starting at time $\tau' < \tau$. Case d occurs with probability $1 - x'$ and results in likely loss of the target mutation. Case c occurs with probability x' and boosts the initial frequency x_0 of the target mutation by a factor $1/x'$ (see Figure S1 in File S1). Treating the interfering mutation as deterministic, its frequency x' at time τ is given by $x' = x_{\det}(\tau - \tau'; \sigma') = 1/[1 + (N-1)\exp(-\sigma'(\tau - \tau'))]$. Hence, we obtain

$$G(\sigma, \tau|\sigma', \tau') = \int_0^1 dx'\ \delta(x' - x_{\det}(\tau - \tau'; \sigma'))x'G_0\left(\frac{1}{Nx'}, \sigma\right), \tag{A2}$$

where δ is Dirac's delta distribution.

Interference by future beneficial mutations

The diagrams in Figure 2, E and F, describe positive and negative interference by a selective sweep starting at time $\tau' > \tau$. Case e occurs with probability $xG_0(x, \tau' - \tau; x_0, \sigma)$ and results in likely fixation of the target mutation by hitchhiking. Here, $G_0(x, \tau' - \tau; x_0, \sigma)$ is the probability that a target mutation of initial frequency x_0 at time τ has reached frequency x at time τ' *without* interference in between. Case f occurs with probability $(1 - x)G_0(x, \tau' - \tau; x_0, \sigma)$ and results in likely loss of the target mutation. Hence, we obtain

$$G(\sigma, \tau|\sigma', \tau') = \int_0^1 dx\ x\ G_0(x, \tau' - \tau; x_0, \sigma)$$
$$= \begin{cases} G_0(x_0, \sigma)/\left(1 + e^{-\hat{\sigma}(\tau' - \tau)}\left(G_0(x_0, \sigma)x_0^{-1} - 1\right)\right) & \text{(for } \sigma > 0), \\ x_0 e^{\hat{\sigma}(\tau' - \tau)} + \left(1 - e^{\hat{\sigma}(\tau' - \tau)}\right)G_0(x_0, \sigma) & \text{(for } \sigma < 0); \end{cases} \tag{A3}$$

see section 2 of File S1 for the evaluation of this integral in the diffusion approximation. The regularized selection coefficient $\hat{\sigma}$ is a shorthand for the crossover from strong to weak selection: $\hat{\sigma} \approx \sigma$ for $N\sigma \gtrsim 1$ and $\hat{\sigma} \simeq 1/2N$ for $N\sigma \lesssim 1$.

Interference by past and future sweeps

First, we compute the conditional fixation probability $G(\sigma, \tau \mid \sigma', \tau', \sigma'', \tau'')$ of a target mutation subject to interference by the closest background sweep (with parameters $\tau' < \tau$ and $\sigma' > \sigma$) and the closest future sweep (with parameters $\tau'' > \tau$ and $\sigma'' > \sigma$). As shown above, the net positive contributions to this probability arise from partial hitchhiking with the past sweep and subsequent full hitchhiking with the future sweep; see Figure 2, C and E. Combining Equations A2 and A3, we obtain

$$G(\sigma, \tau|\sigma', \tau', \sigma'', \tau'') = \int_0^1 dx' \int_0^1 dx\ \delta(x' - x_{\det}(\tau - \tau'; \sigma'))x'xG_0\left(x, \tau'' - \tau; \frac{1}{Nx'}, \sigma\right). \tag{A4}$$

This expression is based on the assumption that the two sweeps act sequentially and independently; that is, the target mutation can fix only if it is free of interference or positively interfered with by both sweeps. Interactions

between the sweeps themselves are neglected (such as rescue of the target mutation by a future sweep, following negative interference by a past sweep). This is in tune with our self-consistent determination of the sweep rate, which absorbs the overlap exclusion between driver mutations [*i.e.*, the condition $\tau'' - \tau' > \tau_{\mathrm{fix}}(\sigma')$] into a reduced uniform or "mean-field" rate $V_{\mathrm{drive}}(\sigma)$ given by Equations 11 and 12. In this approximation, a target mutation of selection coefficient σ is subject to interference by stronger selective sweeps at a total rate $V_>(\sigma) = \int_\sigma^\infty d\sigma' V_{\mathrm{drive}}(\sigma')$. We can now integrate Equation A4 over past and future sweeps (*i.e.*, driver mutations) with an exponential distribution of waiting times $\tau - \tau'$ and $\tau'' - \tau$,

$$G(\sigma) = \int_{-\infty}^{\tau} d\tau' \int_{\tau}^{\infty} d\tau'' \int_{\sigma}^{\infty} d\sigma' \int_{\sigma}^{\infty} d\sigma'' V_{\mathrm{drive}}(\sigma') V_{\mathrm{drive}}(\sigma'') e^{-V_>(\sigma)(\tau''-\tau')} G(\sigma, \tau | \sigma', \tau', \sigma'', \tau'').$$

(A5)

Using Equation A3, the integrations over σ'' and τ'' can be treated analytically, and we obtain

$$G(\sigma) = \int_{-\infty}^{\tau} d\tau' \int_{|\sigma|}^{\infty} d\sigma' V_{\mathrm{drive}}(\sigma') e^{-V_>(|\sigma|)(\tau-\tau')} \int_0^1 dx' \delta(x' - x_{\mathrm{det}}(\tau - \tau'; \sigma'))$$
$$\times \begin{cases} x' G_0\left(\frac{1}{Nx'}, \sigma\right) {}_2F_1\left[1, \; V_>(\sigma)\hat{\sigma}, \; 1 + V_>(\sigma)\hat{\sigma}; \; 1 - Nx' G_0\left(\frac{1}{Nx'}, \sigma\right)\right] & (\text{for } \sigma > 0), \\ \frac{1}{N(|\hat{\sigma}| + V_>(|\hat{\sigma}|))} \left(Nx' G_0\left(\frac{1}{Nx'}, \sigma\right) |\hat{\sigma}| + V_>(|\sigma|)\right) & (\text{for } \sigma < 0), \end{cases}$$

