The Nonstationary Dynamics of Fitness Distributions: Asexual Model with Epistasis and Standing Variation

Guillaume Martin*,[1] and Lionel Roques†

*Institut des Sciences de l'Evolution–Montpellier, (UMR 5554) Centre National de la Recherche Scientifique, 34095 Montpellier, France, and †BioSP, INRA, 84914, Avignon, France

ORCID IDs: 0000-0002-5203-5091 (G.M.); 0000-0001-6647-7221 (L.R.)

ABSTRACT Various models describe asexual evolution by mutation, selection, and drift. Some focus directly on fitness, typically modeling drift but ignoring or simplifying both epistasis and the distribution of mutation effects (traveling wave models). Others follow the dynamics of quantitative traits determining fitness (Fisher's geometric model), imposing a complex but fixed form of mutation effects and epistasis, and often ignoring drift. In all cases, predictions are typically obtained in high or low mutation rate limits and for long-term stationary regimes, thus losing information on transient behaviors and the effect of initial conditions. Here, we connect fitness-based and trait-based models into a single framework, and seek explicit solutions even away from stationarity. The expected fitness distribution is followed over time via its cumulant generating function, using a deterministic approximation that neglects drift. In several cases, explicit trajectories for the full fitness distribution are obtained for arbitrary mutation rates and standing variance. For nonepistatic mutations, especially with beneficial mutations, this approximation fails over the long term but captures the early dynamics, thus complementing stationary stochastic predictions. The approximation also handles several diminishing returns epistasis models (*e.g.*, with an optimal genotype); it can be applied at and away from equilibrium. General results arise at equilibrium, where fitness distributions display a "phase transition" with mutation rate. Beyond this phase transition, in Fisher's geometric model, the full trajectory of fitness and trait distributions takes a simple form; robust to the details of the mutant phenotype distribution. Analytical arguments are explored regarding why and when the deterministic approximation applies.

KEYWORDS asexual evolution; clonal interference; transient dynamics; mutation-selection balance; Fisher's geometric model

EMPIRICAL dynamics of fitness in simple environments are still not quantitatively predicted by evolutionary biology, in spite of a wealth of theoretical progress and an ever-growing corpus of data produced by experimental evolution. To our knowledge, no model exists that was parameterized from independent data and that has proved to predict observed fitness trajectories in either sexual or asexual organisms, from *de novo* mutations, or preexisting standing variance. Patterns of fitness trajectories in microbes (*de novo* mutations in asexuals) have been confronted to and fitted with various theoretical predictions, showing qualitative agreement with models of clonal interference (Tsimring *et al.* 1996; Miralles *et al.* 2000; Gerrish 2001; Desai *et al.* 2007), and suggesting pervasive diminishing returns epistasis among beneficial mutations (Chou *et al.* 2011; Khan *et al.* 2011). However, *fitting* is not *predicting*: several alternative models can be qualitatively consistent with the same data set (Frank 2014). Regarding fitness dynamics during adaptation from standing variance, both theory and data are relatively scarce, at least in asexuals. This limits our knowledge of the transient effects of standing variance, yet these can be critical for short-term adaptive responses to environmental challenges.

Important progress has been made, over several decades, with a rich variety of models predicting fitness dynamics. These models critically depend on (i) a mutation rate, and (ii) a distribution of fitness effects (DFE) of mutations, which is either independent of the background genotype (no epistasis for fitness), or depends on it, typically via its fitness. They differ in the genotype-fitness landscape considered and the

doi: 10.1534/genetics.116.187385
Manuscript received January 21, 2016; accepted for publication October 10, 2016; published Early Online October 21, 2016.
Supplemental material is available online at http://www.genetics.org/cgi/content/full/genetics.116.187385/DC1.
[1]Corresponding author: Institut des Sciences de l'Evolution–Montpellier, ISEM CNRS UMR 5554, Université Montpellier, Bât. 22, Pl. Eugéne Bataillon, 34095 Montpellier, France. E-mail: guillaume.martin@umontpellier.fr

regimes assumed to derive the evolutionary dynamics. Models of mutation and selection in asexuals roughly fall into two (seemingly disconnected) classes: DFE-based models that directly track the distribution of fitness, and trait-based models that follow the distribution of underlying quantitative traits, which determine fitness. The aim of this work is to handle this variety of models into a single analytical framework [in terms of partial differential equations (PDEs)], and to use it to derive new results for these models, regarding nonstationary dynamics or equilibria. We start by briefly summarizing these existing approaches, in a necessarily far-from-exhaustive manner.

Fitness-based models directly follow the dynamics of fitness distributions, typically with a constant mutation rate and DFE over time (no epistasis). Initially based on deterministic equations and diffusive mutation effects (Tsimring *et al.* 1996), they were then refined to include stochasticity and more general DFEs of purely beneficial mutations (Gerrish and Lenski 1998; Rouzine *et al.* 2003; Dwyer 2012; Good *et al.* 2012). More recently, the interplay of a distribution of deleterious and beneficial mutations has been studied in this context, in either low (*e.g.*, Good and Desai 2014) or high (*e.g.*, Neher and Hallatschek 2013) mutation rate limits. As beneficial mutation influx becomes large in asexuals, cosegregating lineages compete for fixation and slow down adaptation; a process further affected by the deleterious mutations that accumulate in each lineage. These "clonal interference" dynamics, in the presence of stochastic fluctuations, are difficult to analyze and often yield complex or nonexplicit formulas; but several models have provided important insight into this process. They have been handled through alternative modeling approaches, accurate in different regimes: low to intermediate mutation rate for the original clonal interference models (Gerrish and Lenski 1998; Gerrish 2001), or higher mutation rate for the more recent "traveling wave" models (Rouzine *et al.* 2003; Good *et al.* 2012; Neher and Hallatschek 2013). Note that in the limit of very large populations, high mutation rates, and weak mutation effects, a simple and explicit Gaussian traveling wave is retrieved for the expected fitness distribution (Neher and Hallatschek 2013).

This rich literature, reviewed elsewhere (*e.g.*, Rouzine *et al.* 2003; Desai and Fisher 2007; Sniegowski and Gerrish 2010; Desai 2013), has a common feature: it describes the stationary regime of a stochastic process. This implies that a full trajectory from given initial conditions (possibly with standing variance) is not available, only the ultimate average rate of steady fitness change. Furthermore, as time goes on, the envelope around this mean fitness prediction typically explodes so that individual populations may lie far from the predicted mean at any time. This limits the comparison to empirical trajectories, which typically start away from a stationary regime and contain a few replicates. Note, however, that this assumption of a steady increase in fitness is often envisioned as reflecting a constant struggle between a steadily changing environment and an adapting population (Neher and Hallatschek 2013). It is possible that in such a regime the envelope may remain narrow and steady state may be reached faster.

Another aspect of the approach is that epistasis must be ignored here, otherwise mutation rates and effects may change over time (as the dominant backgrounds change); impeding the setting of a stationary regime. Recent extensions do include some form of epistasis or deleterious mutations (Kryazhimskiy *et al.* 2009; Dwyer 2012; Good and Desai 2015). However, analytical progress is then difficult beyond the master equation: relatively simple exemplary cases were analyzed in depth, but always in regimes where clonal interference is negligible. Also note that other DFE-based models were devoted to describe mutation-selection balance (another stationarity assumption), ignoring drift and epistasis. General insight into equilibrium fitness distributions has been gained from quasi-species theory (Eigen 1971) or asexual mutation-selection-balance models (Johnson 1999). This literature will not be reviewed here either (see Wilke 2005), but in general, analytical progress has often proved difficult unless simplified forms of DFE are assumed (discussed in Martin and Gandon 2010).

Trait-based models form an equally central body of literature that deals with adaptation affecting a trait or set of traits under selection for an optimum (via some concave phenotype-fitness function). These single peak, trait-based models date back to Fisher's (1930) geometric model (FGM), and also produced a rich literature connected to evolutionary quantitative genetics (Lande 1979). This approach is constrained into a particular form of DFE, but one that does include (i) pervasive epistasis and dominance, and (ii) both beneficial and deleterious mutations. Several patterns of mutant fitness expected in the FGM have been tested on fitness data from mutant lines (Martin *et al.* 2007; Trindade *et al.* 2010, 2012; Manna *et al.* 2011; Sousa *et al.* 2011; Hietpas *et al.* 2013), showing promising overall agreement. The FGM also emerges as the limit of a broader class of genotype-phenotype-fitness landscapes involving highly integrated "small-world" phenotypic networks (Martin 2014). Overall, the FGM seems a reasonable null model for evolutionary predictions (reviewed in Tenaillon 2014). The population genetics of adaptation by mutation and selection, in such trait-based models, has also seen many developments; reviewed extensively elsewhere (*e.g.*, Burger 2000; Orr 2005). It provides a well-studied theory for equilibrium states in various situations (detailed in Roze and Blanckaert 2014); several qualitative properties of equilibria have even been obtained for more general trait-fitness relationships, at least with a single trait (detailed in Burger 1998, 2000). The effect of standing genetic variance has also been studied extensively (from its quantitative genetics heritage), making the FGM an interesting complement to DFE-based models. Furthermore, predictions on trait distributions can be transformed into predictions on measurable fitness distributions under the model (*e.g.*, Martin and Gandon 2010). Yet, in spite of interest in its potential (Barton 1998; Gordo and Campos 2012), analytic progress in situations relevant for

experimental evolution (notably asexuals), has proven equally difficult to obtain. Even equilibrium states are not fully resolved in the FGM. Alternative analytic approximations only exist at each extreme of the mutation rate spectrum: house of cards for a single trait (Turelli 1984) *vs.* Gaussian for arbitrarily many traits (Kimura 1965; Lande 1980) in the low *vs.* large mutation rate limits, respectively. When dealing with the dynamics of adaptation, the classic approach (Lande 1979) focuses on large highly polymorphic sexual populations where the genetic variance of the traits is transiently approximately constant; another stationarity assumption, that is valid this time over finite timescales. However, this option generally breaks down with asexuals. Alternatively, stochastic models of mutation-selection-drift dynamics have been implemented under the FGM for adaptation trajectories (Orr 2000) or mutation-selection-drift balance (Tenaillon *et al.* 2007). However, they apply in a weak-mutation, strong-selection limit (or with unlinked nonepistatic loci in sexuals) where clonal interference is negligible. Finally, it is noteworthy that treatments of trait-based models with high mutational input (Gaussian theories) put less emphasis on drift (often neglected) than their fitness-based counterparts. They do involve multiple cosegregating mutants (clonal interference), but the deterministic predictions prove fairly accurate in this case; suggesting that some difference in the assumptions makes the interplay of drift and other forces less critical.

Aim of this work

Overall, we enjoy a wealth of alternative, complementary approaches of adaptive (or maladaptive) fitness dynamics in the presence of mutation, selection, and possibly drift. Yet, they are not easily connected. They do not provide a readily testable prediction in terms of trajectories of fitness distributions over time, from known initial conditions, in the large asexual populations typical of evolution experiments. To derive such predictions, we extend an approach initially proposed by Burger (1991), who studied trait-based models via the dynamics of the cumulants of the trait distribution, under selection and nonepistatic mutation. We apply this framework to fitness itself. Deterministic dynamics of fitness cumulants/moments have been used previously in nonepistatic fitness-based models; either neglecting drift (Johnson 1999; Desai and Fisher 2011; Gerrish and Sniegowski 2012) or including a stochastic diffusion component and considering the expected cumulants over replicates (Rattray and Shapiro 2001; Good and Desai 2013). Following Burger's (1991) strategy, these studies solved a finite set of cumulant equations numerically, but the system could not be closed as cumulants/moments influence each other in cascade. Here, we focus on the moment generating function (MGF) and cumulant generating function (CGF) of the fitness distribution, which handles all moments (respectively cumulants) in a single function. In a variety of models, this allows us to "close the system" into a single PDE describing the dynamics of the expectation of the fitness distribution, among stochastic replicates, by ignoring the effect of drift. We further include mutational epistasis by considering DFEs that broadly depend on background fitness. Overall, several processes are jointly handled by the PDE (Figure 1): starting from an arbitrary initial fitness distribution, new mutations accumulate on each lineage (with lineage-dependent DFE) and cosegregate under selection (clonal interference). In several classes of models, explicit solutions can be found for the PDE, providing a fully analytic theory in terms of mutational parameters and standing variance. We check the predictions against stochastic, individual-based simulations of various subcases.