(A6)

where ${}_2F_1(a, b, c; z)$ is a hypergeometric function. The remaining integrals in this expression can be evaluated numerically in a straightforward way (see section 4 of File S1 for an iterative procedure). Neglecting the (numerically smaller) integral over past sweeps, we obtain a closed form of Equation A6,

$$G(\sigma) = \begin{cases} G_0\left(\frac{1}{N}, \sigma\right) {}_2F_1\left[1, V_>(\sigma)\hat{\sigma}, \; 1 + V_>(\sigma)\hat{\sigma}; \; 1 - NG_0\left(\frac{1}{N}, \sigma\right)\right] & (\text{for } \sigma > 0), \\ \frac{1}{N(|\hat{\sigma}| + V_>(|\hat{\sigma}|))} \left(NG_0\left(\frac{1}{N}, \sigma\right) |\hat{\sigma}| + V_>(|\sigma|)\right) & (\text{for } \sigma < 0). \end{cases}$$

(A7)

These results also determine the fixation probability of passenger mutations in the decomposition of Equation 13,

$$G_{\mathrm{pass}}(\sigma) = \frac{G(\sigma) - p_{\mathrm{drive}}(\sigma) G_0(\sigma)}{1 - p_{\mathrm{drive}}(\sigma)},$$

(A8)

where $P_{\mathrm{drive}}(\sigma)$ is given by Equation 12 and $G_0(\sigma) = (1 - \exp[-2\sigma])/(1 - \exp[-2N\sigma])$. The neutrality threshold $\tilde{\sigma}$ is obtained by Taylor expansion of Equation A7; we obtain $G(\sigma) = (1/N) + \sigma/[1 + 2NV_>(\sigma)] + \mathcal{O}(\sigma^2)$. Comparing with the corresponding expansion of $G_0(\sigma)$ shows that linkage leads to a linear reduction of the strength of selection by a factor $1 + 2NV_>(\sigma) \approx 1 + 2NV_>(0)$ compared to unlinked sites, as expressed in Equations 14 and 15.

GENETICS

Emergent Neutrality in Adaptive Asexual Evolution

Stephan Schiffels, Gergely J. Szöllősi, Ville Mustonen, and Michael Lässig

File S1

Supporting Text

1   Approximation for pairwise interaction diagrams

In the two diagrams (b) and (c) of Figure 2, we made the approximation that the advantage for the target mutation can be expressed as an increase in the initial frequency. In case of diagrams, (b), the factor of increase is given by $1/(1 - x')$, where x' is the frequency of the deleterious background mutation at time τ, while in diagram (c) the factor of increase is given by $1/x'$, where x' is the frequency of the beneficial driver substitution at time τ. We simulated this situation with a two-locus model in which the beneficial background allele is introduced with frequency x' and the target mutation with frequency $x_0 = 0.01$ *within* the subpopulation carrying the background allele. We then simulate a Wright-Fisher process and measure the fixation probability of the target mutation.

In Figure S1 we show results of a numerical simulation for the case of diagram (c). The validation for diagram (b) can then be read by substituting $x' \to 1 - x'$ in the x-axis. As can be seen, the model is accurate as long as the increase factor $1/x'$ is not too large. For values of $x' < 0.4$, the fixation probability is overestimated. In case of diagram (b) this condition is almost always fulfilled, as deleterious mutants have very small frequencies, i.e. $1 - x' \gg 0.4$. In case of diagram (c) the approximation overestimates the fixation probability if the driver substitution appeared shortly before the target mutation and hence still has small frequency.

In a previously described model [7] the fixation probability in an expanding subpopulation is computed explicitly, which yields a result that is comparable to our approach (see their equation 11). Their result is more accurate than the expression we provide above (expanding the initial frequency in Kimura's formula). In our case, modeling the effect of the expansion as an increase in the initial frequency is accurate enough, in particular since the dominant interference effect is provided by future interfering mutations and not by background mutations.

The effects of background selection have been subject to a large number of articles. Often, these studies find that background selection in fact retains a substantial net effect on the fixation probability of a target mutation [10, 1, 6]. These studies typically assume a mutation-selection balance of many deleterious mutations, with a constant deterministic influx of deleterious mutations. For this case, an argument from Fisher [3] shows that the fixation probability of a beneficial mutation is reduced (see [1]) by a factor $\exp(-U_\mathrm{d}/\sigma_\mathrm{d})$, where U_d is the rate and σ_d is the selection coefficient of deleterious mutations.

In our model, Fisher's argument does not hold, for two reasons: i) because of the presence of adaptive substitutions (selective sweeps), variance in the population is constantly removed, hence a stationary mutation-selection balance is never maintained; ii) because we consider an exponential distribution of selection coefficients, the number of deleterious mutations that are *stronger* in effect than the target mutation, are typically rare enough to be treated stochastically, as done in our pairwise interaction scheme using the diagrams of Figure 2.

2   Single site propagator

In the derivation of the fixation probability (see *Appendix B*, we use the time evolution of the mean allele frequency $M(t, x_0, \sigma) = \int_0^1 dx\, x\, G_0(x, t, x_0, \sigma)$ for which we can derive an analytical expression using the
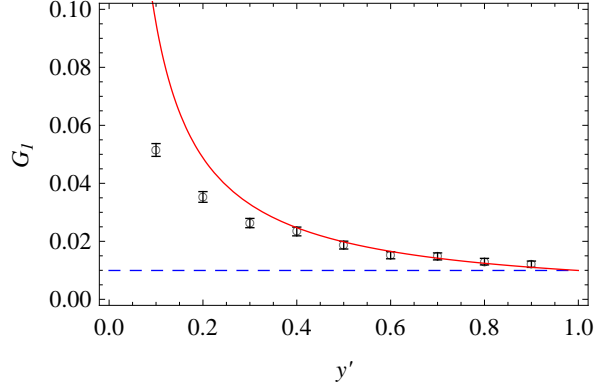
Figure S1: We plot the fixation probability of a beneficial target mutation ($2N\sigma = 10$) on the background of a strongly selected mutation ($2N\sigma' = 100$) initially present with frequency x'. The target mutation has initial frequency $x_0 = 0.01\%$. The black circles are simulation results, the blue dashed line is the expected fixation probability without background selection, the red line is the theory prediction by $G_0(x_0/x', \sigma)$. The population size is $N = 1000$.

diffusion approximation. The Fokker-Planck equation for a single site under drift and selection, using the selection coefficient σ and the population size N reads:

$$\partial_t G = \left[\frac{1}{2N}\partial_x^2(x(1-x)) - \sigma\partial_x(x(1-x))\right] G. \tag{S1}$$

To derive $M(t, x_0, \sigma)$, we multiply equation (S1) with x and integrate by parts, neglecting boundary terms:

$$\partial_t \int_0^1 \mathrm{d}x\, x\, G = \sigma \int_0^1 \mathrm{d}x\, x(1-x) G \tag{S2}$$

or

$$\partial_t M = \sigma(M - \int_0^1 \mathrm{d}x\, x^2 G). \tag{S3}$$

We introduce $M_2 = \sigma(\int_0^1 \mathrm{d}x\, x^2 G - M^2)$ to write

$$\partial_t M = \sigma M(1 - M) - M_2. \tag{S4}$$

We identify two limits of the solution of this equation. First, for $t = 0$, the variance term M_2 must vanish and we see that M evolves logistically with initial value x_0. Secondly, if $\mu N \ll 1$, for times larger than the fixation time $\tau(\sigma) = (2\ln\sigma)/\sigma$ (see chapter *Fixation Time*), G becomes stationary and will consist of two delta peaks at 0 and 1 with weights reflecting the fixation probability G_0:

$$G(x, t, x_0, \sigma) \xrightarrow{t > \tau(\sigma)} (1 - G_0)\delta(x) + G_0\delta(1 - x), \tag{S5}$$

and

$$M_2 \xrightarrow{t > \tau(\sigma)} G_0(1 - G_0). \tag{S6}$$

For the stationary mean allele frequency we get

$$\partial_t M_{\mathrm{stat}} = 0 \quad \Rightarrow \quad M_{\mathrm{stat}} = G_0. \tag{S7}$$

Motivated by the form of S4 and the known limit S7, we can try a logistic ansatz for beneficial and an exponential for deleterious mutations:

$$M_+(t, x_0, \sigma) = \frac{G_0(x_0, \sigma)}{1 + e^{-\hat\sigma t}(G_0(x_0, \sigma)x_0^{-1} - 1)} \tag{S8}$$
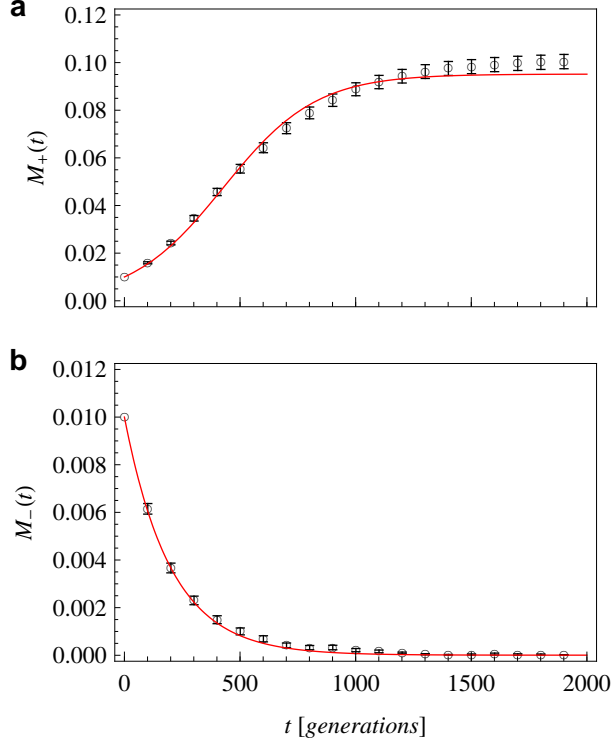
Figure S2: This plot shows the mean frequency of a beneficial (a) and deleterious (b) mutation as a function of time, measured in generations. Black circles are data obtained from simulations, red solid lines are the theory predictions of equation S8 and S9. Simulation data has been obtained from many trajectories started with different random seeds. Error bars indicate the standard error of the mean. Parameters: $N = 1000$, $2N\sigma = 10$, $x_0 = 10/N$.

and

$$M_-(t, x_0, \sigma) = x_0 e^{-\hat{\sigma}t} + (1 - e^{-\hat{\sigma}t})G_0(x_0, -\sigma) \tag{S9}$$

with the standard fixation probability

$$G_0(x_0, \sigma) = \frac{1 - e^{-2N\sigma x_0}}{1 - e^{-2N\sigma}} \tag{S10}$$

and the regularized selection coefficient $\hat{\sigma}$ that has the two limits $\hat{\sigma} = 1/2N$ for $N\sigma \ll 1$ and $\hat{\sigma} = \sigma$ for $N\sigma \gg 1$. The exact form of this crossover is not important. Here we choose

$$\hat{\sigma} = \begin{cases} 1/2N & \text{for } N\sigma \leq 1 \\ \sigma & \text{for } N\sigma > 1 \end{cases} \tag{S11}$$

The predictions from equations S8 and S9 agree with simulations, as shown in Figure S2.

3   Fixation Times

The deterministic equation for the frequency of a mutation under selection σ is

$$x(t) = \frac{1}{1 + \exp(-\sigma'(\tau - \tau'))(x_0^{-1} - 1)} \tag{S12}$$

S. Schiffels *et al.*

Since evolution is always stochastic at the boundaries, we define the fixation time as the time it takes to reach a frequency $1 - 1/(2N\sigma)$, starting from a frequency $x_0 = 1/(2N\sigma)$. This leads to

$$\tau_{\text{fix}}(\sigma) = \frac{2 \log(2N\sigma)}{\sigma}. \tag{S13}$$

For small values of σ, the dynamics is not anymore deterministic and the above expression becomes wrong as can be seen by the limit value for neutral alleles, where we require $\tau_{\text{fix}} = 2N$. For values of $2N\sigma < 3$, we simply set

$$\tau_{\text{fix}}(\sigma) = 2N \quad 2N\sigma < 3. \tag{S14}$$

We typically consider average selection coefficients of $2N\bar{f} \gg 1$, and the average driver mutation has even larger selection coefficients. The exact setting of the above threshold is therefore not important.