Heuristic statements

Before describing the model in more mathematical detail, we first tackle some qualitative aspects of fitness dynamics in the different models above. Let us start by a somewhat technical remark that justifies the use of generating functions here. With any model where the DFE only depends on parental fitness and in an asexual (no recombination/segregation); fitness is the only trait whose distribution fully determines its own evolution. We can thus follow this distribution alone, ignoring the genetic or phenotypic details underlying its variation, namely the number and effects of the mutations carried by different genotypes, over the entire genome. This does not preclude the complications described above: multiple mutations accumulate on each lineage; multiple lineages cosegregate and compete for ultimate fixation; and each lineage may have its own background-dependent DFE (epistasis), as long as this dependence is entirely mediated by the background fitness. Generating functions handle sums of independent variables in a convenient manner, which helps us to study the cumulative effect of multiple mutations accumulating in lineages. It is also known that the effect of selection on fitness distributions takes on a simple form in terms of generating functions (Hansen 1992; Manna *et al.* 2012).

Second, let us consider why and when drift may be ignored in a given finite population, or among replicate finite populations, to describe the average fitness trajectory. The primary impact of drift identified in stochastic fitness-based models lies in its impact on the very fittest edge of the fitness distribution. When this edge represents a small absolute number of individuals, stochastic fluctuations in this subpopulation indirectly bias the future mean fitness dynamics of the whole population, over longer timescales. This effect does not average out if we consider the average mean fitness of replicate populations. However, over a substantial initial period, this fitter edge has little influence on the mean fitness dynamics (discussed in Gerrish and Sniegowski 2012) for two reasons. First, in a large polymorphic population, the short-term mean fitness dynamics are driven by selection and mutation in the bulk of the population, which behaves roughly deterministically. Second, even in a smaller population, drift, of itself, only slightly alters the average frequency dynamics of genotypes, roughly by an order of $-s\langle p^2\rangle/N$ where N is population size and p and s are the allele's frequency and fitness effect, respectively (see, *e.g.*, Otto and Barton 2001). Therefore, any quantity that is linear in genotype frequencies, such as mean

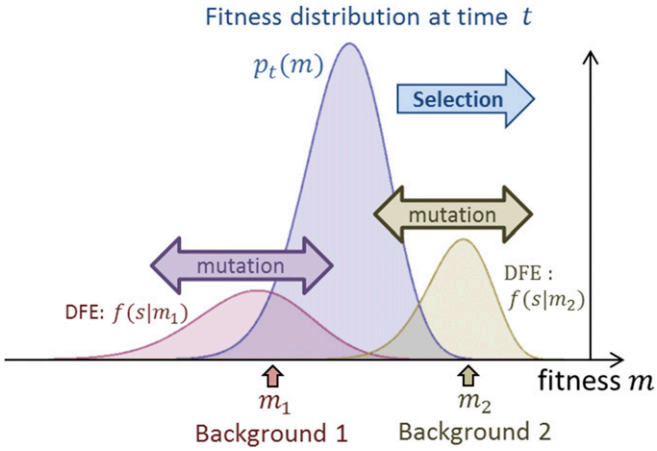

Figure 1 The standing fitness distribution [$p_t(m)$, blue curve] travels to the right by selection. Each genetic background under this distribution (*e.g.* m_1 and m_2 here) mutates to new genotypes with fitness $m_i + s$, where s has the density $f(s|m_i)$ depending on background fitness (red and brown curves for m_1 and m_2, respectively).

fitness or the MGF of the fitness distribution, is only slightly affected by drift over this timescale. It is only once new mutants establish (or not) that the future of the fitness dynamics is inaccurately predicted by a deterministic model: ignoring the stochastic loss of these fitter genotypes leads to an overestimate of mean fitness over longer timescales. Finally, even over longer timescales, the bias induced by drift is only visible if it accumulates over time, as the fittest edge stochastically moves toward fitter classes (at a speed overestimated by the deterministic model). If the set of all possible fitnesses is bounded by some maximal value, stochastic fluctuations should become less important, as the edge cannot spread forward forever: the delay between the edge and the bulk is bounded and tends to decay over time (as the bulk adapts). Most trait-based models consider adaptation toward a phenotypic optimum, implying a form of diminishing returns epistasis, where fitness is bounded on the right by the fitness of this optimum. This may explain why the evolutionary dynamics in these models has been accurately captured by deterministic theories. The same applies for purely deleterious models, where fitness cannot travel beyond the unloaded fitness class. In this case, however, loss of the fitter class also occurs and affects the long-term dynamics [Muller's (1932) ratchet]. Yet, this happens over much longer timescales, as the edge is a large subpopulation and as each "click" of the ratchet has a small impact (especially with continuous DFEs, where the new fittest class typically lies close to the previous one). This argument suggests that, in the presence of a fitness upper bound, it may be possible to accurately capture fitness dynamics by a mere deterministic model, even if clonal interference is involved and even over long timescales. It also suggests that nonepistatic models with beneficial mutations (where deterministic models fail in the long run) could still show transient fitness dynamics in which the average (over replicates) is captured by a deterministic model. Deriving such predictions (and justifying the above heuristic statement), as well as testing their accuracy with stochastic simulations is the central aim of this article.

Methods

General setting

We assume finite haploid asexual populations and follow the expected fitness distribution among replicates, starting from the same initial fitness distribution. We consider a continuous time model (overlapping generations), measured in arbitrary units (hours, days, *etc.*). This setting can also approximate a discrete time model (nonoverlapping generations) when effects are small per generation, the time t is then measured in generations: this will be our simulation scheme. We follow the dynamics of the distribution of the Malthusian fitness m (hereafter "fitness"). In continuous time, this is the expected exponential growth rate of a given genotype. In a discrete time approximation, m is the log of the Darwinian fitness ($m = \log W$), namely, the log of the expected geometric growth rate of a genotype. We define fitness relative to a reference, set at $m = 0$, without loss of generality. This reference is arbitrary as we consider evolutionary dynamics (relative fitness) without coupling to demography. In those models that include some fitness upper bound (*e.g.*, single-peak landscape models or models with only deleterious mutations), we set the optimal genotype (with fitness equal to this maximum) to be the reference $m = 0$ for convenience (so that all $m \leq 0$). In other models (*e.g.*, models with context-independent beneficial mutations), the reference is just an arbitrary point in fitness space. At any time t, an arbitrary set of K_t genotypes, with constant fitnesses $\{m_i\}_{i\in[1,K_t]}$, coexist in relative frequencies $p_t(m_i)$, satisfying $\sum_{i=1}^{K_t} p_t(m_i) = 1$. This approach can describe discrete classes (K_t finite) or infinite countable classes in the limit $K_t \to \infty$ (with convergence to a continuous distribution of fitness). Genotypes compete by frequency-independent selection and mutate according to a Poisson process with fixed rate U per capita per unit time. The fitness of a mutant whose parent has fitness m is $m + s$, where s is the selection coefficient of the mutation relative to the parent, and is drawn from an arbitrary distribution with probability distribution function (pdf) $f(s|m)$ (a probability density function if the distribution is continuous), depending on the parent fitness m.

Notations

We must define various expectations and means. We use an overbar $\bar{X}$ to describe any variable $X(m)$, averaged over the current distribution of genotypes within a focal population: $\bar{X} = \sum_{i=1}^{K_t} p_t(m_i)X(m_i)$. We define the expectation $E(Y|m)$ of any function $Y(s)$ of the DFE in background $m : E(Y|m) = \int Y(s)f(s|m)ds$, and we denote the mean DFE by $\mu_s = E(s)$ whenever it does not depend on m.

Generating functions

The distribution of m at time t can be characterized by its MGF: $M_t(z) = \overline{e^{m\ z}} = \sum_{i=1}^{K_t} p_t(m_i) e^{m_i z}$. For any finite population (K_t finite), this MGF is always defined over the full line $z \in \mathbb{R}$, but we may study it on a compact subset spanning 0 (here, $z \in \mathbb{R}^+$) without loss of generality: this helps handle several continuous class limits (when $K_t \to \infty$). This generating function provides essential information on the distribution at time t: its derivatives at $z = 0$ are the raw moments of the fitness distribution, notably the mean fitness $\bar{m}_t = M_t'(0)$ (the prime refers to differentiation with respect to z). For mathematical convenience, we mostly focus on the natural logarithm of the MGF, which is the CGF: $C_t(z) = \log M_t(z)$. Its derivatives at $z = 0$ are the cumulants of the distribution: in particular, the first three derivatives are the mean $\bar{m}_t = C_t'(0)$, variance $V_t = C_t''(0)$, and third central moment (related to skewness) $\kappa_3 = C_t'''(0)$. Additionally, the maximum of the distribution is given by $C_t'(\infty)$ and the weight of the class $m = 0$ is given by $\rho_t = e^{C_t(\infty)}$; we say that the distribution has a "spike" at $m = 0$ when this quantity ρ_t is strictly positive. It should also be noted that the full distribution of m at time t can be retrieved by applying an inverse Laplace transform to $M_t = e^{C_t}$.

Because each replicate population has its own trajectory of genotypic frequencies, the generating functions $M_t(\cdot)$ and $C_t(\cdot)$ are stochastic functions of z over time. We seek to predict the behavior of the *expectation* of such variables over stochastic replicates, so we use $\langle X \rangle$ to denote any such expectation of X. In particular, $\langle M_t(z) \rangle$ and $\langle C_t(z) \rangle$ are the expected MGF and CGF, which are deterministic functions of z and t, while $\langle \bar{m}_t \rangle$ and $\langle V_t \rangle$ are the expected mean fitness and variance in fitness within populations. These are deterministic functions of time.

Organization of the article

All appendices are included in Supplemental Material, File S1. In Appendix A, we derive *exact* dynamics for the expected generating functions, which do not close. Then we describe *approximate* closed dynamics for these quantities under a deterministic approximation ignoring drift. In Appendix B, we derive general properties of the approximate dynamics, and Appendices C, D, and E provide detailed applications to particular classes of mutational models. In the *Model* section below, we summarize our results on the expected CGF $\langle C_t(z) \rangle$, and its approximate deterministic counterpart, denoted $\mathcal{C}_t(z) \approx \langle C_t(z) \rangle$ (the $\approx$ sign is a reminder that the result is approximate). The *Application* section then illustrates applications to several classes of mutation models, evaluating the accuracy of the approximation on stochastic simulations. A last section summarizes some analytic results on the error involved by the approximation, and hints on why and when it applies. All notations are summarized in Table 1.

Dynamics of the expected CGF under selection, drift, and mutation

Using a multi-type Wright–Fisher diffusion approximation to genotype frequency dynamics (Section II in Appendix A, File S1), it can be shown that the change by selection and drift (SD) over Δt, in the expected CGF $\langle C_t(z) \rangle$ satisfies

$$\begin{aligned} \frac{\underset{\text{SD}}{\Delta} \langle C_t(z) \rangle}{\Delta t} &= \langle C_t'(z) \rangle - \langle C_t'(0) \rangle + \delta_t(z), \\ \delta_t(z) &= \frac{1 - \left\langle e^{\mathcal{C}_t(2z) - 2\mathcal{C}_t(z)} \right\rangle}{2N_e}. \end{aligned} \tag{1}$$

Here, $\delta_t(z)$ is the contribution generated by drift (it vanishes if $N_e \to \infty$), which is essentially the same as given in Good and Desai's (2013) equation D.4. This dynamic term does not allow us to close the system as $\delta_t(z)$ does not depend directly on $C_t(z)$. We thus rely on a deterministic approximation (that we will use throughout), which simply ignores $\delta_t(z)$ in the dynamics, yielding an approximate expected CGF ($\mathcal{C}_t(z) \approx \langle C_t(z) \rangle$), with closed dynamics $\underset{\text{SD}}{\Delta} \mathcal{C}_t(z)/\Delta t = \mathcal{C}_t'(z) - \mathcal{C}_t'(0)$.

Mutation (see the *General setting* section above) generates a DFE whose MGF is denoted $M^S(z, m) = \int_{\mathbb{R}} f(s|m) e^{sz} ds$. It is assumed to have known analytical form, over some positive domain $z \in [0, z_{\max}] \subset \mathbb{R}^+$, determined by the model considered. This may include continuous or discrete distributions, but it does require a DFE with finite higher moments (so that an MGF can be analytically defined). The change in $(C_t(z))$ [and $\mathcal{C}_t(z)$] by mutation (mut), over Δt, takes the general form (Section III.1 in Appendix A, File S1):

$$\frac{\underset{\text{mut}}{\Delta} \langle C_t(z) \rangle}{\Delta t} = \frac{\underset{\text{mut}}{\Delta} \mathcal{C}_t(z)}{\Delta t} = U \left(\left\langle \frac{\overline{e^{m\ z} M^S(z, m)}}{\overline{e^{m\ z}}} - 1 \right\rangle \right), \tag{2}$$

where we recall that $\langle \cdot \rangle$ is the expectation over replicate populations; while the overbar refers to the averaging with respect to m, within a given population, at current time t. The limit, as $\Delta t \to 0$, of $(\underset{\text{SD}}{\Delta} \mathcal{C}_t(z) + \underset{\text{mut}}{\Delta} \mathcal{C}_t(z))/\Delta t$ from Equation 1 and Equation 2, yields the continuous time dynamics of the expected CGF, under the deterministic approximation:

$$\begin{aligned} \partial_t \langle C_t(z) \rangle &\approx \partial_t \mathcal{C}_t(z) \\ &= \mathcal{C}_t'(z) - \mathcal{C}_t'(0) + U \left(\left\langle \frac{\overline{e^{m\ z} M^S(z, m)}}{\overline{e^{m\ z}}} \right\rangle - 1 \right). \end{aligned} \tag{3}$$

This is our central result, from which all the following dynamics are derived. In general, the mutation kernel in Equation 3 does not generate a closed system, even under the deterministic approximation, as the mutational term cannot be expressed in terms of $\mathcal{C}_t(\cdot)$. Fortunately, this term simplifies in several general classes of models, which we detail below, summarize in Table 2, and implement in File S3 (see below).

Table 1 Main notations used throughout the article

Notation	Description	Formula
m	Malthusian fitness	
$\{m_i\}_{i \in [1,K_t]}$	Fitness classes within a population	
$p_t(m_i)$	Frequency of the fitness class m_i at time t	
N, N_e	Population size, effective size	
K_t	Number of fitness classes at time t	
$\bar{m}_t$	Mean fitness at time t	$\sum_{i=1}^{K_t} p_t(m_i)m_i$
V_t	Variance in fitness at time t	$\sum_{i=1}^{K_t} p_t(m_i)m_i^2 - \bar{m}_t^2$
ρ_t	Weight of the class $m = 0$	$p_t(m_i = 0)$
$\bar{X}$	Mean value of any variable $X(m)$, averaged over the current distribution of genotypes within a population	$\sum_{i=1}^{K_t} p_t(m_i)X(m_i)$
$\langle\rangle$	"Ensemble expectation" of any random variable, averaged over replicate (finite) populations	
$M_t(z)$	"Empirical" MGF of m in a given population, at time t	$\sum_{i=1}^{K_t} p_t(m_i)e^{m_i z}$
$C_t(z)$	"Empirical" CGF of m in a given population, at time t	$\log M_t(z)$
$\mathcal{M}_t(z)$	Expected MGF under the deterministic approximation	$\mathcal{M}_t(z) \approx \langle M_t(z)\rangle$
$\mathcal{C}_t(z)$	Expected CGF under the deterministic approximation	$\mathcal{C}_t(z) \approx \langle C_t(z)\rangle$
DFE	Distribution of fitness effects of mutations	
s	Selection coefficient of a mutation relative to its parent	
$f(s\|m)$	Probability distribution function of s in background m	
$E(Y\|m)$	Expectation of any variable $Y(s)$ over the DFE in background m	$\int_{\mathbb{R}} Y(s)f(s\|m)ds$,
μ_s	Mean effect of mutations on fitness in the background with fitness $m = 0$ (or any background in nonepistatic models)	$\int_{\mathbb{R}} s\, f(s\|m = 0)ds$
$M^S(z,m)$	MGF of the DFE	$\int_{\mathbb{R}} f(s\|m)e^{s\,z}ds$
$M_*(z)$	MGF of the DFE in the background with fitness $m = 0$	$M^S(z,0)$
$\omega(z)$	Linear effect of m on the CGF of the DFE	$\partial_m \log M^S(z,m)\big\|_{m=0}$
s_H	Harmonic mean in absolute value of the DFE in the background $m = 0$	$1/E(1/\|s\|)$
U	Genomic mutation rate	
L	Mutation load (with an optimal fitness class at $m = 0$)	$L = -\langle\bar{m}_\infty\rangle$
FGM	Fisher's geometric model	
n	Dimension of the phenotypic space	
λ	Mutational variance at each trait	

Linear background dependence

The first important situation is when the MGF of the DFE can be (exactly or approximately) written as a linear function of $m : M^S(z,m) = a(z)m + M_*(z)$, with some function a and with $M_*(z) = M^S(z,0)$ being the MGF of the DFE in the background with fitness $m = 0$. As an MGF, M_* is continuous on a domain including 0 and must satisfy $M_*(0) = 1$ and $M_*''(z) \geq 0$. The function a must satisfy $a(0) = 0$ and either $a''(z) \leq 0$ over $z \in \mathbb{R}^+$, if fitnesses are bounded on the right so that all $m \leq 0$; or $a''(z) = 0$, if fitnesses are unbounded on the right. This is required for M^S to satisfy the basic MGF properties $M^S(0,m) = 1$ (conservation of probability) and $M^{S''}(z,m) \geq 0$ (convexity for all m and z). Linear background dependence (see Section III.2 in Appendix A, File S1) implies a mutation kernel (Equation 2) of the form $\Delta_{\text{mut}} C_t(z)/(U\Delta t) = M_*(z) - 1 + a(z)C_t'(z)$. The (approximate) expected CGF $\mathcal{C}_t(\cdot)$ then satisfies a first order, linear nonlocal PDE:

$$\partial_t \mathcal{C}_t(z) = \alpha(z)\mathcal{C}_t'(z) - \mathcal{C}_t'(0) + \beta(z), \tag{4}$$

where the functional coefficients are $\alpha(z) = 1 + U\ a(z)$ and $\beta(z) = U(M_*(z) - 1)$, with $\alpha(0) = 1$ and $\beta(0) = 0$. This PDE has the boundary condition $\mathcal{C}_t(0) = 0$ and initial condition $\mathcal{C}_0(z) = C_0(z)$ (initial fitness distribution) and can be solved analytically (Section II.1 in Appendix B, File S1). Define the function y as the solution of the ordinary differential equation $y'(z) = \alpha(y(z))$ with initial condition $y(0) = 0$ and its functional inverse $y^{-1}(z) = \int_0^z 1/\alpha(v)dv$, such that $y(y^{-1}(z)) = z$, defined on $[0,\ z_1)$, where z_1 is the first positive root of α. The unique solution of Equation 4 from initial condition $C_0(z)$ is

$$\mathcal{C}_t(z) = C_0\big(y\big(y^{-1}(z) + t\big)\big) - C_0(y(t)) + \int_0^t \beta\big(y\big(y^{-1}(z) + v\big)\big) - \beta(y(v))dv. \tag{5}$$

The corresponding trajectory of the expected mean fitness is (under the deterministic approximation)

$$\langle\bar{m}_t\rangle \approx \mathcal{C}_t'(0) = \alpha(y(t))C_0'(y(t)) + \beta(y(t)), \tag{6}$$

for all $t \geq 0$. A similar explicit expression is given in File S1, Appendix B (Equation B31) for the trajectory of the expected

Table 2 Various mutational models handled by the proposed framework

Model	Background dependence	Timescale of applicability	$\omega(z)$ or $a(z)$	$M_*(z)$
Nonepistatic deleterious	None	$t \leq T \sim N_e\ e^{-U/s_H}$	0	Arbitrary < 1
Nonepistatic deleterious and beneficial	None	$t \leq T \approx 100 - 1000$[a]	0	Arbitrary
House of cards	Log-linear	$t \in \mathbb{R}^+$	$-z$	Arbitrary
Binary model[b]	Linear	$t \in \mathbb{R}^+$	$-2\sinh(\delta z)/(\Lambda\delta)$	$e^{-\delta z}$
Gaussian FGM	Log-linear	$t \in \mathbb{R}^+$	$-\lambda\ z^2/(1+\lambda z)$	$(1+\lambda\ z)^{-n/2}$
Generalized FGM	$\approx$ Linear($U \gg U_c$)	$t \in \mathbb{R}^+$	$\approx -2\lvert\mu_s\rvert z^2/n$	$\approx 1 - z\ \lvert\mu_s\rvert$
Diminishing return	$\approx$ Linear (near equilibrium)	At equilibrium	$\in [-z, 0]$	Arbitrary < 1

These models only apply when $N_e U\lvert\mu_s\rvert \gg 1$. For each model, each column gives (i) the model type, (ii) the type of background dependence, (iii) the timescale (sometimes approximate) over which the prediction applies (in that it is expected to be reasonably to very accurate), (iv) the background-dependence function [$\omega(z)$ for log-linear background dependence or $a(z)$ for linear background dependence], and (v) the MGF $M_*(z)$ of the DFE in the background with fitness $m = 0$ (fittest background in models with a maximum fitness). In some models the "$\approx$" notifies that this is an approximate result or a conjecture; "$T \sim N_e\ e^{-U/s_H}$" means that the two quantities have the same order of magnitude.

[a] Conjecture and timescale based on observations in our simulations.

[b] Simplified version of Rouzine *et al.*'s (2003) model, detailed in Section III.2, Appendix A, File S1. Here, Λ is the number of sites ('L in the original paper) and δ is the constant deleterious effect of "mutant" alleles (s in the original paper).

variance $\langle V_t\rangle \approx \mathcal{C}_t''(0)$. More generally, Equation 5 gives the trajectory of the whole fitness distribution, for several classes of models described in the *Application* section.

Examples of linear background-dependence models

Nonepistatic models: An obvious case of linear background dependence is for any nonepistatic model (whose DFE has finite moments, so that its MGF exists). In these, we have $M^S(z,m) = M_*(z)$ for all backgrounds, so that $a(z) = 0$ and $\alpha(z) = 1$.

Simplified version of the "binary model": In Rouzine *et al.*'s (2003) binary model (detailed in Section III.2, Appendix A, File S1), genotypes consist of Λ bins representing sites (we use notations different from the original article to avoid confusions with other quantities in this article). Each bin codes for a wild-type ("0") or mutant ("1") allele (with constant deleterious effect $-\delta < 0$). Mutation, at rate u per site (genomic rate $U = u\Lambda$), randomly creates shifts between allele states and allelic effects add up across the genome. This model shows mutational epistasis (the DFE depends on the background m), although fitness is still a sum of allelic effects over the genome. It also implies an upper bound $m = 0$ to all possible fitnesses (*i.e.*, the unloaded wild-type with only 0 bins) and has linear background dependence (see Table 2 and Equation A10, Appendix A in File S1). It can be checked that $a(0) = 0$ and $a''(z) \leq 0$ over $\mathbb{R}^+$. We do not explore this model further here, except in File S3 (see below).