4 Iterative Solution

As discussed in the main text, the following set of equations has to be solved:

$$p_{\text{drive}}(\sigma) = \exp(-\tau_{\text{fix}}(\sigma)V_>(\sigma)), \tag{S15}$$

$$V_{\text{drive}}(\sigma) = U(\sigma)G(\sigma)p_{\text{drive}}(\sigma), \tag{S16}$$

$$V_>(\sigma) = \int_\sigma^\infty d\sigma' \, V_{\text{drive}}(\sigma'), \tag{S17}$$

$$U(\sigma) = LN\mu\lambda_{\text{d}}(\sigma)\rho(\sigma), \tag{S18}$$

$$\lambda_{\text{d}}(f) = \frac{G(-f) + \gamma/(\mu N)}{G(f) + G(-f) + 2\gamma/(\mu N)}, \tag{S19}$$

$$\begin{aligned} G(\sigma) = \int_{-\infty}^\tau d\tau' \int_\sigma^\infty d\sigma' \, V_{\text{drive}}(\sigma')e^{-V_>(\sigma)(\tau - \tau')} \int_0^1 dx' \, \delta(x' - x_{\text{det}}(\tau - \tau'; \sigma')) \\ \times \, x'G_0\left(\frac{1}{Nx'}, \sigma\right){}_2F_1\left[1, \frac{V_>(\sigma)}{\sigma}, 1 + \frac{V_>(\sigma)}{\sigma}; 1 - Nx'G_0\left(\frac{1}{Nx'}, \sigma\right)\right], \quad \sigma > 0 \end{aligned} \tag{S20}$$

and

$$\begin{aligned} G(\sigma) = \int_{-\infty}^\tau d\tau' \int_{|\sigma|}^\infty d\sigma' \, V_{\text{drive}}(\sigma')e^{-V_>(|\sigma|)(\tau - \tau')} \int_0^1 dx' \, \delta(x' - x_{\text{det}}(\tau - \tau'; \sigma')) \\ \times \, \frac{1}{N(|\sigma| + V_>(|\sigma|))}\left(Nx'G_0\left(\frac{1}{Nx'}, \sigma\right)|\sigma| + V_>(|\sigma|)\right), \quad \sigma < 0, \end{aligned} \tag{S21}$$

where ${}_2F_1(a, b, c; z)$ is a hypergeometric function.

The selfconsistent solution $G(\sigma)$ of equations S16, S18, S19, S17, S20 and S21 is found by the following numerical procedure: We initialize the iteration by setting $V_{\text{drive}}(\sigma) \equiv 0$. We then iterate the following steps:

1. Use equation S17 to compute the cumulative rate of drivers $V_>(\sigma)$.

2. Use $V_>(\sigma)$ to compute the fixation probabilities $G(\pm\sigma)$ from equations S20 and S21.

3. Use $G(\pm\sigma)$ to compute the stationary state $\lambda_{\text{d}}(f)$ from equation S19 and hence the mutation rate $U(\sigma)$ and the driver rate $V_{\text{drive}}(\sigma)$ from equations S18 and S16.

Steps 1. and 2. involve standard numerical integration methods, such as the function *NIntegrate* in Wolfram Research's *Mathematica*. To speed up the iterations, we evaluated all of the involved functions at 40 equidistant discrete values of σ between 0 and $\sigma_{\max} = 10\bar{f}$ and linear interpolations between the evaluated points. We used 10 iterations and observe a quick convergence after six iterations of the above algorithm.

5   Comparison with Gerrish-Lenski theory

Gerrish and Lenski (GL) [4] computed the fixation probability of a beneficial mutation in a manner similar to how we compute the rate of driver mutations. In GL-theory, the fixation probability is

$$G_{\mathrm{GL}}(\sigma) = U(\sigma)G_0(\sigma)e^{-\frac{1}{2}\tau(\sigma)\int_\sigma^\infty U(\sigma)G_0(\sigma)}. \tag{S22}$$

The factor $1/2$ in the exponent follows from counting only those stronger mutations that appear on the background *not carrying* the target mutation. Only these mutations decrease the fixation probability of the target mutation. Equation S22 is analogous to equation S16, in which we require that a driver mutation is not interfered with by any stronger driver mutation. There are, however, two important differences: i) Equation S16 has no factor $1/2$, since we exclude from the driver rate those mutations that fix by positive interference (hitchhiking). In contrast, in GL-theory, the only mode of fixation are driver mutations that are not suffering negative interference. ii) Equation S16 reflects a self-consistent closure, in which each interfering driver must itself be free from even stronger interfering drivers. Therefore, apart from the factor $1/2$, equation S22 can be seen as a first iteration loop of the rate of driver mutations. Taking into account hitchhiking as a positive outcome of interference dramatically enhances the fixation probability of weakly beneficial mutations and in particular allows us to compute the influence on deleterious alleles, a case which is not covered by GL-theory.

Our genomic model yields beneficial mutation rates and their distribution as an outcome, while GL-theory takes distribution and rate of beneficial mutations as an input. In GL-theory all stronger driver mutations in the exponent of equation S22 are assumed to be free of interference. In our model this corresponds to assuming the unlinked rate of beneficial mutations, using results from single site theory, to be

$$U_0(\sigma) = \lambda_{0,\mathrm{d}}(f)\rho(f) \tag{S23}$$

and

$$\lambda_{0,\mathrm{d}}(f) = \frac{G_0(-f) + \gamma/(\mu N)}{G_0(f) + G_0(-f) + 2\gamma/(\mu N)}. \tag{S24}$$

A comparison with GL-theory is shown in Figures 3a, 3c and S3b as a brown curve.