Log-linear background dependence

Alternatively, the MGF of the DFE may be log linear in m : $M^S(z,m) = M_*(z)e^{\omega(z)m}$. Here, again, $M_*(z) = M^S(z,0)$ is convex and satisfies $M_*(0) = 1$, while ω must be concave (with bounded fitness set $m \leq 0$) and $\omega(0) = 0$. Plugging this form into the mutational kernel in Equation 2 yields another nonlocal first order PDE for the (approximate) expected CGF, but this time it is nonlinear (Section III.4 in Appendix A, File S1):

$$\partial_t \mathcal{C}_t(z) \approx \mathcal{C}_t'(z) - \mathcal{C}_t'(0) + U\left(M_*(z)\ e^{\mathcal{C}_t(z+\omega(z))-\mathcal{C}_t(z)} - 1\right), \tag{7}$$

for $t \geq 0$ and $z \geq 0$, with the boundary condition $\mathcal{C}_t(0) = \langle C_t(0)\rangle = 0$. The second term $U(M_*(z)e^{\mathcal{C}_t[z+\omega(z)]-\mathcal{C}_t(z)} - 1)$ in Equation 7 describes the effect of mutations accumulating on each background, with a dependence on the standing distribution of background fitnesses (on $\mathcal{C}_t$) mediated by $\omega(z)$. Note that this time this term is only approximate, under similar conditions as the deterministic approximation used all along (detailed in Section III.3 of Appendix A, File S1).

The well-posedness of Equation 7 requires that $0 \leq z + \omega(z)$, so that the nonlocal term remains within the domain under study. This is the case for any epistatic model ($\omega \not\equiv 0$) showing log-linear background dependence, with a fitness optimum at $m = 0$ (see Section I.1 in Appendix B, File S1). Although we were not able to get an explicit solution of Equation 7, which is a nonstandard PDE problem due to the two nonlocal terms $\mathcal{C}_t'(0)$ and $e^{\mathcal{C}_t(z+\omega(z))}$, we were able to get some insight into the behavior of the solution. First, $\mathcal{C}_t'(\infty) = 0$ for all positive times (Section I.2 in Appendix B, File S1) with the epistatic model ($\omega \not\equiv 0$). This means that the support of the fitness distribution instantaneously reaches the optimum $m = 0$, whatever the initial fitness distribution. It implies a memoryless property in the sense that the long-time behavior of the solution is not affected by the initial fitness distribution, which is not obtained in nonepistatic models ($\omega = 0$). Second, analytical expressions are derived (Section I.3 of Appendix B, File S1) for the k^{th} cumulants of the equilibrium distribution ($k \geq 0$) and a dichotomy for the value of the equilibrium mean fitness; namely, either $\langle \bar{m}_\infty \rangle = -U$ or $\langle \bar{m}_\infty \rangle = -U(1-B)$, for some positive constant B. Third, the existence of a spike implies that $\langle \bar{m}_\infty \rangle = -U$ (Section I.4 of Appendix B, File S1). These results were obtained under any of the two general properties (Section I.2 in Appendix B, File S1): H, any background can mutate to the optimal background; or H', any background can at best mutate to some fitter but suboptimal class.

Biologically, this simply means that some form of compensation of deleterious mutations exists.

Examples of log-linear background-dependent models

As an example, we describe two classic models of context-dependent DFEs where log-linear background dependence applies (see also Table 2).

Fisher's geometric model: Fisher's (1930) geometric model (FGM) assumes that each genotype is characterized by a (breeding value for) phenotype at n traits ($\mathbf{g} \in \mathbb{R}^n$) (possibly with some environmental variance effects). An optimal phenotype corresponds to maximal fitness and sets the origin of phenotype space ($\mathbf{g} = \mathbf{0}$). Fitness decreases away from this optimum, and mutation creates random independent and identically distributed (iid) variation $\mathbf{dg}$ around the parent, for each trait. In all our examples, we will consider a quadratic fitness function: in continuous time models, Malthusian fitness is a quadratic function of the breeding value $[m(\mathbf{g}) = -\|\mathbf{g}\|^2/2]$; and in discrete time versions, Darwinian fitness is a Gaussian function of $\mathbf{g}$ $[W(\mathbf{g}) = e^{m(\mathbf{g})} = \exp(-\|\mathbf{g}\|^2/2)]$. A classic version of this model is the "Gaussian FGM," where mutation phenotypic effects are multivariate normal: $\mathbf{dg} \sim N(\mathbf{0}, \lambda \mathbf{I}_n)$, where $\lambda > 0$ is the mutational variance at each trait, and $\mathbf{I}_n$ is the identity matrix in n dimensions. This Gaussian FGM is also the standard model of evolutionary quantitative genetics, dating back to Kimura's (1965) and Lande's (1980) work on mutation and selection on traits with a complex genetic basis (infinitely many possible alleles). The Gaussian FGM shows exact log-linear context-dependence (Martin 2014): $M^S(z, m) = M_*(z)e^{m\,\omega(z)}$ with $M_*(z) = (1 + \lambda\ z)^{-n/2}$ and $\omega(z) = -\lambda\ z^2/(1 + \lambda\ z)$. We study this model in depth in the *Application* section.

House-of-cards model: Kingman's (1978) house-of-cards (HOC) model assumes that mutants have absolute fitness that follows a unique distribution, independent of the background in which they arise. This model is epistatic in that the DFE depends on the background $f(s|m) = g(s + m)$, so that mutant absolute fitnesses X have a given fixed fitness distribution with pdf $g(x)$. Versions of the HOC were used, *e.g.*, in Kryazhimskiy *et al.* (2009) and McCandlish *et al.* (2014) with an exponential or Gaussian distribution g, respectively, and focusing on a regime of low NU where substitutions occur sequentially (no clonal interference). In this model, the MGF of the DFE is $M^S(z, m) = E(e^{s\ z}|m) = M_X(z)e^{-z\ m}$ where $M_X(z) = \int e^{x\ z} g(x)dx$ is the MGF of the chosen distribution of X with pdf g. Thus this model, in its general version, implies log-linear background dependence with $M_* = M_X$ and $\omega(z) = -z$. We do not explore this model further here, except in File S3 (see below).

Individual-based simulations

Individual-based, discrete-time simulations were used to check the validity of the approximations in finite populations for various mutational models. Individuals were sampled every generation according to their fitness $W = e^m$ (Wright–Fisher model of genetic drift and selection). Mutation was simulated in every generation in each individual by randomly drawing a Poisson number of mutations, each with effects drawn into a given DFE, and summing their effects to produce the mutant offspring. When considering trait-based models, genotypes where characterized by their breeding value in n dimensions $\mathbf{g} \in \mathbb{R}^n$. Mutation effects on traits were drawn into a given multivariate distribution and the fitness was computed as $m(\mathbf{g}) = -\|\mathbf{g}\|^2/2$ (quadratic landscape models, or "generalized FGM", see *Application* section).

Data availability

File S1 contains Appendices A–E describing all analytical derivations. File S2 provides a numerical solver of Equation 7, applied to the FGM as a MATLAB source code, together with a MATLAB graphical user interface and code for individual-based simulation. The solver is based on a finite-difference method with variable step sizes in z (smaller steps near $z = 0$, to get accurate values of the derivatives $\mathcal{C}_t'(0), \mathcal{C}_t''(0)$) and an implicit scheme in time. Because of the transport term $\mathcal{C}_t'(z)$, which tends to translate the solution toward the left with speed 1, the solution was computed on a finite interval of $z \in [0, t_{\max}]$, where $t_{\max}$ is the duration of the simulation. See Section V in Appendix D, File S1 for more details. File S3 gives examples of each subcase in the six first models in Table 2: analytical trajectories, solver for the PDE (Equation 7; method of lines), simulation code, and illustrations of theory *vs.* simulation results. File S4 contains movies illustrating the dynamics of the distribution of fitness (and trait when applicable, *i.e.*, in FGM) over time, compared to the theoretical distributions.

Application

Here we study various models for which the PDEs in Equation 4 and/or Equation 7 apply. We distinguish three main applications: (1) nonepistatic models of general form; (2) epistatic models of general form, nearing equilibrium; and (3) epistatic models generated by quadratic fitness functions of phenotypes (FGM). Throughout, we use the deterministic approximation, so we write $\approx$ to recall the approximate nature of our results.

Nonepistatic models

Before tackling epistatic models, we first focus on context-independent mutation models, mostly to check that we retrieve previously known properties and to provide some new results. Because several results on nonepistatic models are already known, we put most of the results on this section in a dedicated Appendix C, File S1, and focused on new insights. As we have seen, any nonepistatic model is a trivial subcase of Equation 4 with $\alpha(z) = 1$ $(y(z) = y^{-1}(z) = z)$ and $\beta(z) = U(M_*(z) - 1)$ Equation 5 yields

$$\begin{aligned}\mathcal{C}_t(z) &= C_0(z+t) - C_0(t) + U\int_0^t M_*(z+v) - M_*(v)dv,\\ \langle \bar{m}_t\rangle &= \langle C_t'(0)\rangle \approx \mathcal{C}_t'(0) = C_0'(t) + U(M_*(t)-1),\\ \langle V_t\rangle &= \langle C_t''(0)\rangle \approx \mathcal{C}_t''(0) = C_0''(t) + U(M_*'(t)-\mu_s),\end{aligned} \quad (8)$$

where we recall that $\mu_s = E(s)$. This result essentially retrieves an alternative formulation of equation 10 of Desai and Fisher (2011), itself a continuous time version of Johnson's (1999) equation 13. These previous results both assumed purely deleterious mutations, which proves unnecessary in the derivation of Equation 8. Equation 8 further allows for arbitrary standing variance in fitness via the additional term $C_0(z+t) - C_0(t)$, previously obtained for an infinite asexual population without mutation (Hansen 1992; Manna *et al.* 2012). As such, results in terms of CGFs or MGFs provide valuable information on the trajectory of moments, but are not so easy to fit on observed empirical distributions, which require an explicit distribution function. In Appendix C II, File S1, we derive the stochastic representation of fitness from Equation 8 to help derive such functions. Movie 1, A and B, in File S4, illustrates the dynamics of the full fitness distribution for a negative gamma DFE and a constant DFE, respectively. In the parameter range chosen, the prediction from Equation 8 accurately fits the observed distribution from the simulation of a single finite population of size $N = 10^5$. Other illustrative examples are given in File S1, Appendix C.

Retrieving previous results: Several key known results on nonepistatic deleterious mutation models are readily obtained from Equation 8 (detailed in Appendix C, File S1), such as properties of nonepistatic mutation-selection balance with arbitrary DFEs. In particular, Johnson's (1999) result for discrete fitness classes straightforwardly extends to continuous DFEs: the equilibrium fitness distribution is a negative compound Poisson, with Poisson parameter U/s_H where $s_H = 1/E(1/|s|)$ is the harmonic mean of the DFE in absolute value. Note that allowing for continuous distributions implies that the harmonic mean may be zero ($E(1/|s|) \to \infty$), in which case the spike of fittest genotypes (with weight e^{-U/s_H}) is *de facto* absent and the fitness distribution converges to a Gaussian (Equation C10 in Appendix C, File S1). Equation 8 also implies that with arbitrary deleterious DFE and mutation rate U, the mutation load is $L = -\langle \bar{m}_\infty \rangle = U$, and the equilibrium variance in fitness is $\langle V_\infty \rangle = U|\mu_s|$. This extends a result previously derived as a low mutation rate limit (Burger and Hofbauer 1994) to the full mutation rate spectrum.

Timescales of load build-up vs. loss of accuracy with purely deleterious mutations: Equation 8 allows us to derive the "characteristic time" t_q that it takes to reach some proportion q of the ultimate equilibrium ($\langle \bar{m}_{t_q}\rangle = -q\,U$). Neglecting standing variance this time, t_q is the solution of $M_*(t_q) = 1 - q$: notably, it is independent of the mutation rate. This time can be computed for any given DFE, and admits simple bounds in the general case (see Appendix C III, File S1). For example, $3/|\mu_s| \le t_{0.95} \le 8/s_H$, it takes between $3/|\mu_s|$ and $8/s_H$ generations to reach 95% of the load. We recall that $|\mu_s|$ and s_H are the arithmetic and harmonic means of the DFE in absolute value, respectively.

Nonepistatic models with beneficial mutations: When the kernel includes a portion of beneficial mutations ($M_*(\infty) = \infty$), mean fitness increases indefinitely ($\langle \bar{m}_t\rangle \to \infty$ in Equation 8) and our approach overestimates this increase after some time (see *Application*). For any nonepistatic model, the long-term fitness dynamics are best described by stochastic origin-fixation models (with or without clonal interference), once a stationary regime of fitness change has set. However, we propose that Equation 8 can provide some connection between the transient and stationary regimes and predict the fitness trajectory before stationarity (Section III in Appendix C, File S1). Assume a given rate ν of fitness change is predicted at stationarity. If we assume a sharp transition from a deterministic to a stochastic stationary regime, this transition must then occur when the deterministic and stochastic models have equal rates of mean fitness change, namely at some time $t = \tau$ such that $\partial_t\langle \bar{m}_t\rangle = \nu = U\,M_*'(\tau)$ (Equation 8, ignoring the contribution from standing variance). Up to this time, mean fitness is assumed to be given by the deterministic theory ($\langle \bar{m}_t\rangle = U\,M_*(t)$) while it increases steadily at rate ν afterward. This conjecture proves reasonable, as illustrated in Figure 2. In Figure 2A, the DFE consists of purely beneficial, exponentially distributed, mutation effects ($s \sim \mathrm{Exp}(1/\mu_s)$, with $\mu_s > 0$) and ν is given by clonal interference theory (equation 16 in Good *et al.* 2012). In Figure 2B, a shifted gamma DFE is considered: $s \sim s_0 + x$, with $s_0 > 0$ and $x \sim -\Gamma(a,b)$ and the stationary rate ν is computed empirically, based on the adaptation rate that is observed at large times in the individual-based simulations. Using only this rate ν as input, the transition time τ is computed and the full trajectory of expected mean fitness is predicted (see also Figures C2–C5 in Appendix C, File S1, for other parameter values and other DFEs). By construction, theory (lines) and average from simulations (circles) should have the same slope ν in the late linear increase phase. However, they need not be superposed, especially over the full timescale studied. Coarse grain observation indeed suggests that the whole trajectory is surprisingly well captured by this simple heuristic technique. However, a transiently oscillating behavior (of the *average* trajectories) arises around the inferred transition time τ in all our simulations. This shows that the actual behavior is more complex than a simple transition from nonstationary/deterministic to stationary/stochastic (discussed in Desai and Fisher 2007).

In any case, the simulations in Figure 2 and Figures C2–C5 in Appendix C, File S1, show that the simple deterministic approximation does capture the dynamics over possibly several 100 (Figure 2A) or 1000 (Figure 2B) generations (all the more as the proportion of beneficial mutations is small, apparently).

Furthermore, recall that this treatment only applies to thin-tailed DFEs (that fall off faster or as an exponential), otherwise

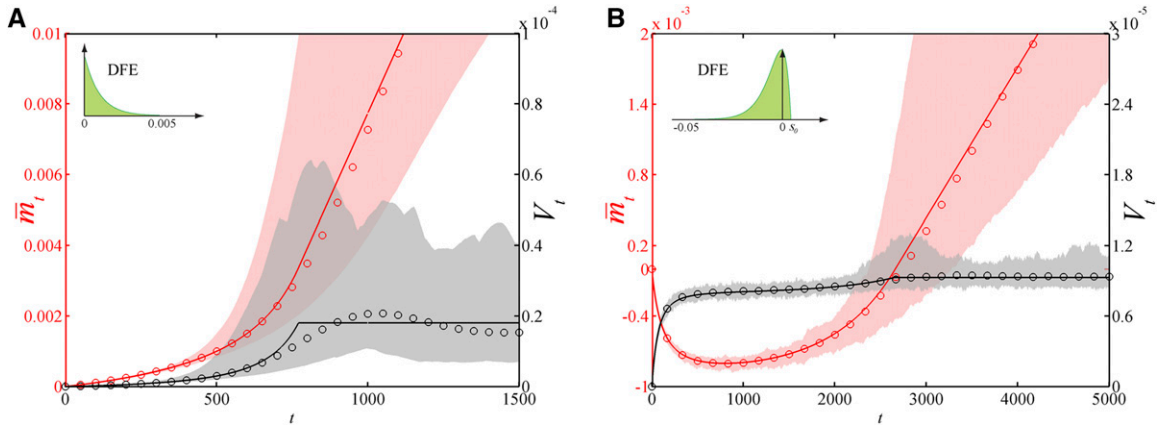

Figure 2 Mean fitness $\bar{m}_t$ and variance V_t trajectories in nonepistatic models including beneficial mutations. (A) Exponential DFE: $s \sim \text{Exp}(1/\mu_s)$ with mean effect $\mu_s = 0.001$. (B) Shifted gamma DFE: $s \sim s_0 + x$, with $s_0 > 0$ and $x \sim -\Gamma(a, b)$, with $a = 2$, $b = 5 \cdot 10^{-3}$ and $s_0 = a \cdot b/5$. In both cases, $U = 10^{-3}$. Solid lines: for $t < \tau$, the expected trajectories $\langle \bar{m}_t \rangle$ and $\langle V_t \rangle$ are given by our analytical theory (Equation 8); for $t \geq \tau$, the slope $\nu = \langle \partial_t \bar{m}_t \rangle$ and the variance $\langle V_t \rangle$ are kept constant. In (A), the transition time τ ($\simeq 770$) is such that ν equals the theoretical asymptotic slope given by equation 16 in Good *et al.* (2012); in (B), the transition time τ ($\simeq 2650$) is such that ν equals the empirical slope observed in the individual-based simulations during the interval $t \in (4000, 6000)$. ○: empirical mean fitness and variance given by individual-based simulations, averaged over (A) 10^3 populations or (B) 10^2 populations, with $N = N_e = 10^6$. Shaded regions: 99% C.I.'s for the mean fitness (in red) and the variance (in gray). We assumed initially clonal populations with $m_0 = 0$.

the MGF is not analytic and $\tau \to 0$. The limiting case is an exponential tail, for which τ becomes smaller as the tail falls slower (larger mean). Yet, the simple heuristic technique did show good accuracy when simulating exponential DFEs with $\mu_s = 0.01$ or 0.001 (Figure 2; Figures C2 and C3 in File S1).

Finally, Figure 2 and Figures C2–C5 in Appendix C, File S1, also illustrate that variation around the expected mean fitness explodes over time (red envelopes), especially after the transition to stationarity (late linear phase). Therefore, the empirical insight gained from the sole prediction of the expected mean fitness dynamics (without its envelope) can be *de facto* limited in this regime.

Equilibrium in the presence of diminishing returns epistasis

Now consider an epistatic model ($M^s(z, m) \neq M_*(z)$), where beneficial mutations become less frequent and of smaller effect as the population adapts; corresponding to a form of "diminishing returns" epistasis. More precisely, we assume that (i) fitnesses are bounded on the right (the maximum fitness is then set at $m = 0$), and (ii) there is compensation (suboptimal backgrounds produce a portion of beneficial mutations). In this case, near equilibrium, the fitness distribution shrinks toward the maximum, and a first order Taylor series of $C_S(z, m) = \log M^S(z, m)$ in small m yields $M^S(z, m) = M_*(z) e^{\omega(z) m}(1 + O(m^2))$. Here $\omega(z) = \partial_m C^S(z, m)\big|_{m=0}$ is the slope of the change with m of the CGF of the DFE, in the vicinity of $m = 0$, while $M_*(z) = M^S(z, 0)$ is the MGF of the DFE in the optimal background. Arbitrary models with diminishing returns epistasis (and a fitness upper bound) thus converge to log-linear background dependence near equilibrium. Then, by the memoryless property of log-linear background-dependent models (see Equation B3 in Appendix B I, File S1), the CGF converges as $t \to \infty$ to a unique equilibrium, independently of the initial CGF (the equilibrium cumulants are detailed in Section I, Appendix B, File S1). Overall, mutation-selection balance is therefore a local attractor for this class of models and a global attractor for models with exact background dependence (such as the FGM).

In order to get further insight into the equilibrium fitness distribution, we now use a linear approximation to the MGF with small m, yielding $M^S(z, m) = a(z)m + \beta(z) + O(m^2)$, where $\beta(z) = U(M_*(z) - 1)$ and $\alpha(z) = 1 + U M_*(z)\omega(z)$. The asymptotic properties of Equation 5 as $t \to \infty$ (Section II.2 and II.3 in Appendix B, File S1) then yield a general theory for mutation-selection balance in the presence of diminishing return epistasis.

Mutation load: In particular, mean fitness stabilizes to $\langle \bar{m}_\infty \rangle = \beta(z_1)$, where z_1 is the smallest positive root of α. Therefore, the mutation load is $L = 0 - \langle \bar{m}_\infty \rangle = -\beta(z_1) = U(1 - M_*(z_1))$. Two situations can occur: either α has no such root ($z_1 = \infty$), in which case $L = U$; or it has a root $0 < z_1 < \infty$, in which case $L = U(1 - M_*(z_1)) = U + 1/\omega(z_1)$. As $0 < M_*(z_1) < 1$, the load is then smaller than the mutation rate $0 < L < U$. The first situation ($L = U$) always arises as $U \to 0$. We thus have some form of phase transition in the dependence of the load on mutation rate, as U increases.

Equilibrium fitness variance: We have $\langle V_\infty \rangle = U\, L\, \omega'(0) + U|\mu_s|$, where $\mu_s = M_*'(0)$, as above. Note that the term $\omega'(0) = \partial_m E(s|m)|_{m=0}$ is the slope of the change in the mean of the DFE with m in the vicinity of $m = 0$. It seems likely that in most models, this slope is of the same order as the mean itself: $\omega'(0) = O(|\mu_s|)$. We thus have, *a priori*, $\langle V_\infty \rangle = U|\mu_s|\ (1 + O(L))$ where we have seen that $L \leq U$; therefore, the fitness variance is close to $U|\mu_s|$ at equilibrium

in a vast variety of models (epistatic or not), as long as $U \ll 1$. It is easily checked that the equilibrium for a nonepistatic model with deleterious mutations (see above) is retrieved as a subcase: $\omega(z) = 0$ ($\alpha(z) = 1, z_1 = \infty$) so that $L = U$ and $\langle V_\infty \rangle = U|\mu_s|$.

Spike of optimal genotypes: A spike may exist (Section II.4 in Appendix B, File S1), but only provided the load is $L = U$ and if $\omega'(0) \geq 0$, namely, when maladaptation at most aggravates the mean deleterious effect of mutations (they become more or equally deleterious as the background becomes suboptimal). The spike converges as $U \to 0$ to that of the corresponding nonepistatic model with the same $M_*(\cdot)$. We have $\langle \rho_\infty \rangle \to e^{-U\int_0^\infty M_*(u)du} = e^{-U/s_H}$ where s_H, as before, is the harmonic mean (in absolute value) of the DFE in the optimal background at $m = 0$. Furthermore, whenever $s_H = 0$, the spike is vanishing at equilibrium, for any U. Finally, when $\omega'(0) = 0$ (as in the FGM), the weak mutation limit is also the upper bound $\langle \rho_\infty \rangle \leq e^{-U/s_H}$ for any U.

Some of the qualitative results above are reminiscent of Burger's (2000) propositions 2.1 on p.127 and 5.1 on p.145, proven for a single continuous trait, by a very different approach. It states that, independently of the trait mutational kernel or the trait-to-fitness function, the load (i) converges to $L = U$ as $U \to 0$, (ii) is exactly equal to U whenever a spike exists at the optimum, and (iii) is always less or equal to this limit ($L \leq U,\ U \in \mathbb{R}^+$). This section thus extends this result by providing a general approach to analytically compute these mutation loads, spike heights, and higher moments for *all* U.

FGM

Let us now consider a classic model with diminishing returns epistasis: FGM (Fisher 1930), described in the *Model* section, as an example of log-linear background dependence.