6   Computer simulations

We simulate the Wright Fisher model with N individuals (N fixed throughout). Each individual consists of a sequence of L alleles, denoted by bits $a_{ij} = \{1,0\}$ with $i = 1\ldots N$ and $j = 1\ldots L$. At every generation, alleles at any individual can mutate from 1 to 0 and vice versa with probability μ. The fitness F_i of individual i is obtained by a sequence of additive selection coefficients f_j multiplied by the time-dependent direction of selection $\eta_j(t) = \{-1, 1\}$:

$$F_i = \sum_{j=1}^{L} a_{ij}\eta_j(t)f_j. \tag{S25}$$

The local selection coefficients $f_j \geq 0$ are kept constant throughout evolution and fixed at values:

$$f_j = \frac{\bar{f}}{\Gamma(1+1/\kappa)} \left(-\log\left(1 - \frac{j-1}{L} \right) \right)^{1/\kappa} \tag{S26}$$

The numbers f_j reflect an ordered set of random variates from the Weibull distribution (as suggested in [9]):

$$\rho(f) = \frac{\kappa}{\zeta} \left(\frac{f}{\zeta} \right)^{\kappa-1} e^{-(f/\zeta)^\kappa} \tag{S27}$$

with $\zeta = \bar{f}/\Gamma(1+1/\kappa)$ such that the mean of this distribution is $\bar{f}$. For $\kappa = 1$ this distribution is exponential.

S. Schiffels *et al.*

The direction of selection $\eta_j(t)$ is a stochastic random variable that flips its sign on average every $1/\gamma$ generations with the autocorrelation

$$\langle \eta_j(t)\eta_k(t') \rangle = e^{-\gamma|t-t'|}\delta_{jk}. \tag{S28}$$

A flip of direction affects the fitnesses of all individuals in the population. Each generation is stochastically sampled from the previous generation, using a multinomial sampling process, in which the probability p_i of picking individual i is given by:

$$p_i = e^{F_i - \langle F \rangle}, \tag{S29}$$

with the mean fitness

$$\langle F \rangle = \frac{1}{N}\sum_{i=1}^{N} F_i. \tag{S30}$$

A substitution is observed if a site that was monomorphic at some allele becomes polymorphic and then fixed at the other allele. Each substitution can be categorized as "beneficial", if the new allele is the currently fitter, or "deleterious" if it is the currently less fit of the two alleles. We also keep track of the fixation state: Over a sufficiently large number of generations, we can compute λ_j of site j as the fraction of time at which the frequency of the currently fitter allele was below 0.5. Since we generally consider low local mutation rates $\mu N \ll 1$, this fraction yields approximately the fraction of time the population was *fixed* at the locally less fit allele. At every site, the origination rate of new beneficial mutations can be computed as $\mu\lambda$ and of new deleterious mutations as $\mu(1 - \lambda)$. Fixation probabilities can be computed by dividing the rate of beneficial/deleterious substitutions by the rate of beneficial/deleterious originations.

For the stationary adaptation simulations, we initialize the program with a monomorphic population where all alleles are 0. In each run of the simulation, we let the population equilibrate for $\min(1/\mu, 1/\gamma)$ generations before initiating any measurement to ensure that the population is in the stationary state.

In case of the approach to equilibrium, we first run the above protocol for stationary adaptation (for some parameter γ as given in the particular presentation of the results) for sufficiently long time to ensure stationarity. Then we set the flip rate $\gamma = 0$ and start obtaining measurements as described above. All results are now time-dependent. To obtain averages, we therefore have to repeat this program many times and average over the full ensemble of simulations for each time point.

The theoretical derivations for our model suggest some simple scaling laws. As can be seen, the population size N and the genome length L are only relevant in the parameter combinations $N\mu L$, $N\gamma L$ and $N\bar{f}$. As long as we keep these parameter combinations fixed, we can use smaller values for N and L to speed up the simulations. This scaling holds up to the following conditions: i) $\mu L \ll 1$ to avoid two mutations in the same individual in the same generation, ii) $\mu N \ll 1$ so that sites follow substitution dynamics with short polymorphic times and iii) $N\gamma \ll 1$ so that the time between selection flips is larger than the time needed for a fixation. These conditions can always be fulfilled for given parameters $N\mu L$, $N\gamma L$ and $N\bar{f}$.

7 Supplementary Results

7.1 Mutation and Fixation rate

As seen in equation S18, the rate of beneficial and deleterious mutations in our model depends on the dynamics. In Figure S3a) we plot the rate of mutations compared to the expectation without interference from linkage. The rate of beneficial mutations is enhanced, while the rate of deleterious mutations is decreased in comparison to single site theory. This increase is reflecting the fact that the population under linkage is less adapted to its environment and many sites in the genome are not fixed at the locally fitter allele. Hence mutations at these maladapted sites emit more beneficial mutations. It turns out that for beneficial mutations, $\sigma > 0$, the distribution $U(\sigma)$ is exponential. In Figure S3b) we show the rate of fixations, which is a product of the mutation rate and the fixation probability. The comparison with GL-theory emphasizes the point made in section 5: the rate of weakly beneficial fixations is underestimated due to neglecting

hitchhiking. Figure S3c) shows the theory prediction of the substitution rate (see S3b) together with its partitioning in passengers (gray shading) and drivers (green shading).

7.2 Positive and negative fitness flux

Since in our model we explicitly include the dynamics of deleterious mutations, we can disentangle the fitness flux into a positive fitness flux, constituted by the beneficial mutations, and a negative part, caused by the fixation of deleterious mutations:

$$\Phi(f) = \Phi_+(f) - \Phi_-(f) = fV(f) - fV(-f). \tag{S31}$$

In figure S4 we show this decomposition. As can be seen, the two terms of the fitness flux have their main contributions coming from different parts of the spectrum of selection coefficients: While the positive flux is mainly carried by strongly beneficial mutations, the negative flux consists of weaker deleterious fixations. In total, the positive flux is always much larger than the negative one, but the decomposition reveals an interesting pattern: Remarkably, even very strongly selected genomic sites provide a significant contribution to the negative flux, reflecting ubiquitous (strongly) deleterious passenger mutations.