Gaussian FGM: Recall that we denote Gaussian FGM the classic version with a multivariate normal distribution for mutation phenotypic effects, which shows exact log-linear context dependence (Table 2) so that Equation 7 applies.

Trajectories: The fitness mean and variance trajectories over time (predicted by numerically solving Equation 7) are illustrated for a small mutation rate in Figure 3A ($U < U_c$, see below for the definition of the critical value U_c) and a high mutation rate ($U > U_c$) in Figure 3B. They are compared with the average fitness mean and variance in simulations (population size $N = 10^5$). Smaller and larger population sizes and other mutation rates are illustrated in Figures D2 and D3, Appendix D, File S1. The deterministic approximation, here, is accurate across the whole mutation rate spectrum (roughly as long as $NU|\mu_s| \gg 1$). Note that, while the two first derivatives at $z = 0$ (expected mean $\langle \bar{m}_t \rangle = \mathcal{C}_t'(0)$ and variance $\langle V_t \rangle = \mathcal{C}_t''(0)$) are accurately retrieved from the numerical solution of Equation 7, the third order derivative is more problematic to obtain (due to limited machine ε) and would require solving the PDE satisfied by $\mathcal{C}_t'(z)$, together with Equation 7.

Equilibrium: The equilibrium for the Gaussian FGM is a global attractor (by the memoryless property of log-linear background-dependence models, Appendix B, File S1). Its properties are readily derived from the framework in *Equilibrium in the presence of diminishing returns epistasis* (detailed in Appendix D, File S1) and summarized in Table 3 (approximate results for $n \geq 3$ are derived in Appendix E, File S1). Three qualitatively distinct situations arise according to the dimensionality n and mutation rate U, which determine the existence of a finite positive root to α. Consistent with the general results in *Equilibrium in the presence of diminishing returns epistasis*, a phase transition can occur (if $n \geq 3$) at a critical mutation rate U_c, which depends on dimension and scale (explicit formulas in Table 3 and File S1 in Appendix D, Section III). The results are consistent with Waxman and Peck's (1998) conclusions: a spike of optimal genotypes only exists at a low enough mutation rate ($U < U_c$) and in $n \geq 3$ dimensions. Here an exact expression is obtained for the critical mutation rate where the spike vanishes, for the spike height below this threshold, and for the mutation load over the full range of U. Note that explicit expressions for the spike height in $n = 3$ dimensions were also obtained (for a non-Gaussian FGM) in Waxman and Peck (2006), by a different approach.

A simple approximation emerges for the equilibrium fitness distribution when $U < U_c$ in terms of a mixture of a probability mass of optimal genotypes and a negative gamma distribution of suboptimal genotypes, corresponding to a Gaussian FGM in $n - 2$ dimensions (with $n \geq 3$):

$$\text{if } U < U_c: \begin{cases} m = 0, & \text{with probability } \langle \rho_\infty \rangle = e^{-U/s_H}, \\ m \sim -\Gamma\left(\frac{n-2}{2}, \lambda\right), & \text{with probability } 1 - \langle \rho_\infty \rangle \end{cases}. \tag{9}$$

Strikingly, the weight of the spike is *exactly* the same as that in the corresponding nonepistatic model here (gamma DFE), whereas our heuristic analysis (*Equilibrium in the presence of diminishing returns epistasis*) only suggests such convergence in the limit of low mutation rates, in general. The full fitness distribution in Equation 9 is exactly that expected in the absence of epistasis, in the small U/s approximation described in File S1, Appendix C, Section II (Equation C9). A simple pattern thus emerges: for any $U < U_c$, the equilibrium fitness distribution in the FGM is approximately "blind" to the presence of epistasis, and behaves as the equivalent nonepistatic model with DFE given by that of the optimal genotype. We thus retrieve essentially an HOC approximation (Turelli 1984) on fitness itself.

On the other hand, when $U \gg U_c$, a weak selection strong mutation (WSSM) limit (detailed below and in Appendix E, File S1) yields a complementary approximation for the fitness distribution at high mutation rates.

$$\text{if } U \gg U_c : m \sim -\Gamma\left(\frac{n}{2}, \sqrt{U\lambda}\right). \tag{10}$$

Finally, note that the equilibrium higher moments of Equation 7 (exact for the Gaussian FGM) can be studied analytically

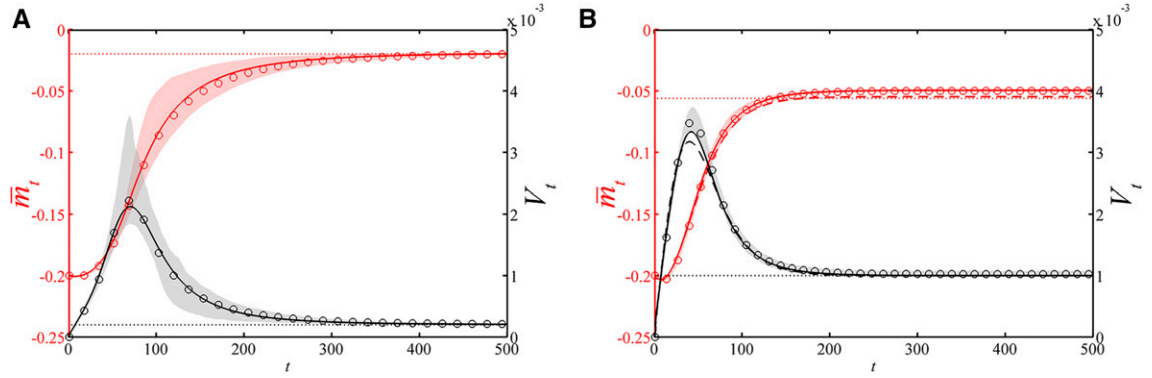

Figure 3 Mean fitness $\bar{m}_t$ and variance V_t trajectories in a Gaussian FGM. (A) $U = 0.02 < U_c$; (B) $U = 0.1 > U_c$. Solid lines: expected trajectories $\langle \bar{m}_t \rangle$ and $\langle V_t \rangle$ given by the numerical solution of Equation 7, with $M_*(z) = (1+\lambda\, z)^{-n/2}$ and $\omega(z) = -\lambda\, z^2/(1+\lambda\, z)$. Dotted lines: equilibria $\langle \bar{m}_\infty \rangle = -L = -U + U(1+\lambda\, z_1)^{-n/2}$ and $\langle V_\infty \rangle = U|\mu_s| = U\, n\, \lambda/2$ given by the analytical theory. Dashed lines (B): expected trajectories from the WSSM approximation (Equation 12 for $\langle \bar{m}_t \rangle$ and B31 for $\langle V_t \rangle$). ○: empirical mean fitness and variance given by individual based simulations, averaged over 10^3 populations ($N = N_e = 10^5$). Shaded regions: 99% C.I.'s for the mean fitness (in red) and the variance (in gray). The parameter values are $n = 6$ traits and $\lambda = 1/300$ ($|\mu_s| = 0.01$), leading to a critical mutation rate $U_c = 16\lambda \simeq 0.05$. We assumed initially clonal populations with $m_0 = -20|\mu_s| = -0.2$.

(Appendix B, File S1) and are very close to the general expressions derived from the linearization in *Equilibrium in the presence of diminishing returns epistasis*.

Generalized FGM: Fisher's original formulation did not specify the shape of the fitness function (linear, quadratic. *etc*.) or the distribution of mutation effects on $\mathbf{g}$ (normal, uniform, *etc*.), except that it must be spherically symmetrical (effects are iid across traits) and centered on the parental phenotype. Keeping the quadratic fitness function, we study a "generalized FGM" (see Appendix E, File S1) with arbitrary spherically symmetric distributions of mutation phenotypic effects ($\mathbf{dg} \sim$ D). A given distribution $\mathcal{D}$ determines a given DFE in the optimal background [a given $M_*(z)$]. The function ω is then $\omega(z) = 2\ z^2 M_*'(z)/(nM_*(z))$, thus allowing the study of equilibria for any choice of $\mathcal{D}$ (Section I, Appendix E, File S1). As an example (Section II, Appendix E, File S1), we derive the equilibrium properties of a model with arbitrary dimension n and negative exponential DFE at the optimum: $s \sim -Exp(1/|\mu_s|)$. In particular, the load is $L = \min(U, n\ \sqrt{U\lambda}/2)$, where $\lambda = 2|\mu_s|/n$ is the mutational variance on each trait (for consistency with the Gaussian FGM).

Weak Selection Strong Mutation (WSSM) approximation: More general and simple results (Section III, Appendix E, File S1) are obtained from a WSSM approximation, more precisely whenever $U \gg \tilde{U}_c = n^2\lambda/4$, where $\lambda = 2|\mu_s|/n$ is the mutational variance on each trait. Note that $\tilde{U}_c \approx U_c$ (for substantial n): it is roughly at the same mutation-rate threshold that equilibria (U_c) and transient dynamics ($\tilde{U}_c$) show a qualitative transition. In the WSSM regime, the mutational kernel is approximately linear in m, so that Equation 4 captures the CGF dynamics, even away from equilibrium. The coefficients are $\alpha(z) = 1 - U\,\lambda\, z^2$ and $\beta(z) = -U\,\lambda\, n/2\, z$.

Equilibrium: As was already stated above (Equation 10), the corresponding equilibrium fitness distribution is a negative gamma: $m \sim -\Gamma(n/2, \sqrt{U\lambda})$. Connecting this approximation with the known value of the load at lower mutation rates, $L = U$ provides a simple expression covering all the range of U :

$$L \approx \min(U, n\sqrt{U\lambda}/2), \tag{11}$$

with a phase transition at $U = \tilde{U}_c = \lambda/4n^2$. The accuracy of this simple result is illustrated in Figure 4A, where the load is shown for single replicate simulations over a range of U. We simulated two alternative models (Gaussian FGM with a gamma DFE and an inverse Gaussian DFE), scaled to the same value of $|\mu_s|$ and with the same dimensionality n. Both models yield the same results, accurately captured by Equation 11. The spike of optimal genotypes is shown in Figure 4B for the same simulations: here, all genotypes pertaining to an effectively neutral fitness class relative to the optimum ($-1/N \le m \le 0$) were counted as "under the spike." As expected by theory, the spike weight is approximately e^{-U/s_H}, where s_H differs between the two models (Gaussian or inverse Gaussian).

Trajectories: The analytic solution (Equation 5) applied to the WSSM approximation can be equated, at all times, to a known explicit distribution, depending on the initial condition. The corresponding distribution of the underlying phenotype is also explicit, and, in all cases, happens to be multivariate Gaussian (with time-varying variances and means). Therefore, the WSSM approximation exactly matches Kimura's (1965) and Lande's (1980) Gaussian approximation for traits at equilibrium, and extends it away from equilibrium. Indeed, although obtained in very different manners, these two approaches rely on qualitatively the same WSSM assumption. Lande already conjectured that this approximation was mostly independent of the underlying distribution of mutation effects on phenotype, and should be valid away from equilibrium, as the dynamics of phenotypic

Table 3 Mutation-selection balance properties in the Gaussian FGM

n	U_c	z_1	Load	$V(m)$	Spike height
1	∞	$< \infty$	$< U$	$U\lvert\mu_s\rvert$	0
2	λ	$\begin{cases} \infty, & U < U_c \\ 1/(\sqrt{U\lambda} - \lambda), & U \geq U_c \end{cases}$	$\begin{cases} U, & U < U_c \\ \sqrt{U\lambda}, & U \geq U_c \end{cases}$	$U\lvert\mu_s\rvert$	0
≥ 3	$\approx (n^2\lambda/4)$	$\begin{cases} \infty, & U < U_c \\ \approx 1/\sqrt{U\lambda}, & U \geq U_c \end{cases}$	$\begin{cases} U, & U < U_c \\ \approx n/2\sqrt{U\lambda}, & U \geq U_c \end{cases}$	$U\lvert\mu_s\rvert$	$\begin{cases} e^{-U/s_H}, & U < U_c \\ 0, & U \geq U_c \end{cases}$

Here $\mu_s = E(s|m) = E(s|0) = -\ n\lambda/2$ is the arithmetic mean of the DFE and $s_H = 1/|E(1/s|0)| = 1/(\lambda(n/2-1))$ is the harmonic mean of the DFE. $\approx$ notifies that this is an approximate result (Appendix E, File S1).

variance are then independent of the mean (equation 19 in Lande 1980). Here, the result arises explicitly as a WSSM limit of a generalized FGM. The present approach extends these former results to fitness (and trait) dynamics where the phenotypic variance is not constant, and provides an explicit threshold ($U \gg \tilde{U}_c = n^2\lambda/4$), beyond which it is accurate. All results are given in Appendix E, File S1, here we only detail the mean fitness trajectories.

Adaptation from a clone: For a population started with a clone at given fitness $m_0 \leq 0$, the mean fitness trajectory, given by Equation 6, is

$$\langle \bar{m}_t \rangle = \mathcal{C}_t'(0) \approx -\frac{n}{2}\sqrt{U\lambda}\tanh\left(t\sqrt{U\lambda}\right) + \operatorname{sech}\left(t\sqrt{U\lambda}\right)^2 m_0. \quad (12)$$

This WSSM approximation is illustrated in Figure 3B (dashed lines) and proves fairly accurate even when U is only mildly superior to $\tilde{U}_c$ ($U = 2\,\tilde{U}_c$ in this example). The corresponding trajectory of the full fitness and phenotype distributions are illustrated in File S4 (Movie 2, A and B, respectively), showing the agreement between simulations and theory, for a single replicate. The characteristic time of this fitness trajectory is the time $t_{0.99}$ taken to fulfill 99% of the trajectory. Strikingly, it is independent of the details of the model: $t_{0.99} \approx 3/\sqrt{U\lambda}$. In particular, it is independent of the distance to the optimum (m_0) : it takes roughly the same time to reach equilibrium from an optimal or a highly suboptimal clone in the WSSM regime.

Adaptation from an equilibrium population: In a similar manner, we may consider a population starting at equilibrium, undergoing a sudden shift in the optimum and responding to this new environment, this time with standing variance available. Here too, the whole fitness and phenotype distributions are explicit over time, including with a change in U or λ between the former and new environments. If the shift only affects the optimum (not U or λ) and is such that the mean phenotype shows a fitness lag m_0 (mean fitness is then $\bar{m}_0 = m_0 - \sqrt{U\lambda}n/2$), then

$$\langle \bar{m}_t \rangle = \mathcal{C}_t'(0) \approx m_0\, e^{-2t\sqrt{U\lambda}} - \sqrt{U\lambda}\, n\Big/2. \quad (13)$$

The trajectory of the fitness and phenotype distributions are illustrated in File S4 (Movie 3, A and B, respectively), with an additional doubling of the mutation rate in the new environment. Here too, the characteristic time is independent of the distance to the optimum $t_{0.99} \approx 2.3/\sqrt{U\lambda}$, and it is only mildly shorter than the characteristic time in the absence of standing variance. In all cases, the characteristic times scale with $1/\sqrt{U\lambda}$, showing that the "cost of complexity" well known in the FGM (Orr 2000), is only mediated by $\sqrt{U\lambda} = \sqrt{2U|\mu_s|/n}$ in this regime. When comparing different dimensionalities n, if we scale λ to the same $|\mu_s|$, complexity slows down adaptation as $1/\sqrt{n}$. Otherwise, simply adding traits with the same variance λ does not affect the characteristic time, it simply increases the mutation load as $L = n\sqrt{U\lambda}/2$.

Convergence to the deterministic approximation

Our simulations, which included full stochasticity (individual-based model) showed good agreement with the theory in Equation 3, which ignores drift. This seems to hold over either effectively infinite timescales (*e.g.*, FGM, Figure 3 and Figure 4, and other models illustrated in File S3), over very long timescales (nonepistatic models with purely deleterious mutations, Figure C1, File S1), or over only a few 100 or 1000 generations (nonepistatic models with beneficial mutations, Figures C2–C5, File S1). Accuracy also seems to increase as NU gets larger for the models and parameters we explored. It has indeed long been observed that deleterious mutation models or models with an optimum could be handled reasonably well by deterministic population genetics. This then raises the question of why the deterministic approximation ultimately breaks down with nonepistatic models, whereas it does not seem to do so with diminishing returns epistasis.

This can obviously be tackled by individual-based simulations for any given model. In the case of nonepistatic models, analytical studies have also pointed to a complex interplay of drift and other forces in the mid- to long-term behavior of asexual models (*e.g.*, Desai and Fisher 2007): the importance of the "stochastic edge" of the fitness distribution (Brunet *et al.* 2008) depends on whether this edge is stochastic or not (highly populated or not). The present treatment provides some hint on the issue by looking at the term neglected in the deterministic dynamics: $\delta_t(z)$ in Equation 1. This "stochastic source term" is negative, vanishes at $z = 0$ but increases with z (see Section III, Appendix B, File S1). That

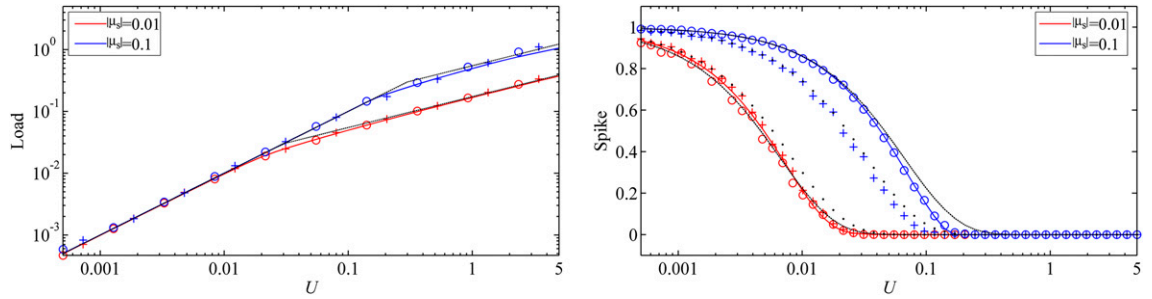

Figure 4 (A) Mutation load L and (B) spike ρ_∞ as a function of mutation rate U : with two values of $|\mu_s|$ (0.1 and 0.01) and with the standard Gaussian FGM or an FGM with inverse Gaussian DFE at the optimum (IG FGM). Solid red and blue lines: numerical values obtained with Equation 7 for the Gaussian FGM (estimated at a large time $T = 10^3$); the (A) load was computed as $-C_T'(0)$ and (B) the expected spike as $e^{C_T(900)}$. (A) Black dashed lines: analytic approximations $L \simeq \min(U, n/2\sqrt{U\lambda})$ (Equation 11). (B) Black dashed or dotted lines: $\langle \rho_\infty \rangle \simeq e^{-U/s_H}$, where s_H is the harmonic mean of the DFE (in absolute value) at the optimum (Gaussian or IG FGM respectively, Equation D8, File S1). Circles (Gaussian FGM) and crosses (IG FGM): simulated values of the mutation load and of spike at time $T = 10^3$ given by individual-based simulations of a single population ($N = N_e = 10^5$). The parameter values are $n = 6$ traits and $|\mu_s| = 0.01$ or 0.1. The inverse Gaussian distribution has mean $|\mu_s|$ and shape parameter 0.05.

$\delta_t(z) < 0$ means that the deterministic prediction overestimates the cumulants (for example, the expected mean fitness is actually bellow the deterministic prediction). That $\delta_t(z)$ is small around $z = 0$, means that the current error on the bulk of the distribution (the first derivatives of $\delta_t(\cdot)$ at $z = 0$) is limited. On the other hand, because of the transport term $C_t'(z)$ in Equation 3, the larger error $|\delta_t(z)|$ for a large z progressively affects the accuracy of the deterministic approximation around $z = 0$ (hence the bulk itself) at later times. Intuitively, this reflects the fact that sampling (drift) induces relatively more stochastic variation in the extrema than in the mean and variance of a distribution: the maximum can be very important for the long-term rate of adaptation (stochastic edge; Brunet *et al.* 2008), while the mean and variance influence the short-term "bulk" dynamics.

Whether and when a substantial deviation will accumulate depends on the details of the model, and can be difficult to quantify. However, in the case of linear background dependence (Equation 4) some general insight can be obtained, focusing on mean fitness trajectories. The relative deviation between the "exact" expected mean fitness $\langle \bar{m}_t \rangle = \langle C_t'(0) \rangle$ and that predicted by the deterministic approximation $C_t'(0)$, has an explicit upper bound at all times:

$$\frac{|C_t'(0) - \langle \bar{m}_t \rangle|}{|\langle \bar{m}_t \rangle|} \leq \frac{1}{|\langle \bar{m}_t \rangle|} \int_0^t w_t(v) \left\langle \frac{|\bar{m}_v|}{N_e \; p_{\max}(v)} \right\rangle dv. \qquad (14)$$

Here, $w_t(v) = \alpha(y(t-v)) = y'(t-v)$ is a weight which depends on the form of epistasis (via y), see Equation 4 and the paragraph below. Roughly, if $|\bar{m}_v|$ and $|\bar{m}_t|$ are of comparable order, the relative error is proportional to (i) t/N_e, and (ii) to a weighted mean over the period $(0, t)$ of the expected inverse frequency (across replicates) of the fittest class. Equation 14 provides some intuition on how and why different mutation models deviate from the deterministic prediction. We treat each in turn.

Nonepistatic models

The weights are then $w_t(v) = \alpha(y(t-v)) = 1$, so the error must accumulate over time. With purely deleterious mutation models, $p_{\max}(t)$ remains large for a long time ($p_{\max}(t) \geq e^{-U/s_H}$ in the deterministic approximation), and it can be shown (Section III, Appendix C, File S1) that the relative error in Equation 14 remains ≤ 1 for some "time to loss of accuracy" of order $N_e \; e^{-U/s_H}$ (see Table 2). The "characteristic time" to reach 95% of equilibrium ($t_{0.95} \leq 8/s_H$, see *Application*, *Nonepistatic models*) is therefore often much less than the timescale over which the deterministic approximation breaks down and Muller's ratchet starts to click (of order $N_e \; e^{-\; U/s_H}$, Table 2).

With beneficial mutations, however, the fittest class consists of a small number of fit mutants so the error accumulates much faster. Furthermore, as the error depends on inverse frequencies of the fittest class, the fluctuations of this stochastic edge (across replicates and times), especially through smaller values, are important; a fact already pointed out for these models (Hallatschek 2011).

Diminishing returns epistatic models

With diminishing returns, two effects alleviate the deviation. First, mere intuition suggests that, as there is a reachable fitness upper bound, this fitness edge should ultimately become highly populated ($p_{\max}(t) \gg 1/N$) after sufficient time. This remains a verbal argument. Second, beyond the critical mutation rate threshold [whenever $\alpha(\cdot)$ has a finite root], the weights $w_t(v)$ in Equation14 vanish as $t \to \infty$. This implies that the error ultimately becomes independent of the earlier dynamics of $p_{\max}(v)$ and remains bounded by a constant independent of t (see Appendix B, part III.2, File S1). This explains why these models are always accurately captured by the deterministic approximation at large times (see Figure 4 on equilibrium states), even when a substantial deviation from the deterministic trajectory builds up transiently (as observed, *e.g.*, in Figure D2, File S1, with $U = 0.0002$, $NU = 2$). Intuitively (without formal proof), we expect the transient error to be larger with smaller NU and when starting from a strongly maladapted population, as the fittest class may be small for a long time.

Qualitatively, this absence of accumulation of deviation over large times is a key difference introduced by epistasis. The result is reminiscent of Poon and Otto's (2000), which showed that even a minimal amount of compensating mutations can stop Muller's ratchet. A substantial transient deviation may arise at intermediate times, but it ultimately shrinks again.