7.3 Relation to mutation based models

The self-consistency of genomic state and mutations is one feature that distinguishes our model from most previous studies of adaptation under linkage [4, 12, 2, 8]. These mutation-based models constrain the distribution of selection coefficients for beneficial mutations, $u(\sigma) = (1/U_b)U(\sigma)$, to a fixed shape and use the total rate of beneficial mutations, U_b, and their mean effect, $\sigma_b = \int_0^\infty \sigma u(\sigma)d\sigma$, as independent input parameters. This is a suitable setup to evaluate the speed of adaptation at stationarity, because Φ depends in good approximation only on the distribution of beneficial mutations (see Figure S5a). However, mutation-based approaches of this type cannot predict genomic quantities such as the average degree of adaptation, α, which arguably is the most appropriate measure of the efficiency of the adaptive process over long periods of time. Even at stationarity, the average degree of adaptation is not uniquely determined by U_b and σ_b, but depends on all three genomic parameters $\bar{f}$, γ, and L in a nontrivial way (see Figure S5b). In our model, the rates $U(\sigma)$ of beneficial mutations (and, hence, U_b and σ_b) are dependent quantities, which must be derived from the self-consistent solution of the genome dynamics described in the main text. Changing any of the genomic parameters, say L, will change U_b and σ_b, so that these parameters are not suitable as input if we want to evaluate the dependence of the model on L. Similarly, U_b and σ_b change with time in a non-stationary adaptive process.

7.4 Epistasis

In the model presented in the manuscript, we use a strictly additive fitness model without any epistatic interactions between genomic sites. To check how epistasis affects our results, we employ a simple extension to our model that explicitly realizes pairwise epistatic interactions. This extension is done in the spirit of the study presented in reference [5]. In addition to the strictly additive contribution to an individuals fitness S25, we add another term for the pairwise interactions:

$$F_i = \sum_{j=1}^{L} a_{ij}\eta_j(t)f_j + e\sum_{j<k}(2a_{ij} - 1)(2a_{ik} - 1)f_{jk}, \tag{S32}$$

with the matrix f_{jk} describing the pairwise epistatic interactions between sites. In the simplest case realized here, we use normal distributed random numbers with mean 0 and standard deviation 1, fixed throughout evolution. The impact of the epistatic interactions can be estimated by comparing the magnitude of the additive fitness effects with the magnitude of the epistatic effects. As can be seen from equation S32, the
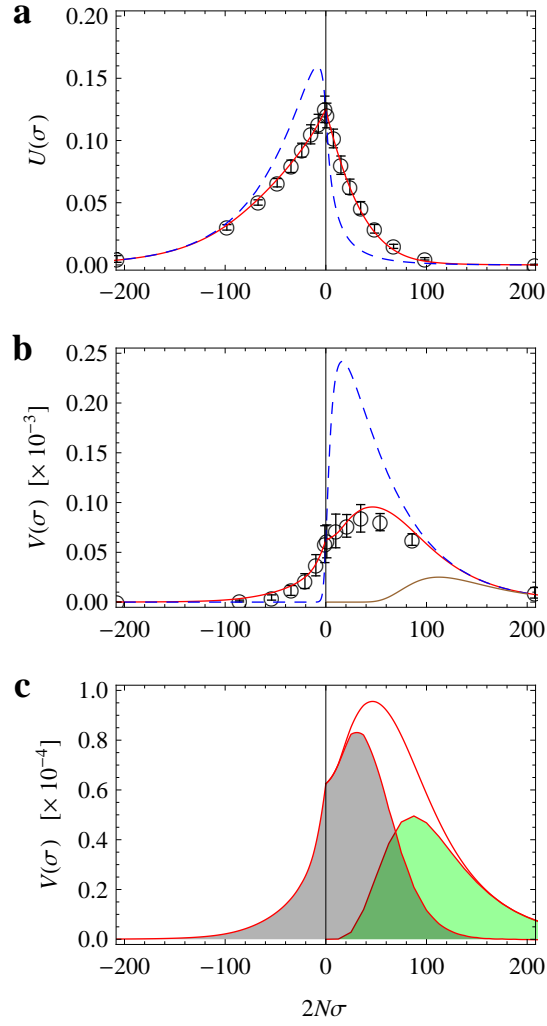
Figure S3: a) rate of mutations, b) rate of substitutions, c) theory prediction of substitutions with partitioning into passengers (gray) and drivers (green). Black circles: Simulation data, blue line: single site theory, red line: self-consistent theory, brown line: GL-theory. Parameters are $N = 2000$, $L = 1000$, $2N\gamma = 0.1$, $2N\mu = 0.025$, $2N\bar{f} = 50$.
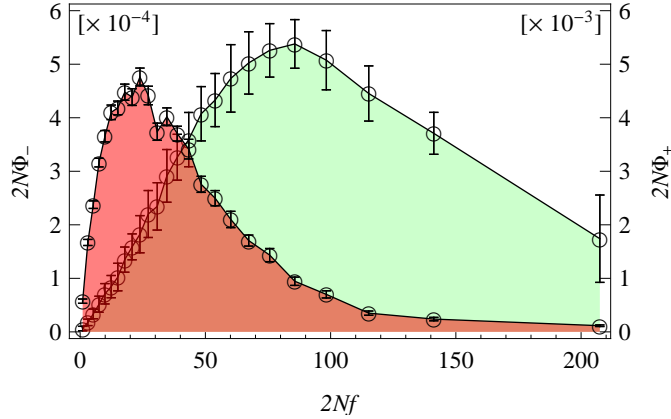
Figure S4: In green: Positive fitness flux, In red: negative fitness flux. For clarity, the negative fitness flux is shown with a ten-fold amplification to give a more direct comparison. Black circles: Simulation data ($N = 2000$, $L = 1000$, $2N\gamma = 0.1$, $2N\mu = 0.025$, $2N\bar{f} = 50$).

epistatic term scales as L^2 (double sum), whereas the additive term scales as L. We therefore have

$$e \sim \frac{\bar{f}}{L} \tag{S33}$$

as the typical crossover at which epistatic interactions dominate the additive fitness.

Figure S6 shows the degree of adaptation as a function of the site selection coefficient f, for both the strictly additive model (blue curve) and the epistatic model at the crossover, i.e. $e = \bar{f}/L$. Clearly, although the epistatic contribution to the fitness function is of the same magnitude as the additive part, the degree of adaptation is very robust under epistasis. Figure S7 shows the total degree of adaptation for varying values of the scaled epistasis parameter $\hat{e} = e(L-1)/\bar{f}$. The predicted crossover is given by $\hat{e} = 1$. In summary, epistasis does not quantitatively change our model and results, as long the epistatic interactions are of the same magnitude as the additive component of the fitness.

7.5 Mutation rate during approach to equilibrium

In the main text we show the degree of adaptation during an approach to equilibrium. Here we show additionally how the mutation rate itself depends on time in such a scenario. Figure S8 shows the distribution of fitness effects of new mutations at three different time points. Initially, the population is poorly adapted, so there are many beneficial mutations available. As the population approaches equilibrium, more mutations become deleterious, as seen in the plot. We use the same simulation protocol as described in the main text: the initial population has been evolved in stationarity with $2N\gamma = 0.1$. To observe the approach to equilibrium, we set $\gamma = 0$.

7.6 Non-exponential distributions of selection coefficients

In addition to the exponential distribution results shown in the main text, we present two additional cases: For $\kappa = 1/2$ the tail of the distribution is a stretched exponential with broadly distributed selection coefficients. For $\kappa = 2$ we recover a gaussian tail and hence more sharply distributed selection coefficients. In Figure S9 we show simulation results with theory predictions for the three distribution shapes $\kappa = 1/2, 1, 2$. While the case $\kappa = 1$ is predicted most accurately by our theory, our prediction slightly deviates from the simulations using the two non-exponential distributions.
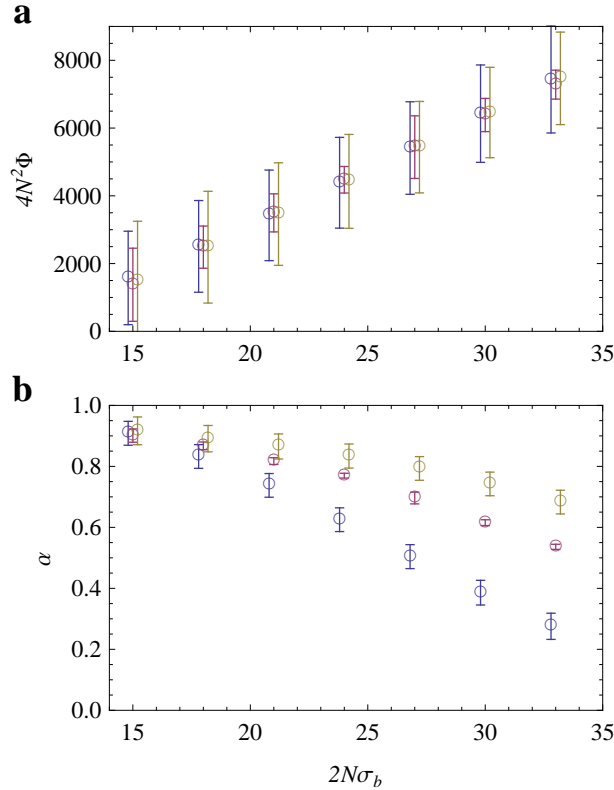
S. Schiffels *et al.*

Figure S5: This plot shows a) the scaled fitness flux $4N^2\Phi$ and b) the degree of adaptation α as a function of σ_b at *fixed $U_b = 30$*. The data points have been obtained by interpolating simulation results with varying γ, L and μ with three different fixed values of $2N\bar{f} = 40, 50, 60$ in blue, red and yellow. Error bars are predicted by the interpolation method (see Numerical Recipes 3rd edition [11] "Kriging"). In a) circles have been shifted horizontally by $\epsilon = \pm 0.2$ to make distinction between the data points possible.

References

[1] Nick H Barton. Linkage and the limits to natural selection. *Genetics*, 140(2):821–41, Jun 1995.

[2] Michael M Desai, Daniel S Fisher, and Andrew W Murray. The speed of evolution and maintenance of variation in asexual populations. *Curr Biol*, 17(5):385–94, Mar 2007.

[3] RA Fisher. The genetical theory of natural selection. *Clarendon Press, Oxford*, 1930.

[4] P J Gerrish and Richard E Lenski. The fate of competing beneficial mutations in an asexual population. *Genetica*, 102-103(1-6):127–44, Jan 1998.

[5] Richard A Neher and Boris I Shraiman. Competition between recombination and epistasis can cause a transition from allele to genotype selection. *PNAS*, 106(16):6866–71, Apr 2009.

[6] H Allen Orr. The rate of adaptation in asexuals. *Genetics*, 155(2):961–8, Jun 2000.

[7] S P Otto and M C Whitlock. The probability of fixation in populations of changing size. *Genetics*, 146(2):723–33, Jun 1997.
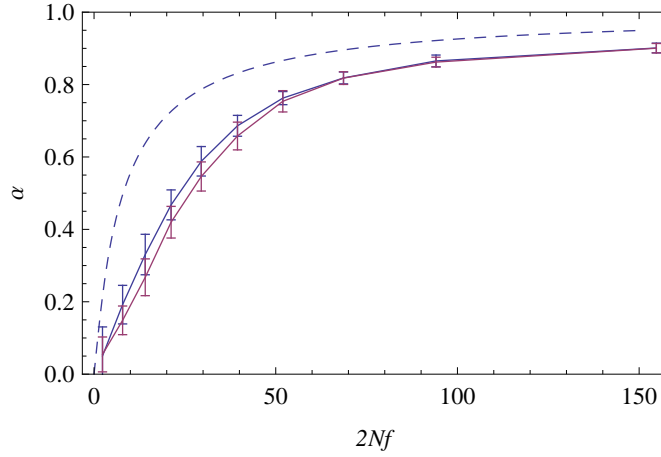
Figure S6: This plot shows the degree of adaptation as a function of the local selection coefficient f_j. The blue curve is the result from a simulation with no epistasis, $e = 0$, while the red curve was simulated with the epistatic fitness component being equal to the additive component: $e = \bar{f}/(L-1)$. Other parameters are $N = 1000$, $L = 1000$, $2N\gamma = 1$, $2N\mu = 0.025$, $2N\bar{f} = 50$.
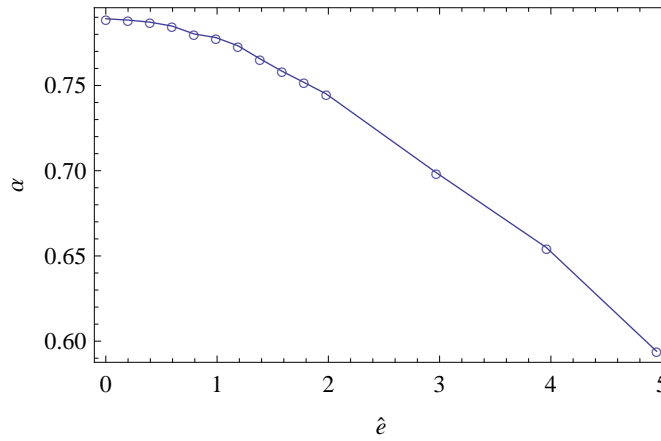

Figure S7: Here we show the total degree of adaptation with the same parameters as in Figure S6, with varying values of scaled epistasis $\hat{e}(L-1)/\bar{f}$. For $\hat{e} > 1$ the epistatic component of the fitness is larger than the additive component.
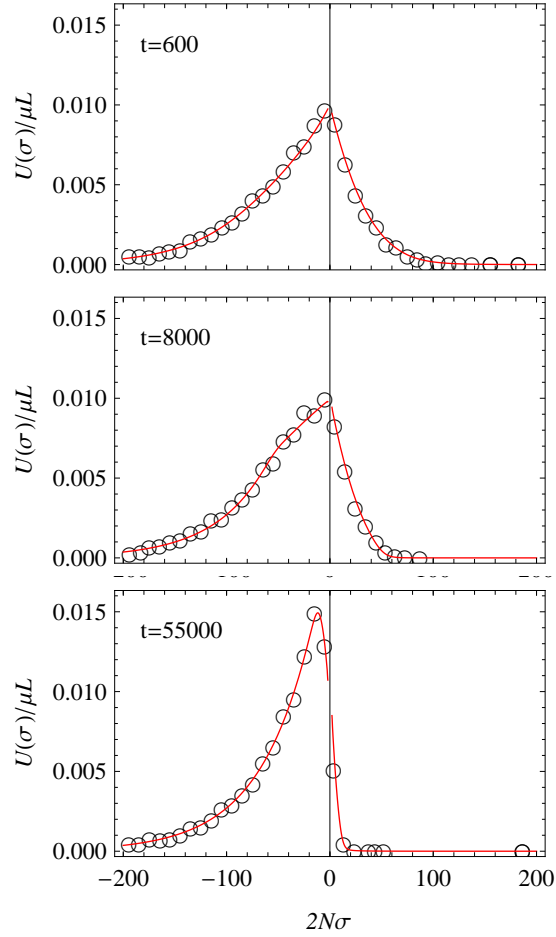
S. Schiffels *et al.*

Figure S8: This Figure shows simulation and theory results for an approach to equilibrium, starting from an initially poorly adapted state (see main Text). We plot the distribution of new mutations for three different time points. Simulation results are shown in circles, theory predictions from numerically solving equation 5 in the main text are shown in red. Parameters are $N = 1000$, $L = 1000$, $2N\bar{f} = 50$, $2N\mu = 0.025$. The initial state is a stationary state with $2N\gamma = 0.1$.
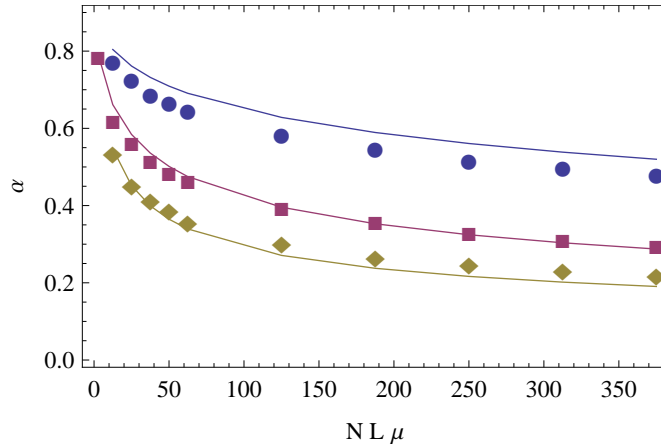
Figure S9: This plot shows simulation results for the degree of adaptation as a function of the total mutation rate. We show results for three values of $\kappa = 0.5$ (Circles), $\kappa = 1$ (Rectangles) and $\kappa = 2$ (Diamonds). Other parameters are $N = 4000$, $2N\bar{f} = 50$, $2N\gamma = 0.1$, $2N\mu = 0.025$.

[8] Su-Chan Park and Joachim Krug. Clonal interference in large populations. *PNAS*, 104(46):18135–40, Nov 2007.

[9] Su-Chan Park, Damien Simon, and Joachim Krug. The speed of evolution in large asexual populations. *J Stat Phys*, 138(1-3):381–410, Jan 2010.

[10] J R Peck. A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. *Genetics*, 137(2):597–606, Jun 1994.

[11] William H Press, Saul A Teukolsky, William T Vetterling, and Brian P Flannery. Numerical recipes 3rd edition: The art of scientific computing. *Cambridge University Press*, 2007.

[12] Claus O Wilke. The speed of adaptation in large asexual populations. *Genetics*, 167(4):2045–53, Aug 2004.