Discussion

The proposed approach models the dynamics of fitness distributions in the presence of selection and mutation (neglecting drift) in large asexual populations with a variety of DFEs. A deterministic PDE arises as an approximation to the dynamics of the expectation (over stochastic replicates) of the CGF of the fitness distribution. This allows to easily handle clonal interference between cosegregating mutants (drawn from various classes of mutation models), and the contribution from standing variance, at or away from stationary regimes.

Main results and possible empirical tests

When considering only the contribution from standing variance (negligible contribution from *de novo* mutation), Equation 8 with $U = 0$ predicts the full fitness distribution over time from an arbitrary initial condition. This provides a versatile model for the response to selection of large polymorphic asexual populations over short timescales, *i.e.*, before new mutations contribute to adaptation. The predicted trajectories are highly testable in experimental evolution: it only requires a measurement of the initial fitness distribution. We hope it may foster empirical tests of adaptive dynamics from standing variance in model asexual organisms, with a potential for faster observable responses than when a single clone adapts by new mutations.

The use of CGFs also simplifies the treatment of nonepistatic models with fairly general DFEs (Figure 2 and Figures C2–C5, File S1). For a nonepistatic deleterious mutation, most previous results are retrieved as a subcase (see *Application*, *Nonepistatic models*). We further find that the fitness distribution admits explicit (testable) form over time (Appendix C, Figure C1 in File S1; and Movie 1, A and B, in File S4) and that the timescale to reach equilibrium from an optimal clone is independent of the mutation rate U (and of order $1/|\mu_s|$), which is easily smaller than that over which the deterministic approximation breaks down.

When nonepistatic beneficial mutations are added, the approach breaks down over shorter timescales (detailed in *Model* section, *Application*, *Nonepistatic models*, and Table 2). In general, the deterministic approximation breaks (after some time) when the fittest class is only represented by a few copies (see Section III, Appendix B, File S1), forming a stochastic edge (Brunet *et al.* 2008). However, in this case, we observe by simulation that Equation 8 still provides a rough connection (Figure 2) between the early regime of adaptation (deterministic), and the ultimate stationary regime (stochastic). Because Equation 8 easily handles a wide variety of DFEs (*e.g.*, including beneficial and deleterious mutations) that are not easily treated in the stationary stochastic regime, it may also be used as a more general null model over shorter empirical timescales (albeit still ignoring epistasis).

The same framework can be applied to mutation kernels showing diminishing returns epistasis (*Equilibrium in the presence of diminishing returns epistasis* and *FGM* in *Application*). In that case, the discrepancy with the deterministic approximation remains bounded (sometimes very small) at all times (Figure 3, Figure 4, and File S3), because the fittest class is rapidly filled with a substantial number of selected mutants. The fitness distribution and the proportion of optimal genotypes at equilibrium then take testable explicit forms (*Equilibrium in the presence of diminishing returns epistasis in Application* and Figure 4) in a variety of diminishing returns epistasis models where beneficial mutations compensate suboptimal genotypes. Overall, our most robust prediction (both with and without epistasis) at equilibrium is that fitness variance should be close to $\langle V_\infty \rangle \approx U|\mu_s| + o(U)$, whenever $U \ll 1$. This is also testable (given a large population at equilibrium), as the product $U|\mu_s|$ can be directly estimated from mutation-accumulation experiments (reviewed, *e.g.*, in Keightley and Eyre-Walker 1999). It is also easier to estimate the fitness variance (and possibly skewness, *etc.*) than the mutation load, as the latter requires an estimate of the maximal fitness. Such an estimate would only be possible if optimal genotypes were frequent (not always the case), or given a particular model for the equilibrium fitness distribution (*e.g.*, Equation 9 and Equation 10), which depends on the assumed DFE at the optimum.

The approach via CGFs is also particularly well suited for the Gaussian FGM with normally distributed mutant phenotypes. This model has recently served as a landmark null model of context-dependent DFE (background and/or environment dependence, Tenaillon 2014). It has also long been a landmark tool in evolutionary ecology and quantitative genetics: most treatments of the adaptive and demographic responses to environmental challenges, or of the distribution of phenotypes under stabilizing selection, are based on its assumptions (Tenaillon 2014). Under this Gaussian FGM, the fitness dynamics (averaged over replicates) are fully captured by a single PDE (Equation 7 and Figure 3) covering the full mutation rate spectrum. Known analytical treatments of this model mostly described equilibria under two extreme regimes: in the limit $U \ll |\mu_s|$ with $n = 1$ dimension (Turelli 1984) or in the limit $U \gg |\mu_s|$ with arbitrary n (Lande 1980). Here, the full fitness distribution at or before equilibrium is predicted (analytically or numerically by solving Equation 7) for all U and arbitrary n (Figure 3, Figure 4; Appendix D in File S1; and Movie 2 and Movie 3 in File S4). This yields a fully testable pattern to fit to observed fitness distributions or mean fitness dynamics.

Finally, the results extend to arbitrary (spherically symmetrical) distributions of mutant phenotypes in a WSSM limit ($U \gg U_c \approx n^2\lambda/4$). In this limit, both traits and fitness, at all

times, converge to simple analytic distributions, independently of the details of mutational effects. This limit (Figure 3B and Appendix E in File S1) arises here as a diffusion approximation in fitness space, and corresponds to normally distributed phenotypes (with time-varying mean and variance), consistent with Kimura's (1965) derivation at equilibrium in one dimension, and Lande's (1980) conjecture for multiple dimensions. The approach extends these theories away from equilibrium and clarifies the threshold mutation rate ($\tilde{U}_c$) where they apply. These trajectories are also highly testable. Indeed, (i) the full distribution is analytic at all times from known initial condition (it may be applied on short experimental timescales), and (ii) the FGM can be parameterized (Martin and Lenormand 2006) from data on deleterious mutation effects ($|\mu_s|$ and n)and rates (U), which are more readily available to the experimenter than beneficial mutation kernels and rates.

The evolutionary process inherent in the FGM is complex in large asexual populations and at high mutation rates: it includes clonal interference, both deleterious and beneficial mutations and pervasive epistasis. Yet, the resulting fitness trajectories in the WSSM limit (Equation 12 and Equation 13) display surprisingly simple and robust patterns, independently of the details of the underlying mutational process. In particular, the mean fitness (at any time away from equilibrium) scales simply with the initial maladaptation: $\langle \bar{m}_t \rangle \approx m_0 \ \mathrm{sech}(\sqrt{U\,\lambda}\ t)^2$, see Equation 12. This latter pattern is, at least qualitatively, in agreement with the observation that the cumulative mean fitness increase (over stochastic replicates and between distant generations) scales almost linearly with initial maladaptation ($\langle \bar{m}_t \rangle \propto m_0$). Couce and Tenaillon (2015) recently showed this empirical pattern to hold across several species and data sets, and suggested that the FGM may be one among several models yielding such linearity. The present analytic treatment might allow us to go beyond qualitative analyses and perform quantitative tests based on known parameters. A test of the FGM and other models would (ideally) require confronting full observed trajectories with (independently parameterized) predictions. We hope that the proposed approach may help such quantitative testing. Deriving (approximate or exact) analytic solutions to Equation 7 away from the WSSM limit would also be useful in this regard, but requires further effort.

Finally, and although not detailed here, other epistatic models can be predicted analytically (Equation 4) or numerically (Equation 7) through the proposed framework. Two such examples are summarized in Table 2: Kingman's (1978) HOC model (Equation 7) and a simplified version of Rouzine *et al.*'s (2003) binary model (Equation 4). Evaluating how accurate the predictions are, depending on the models and parameters, requires extensive simulation work beyond the scope of this article, but the necessary tools (and illustrations of the accuracy) are provided in File S3.

Limits

The model obviously has several limits; first of all, not all equations proposed here can be solved analytically (Equation 7) and we must then rely on numerical solutions. But more fundamental issues can be raised about the approach itself. We detail them below and discuss how to improve these aspects.

Genetic drift and clonal interference: Drift is explicitly modeled in Appendix A, File S1, but only to determine the error implied by neglecting it (Equation 14). Our results suggest that if the fittest class of genotypes quickly reaches (and remains at) a substantial frequency, the deterministic approximation is accurate even over the long term (see the section on *Convergence to the Deterministic Approximation*). This is typically what occurs with diminishing returns epistatic models (where fitness is bounded from above), which also prove to have a memoryless property that makes them less prone to accumulate stochastic deviations.

During adaptation over a single peak landscape, clonal interference is pervasive (multiple asexual lineages compete for fixation); yet, modeling the stochastic fate of each mutant does not prove critical in this model. Conversely, in similar conditions, it proves critical to do so with nonepistatic models of beneficial mutation, at least over long timescales. Overall, clonal interference need not always be described in the presence of drift: nonepistatic models with beneficial mutations, most studied in this context (Sniegowski and Gerrish 2010), happen to be a case where it is particularly important to do so. From an empirical perspective, it is simpler to avoid a theory that requires details of the genetic drift process, as the relevant parameters are notoriously difficult to measure (N_e, the stochastic reproductive variance which may vary between genotypes, *etc.*). However, a proper treatment of the effect of stochastic forces (drift and mutations) would still be useful even in models where the expected trajectory is robust to their effect: it would provide envelopes around the deterministic expectation. Models of stochastic fronts and cutoffs may be used once translated into CGF dynamics, or stochastic PDEs using the results of Appendix A, File S1.

Segregation and recombination: Asexuals are our focus here, because they form the vast majority of model organisms in experimental evolution, for which this work is intended. However, sex is the norm *in natura* and will also likely become increasingly more studied empirically. The approach by CGFs was originally designed to handle recombinant genomes (Burger 1991), as the CGF from independent loci add up, providing simple extensions. Indeed, some of our results naturally extend to sexuals in simplified situations (not detailed here). However, fitness is typically nonadditive across loci, so that simple additive theory may prove inaccurate in more realistic models.

Substitution data: The present model directly follows fitness dynamics, without explicitly modeling substitutions at the molecular level. They do occur (an allele becomes dominant, then another takes over, *etc.*), but their dynamics may be complex (cosegregating alleles). By not requiring an explicit description of these dynamics, fitness trajectories in

nonstationary regimes with complex epistatic models can be handled. Yet, this is at the cost of providing no information on the underlying genetic basis of adaptation (which is now partly available empirically). For some important models, possibly epistatic but with low polymorphism, these underlying dynamics may be inferred from backward modeling. However, regimes with high polymorphism might show more complex molecular signatures, especially away from stationary regimes. The proposed framework may generate alternative coalescent models suited for epistatic, nonstationary models, just as traveling wave models have been successfully used (Good *et al.* 2014) for nonepistatic models at a stationary regime.

More complex environments and landscapes: The models considered here mostly assumed a fixed environment in which adaptation occurs, as is typical in most theories of adaptation (Orr 2005), and as is relevant to many experimental evolution settings. However, more complex situations are of interest: multiple environments connected by migration, a continuously changing environment with a moving optimum, trade-off in life history traits. In some cases, these can be expressed as an adaptive process on multiple fitness components, and may then be handled by considering the dynamics of a multivariate CGF, describing the joint distribution of these components. Also, trait-based landscapes where traits are not equivalent for selection and/or mutation (*e.g.*, anisotropic FGM) are not handled by the model as such. Indeed, the DFE is then not only dependent on the background fitness alone (distance to the optimum), but also on additional details (direction to the optimum). These can also be handled by introducing multivariate CGFs, describing the joint fitness contributions from each phenotypic dimension. We believe PDEs for such multivariate CGF dynamics can be written for many important classes of models where multiple fitness components interact. The open question will more likely be whether they can yield analytical insight.

Conclusion

We believe theoretical tools are now available that provide "null" adaptation models, which may be quantitatively confronted to experimental evolution data (including those with standing variance, rarely studied in these experiments). Such tests of basic process predictions are necessary if we are ever to apply our theories quantitatively, into the wild, or into the human body.

Acknowledgments

This work was fostered by discussions with François Hamel, Marie-Eve Gil, Phillip Gerrish, Thomas Lenormand, David Waxman, Sylvain Gandon, Ophélie Ronce, Jimmy Garnier, and Luis Miguel Chevin. This work was funded by an Institut National de la Recherche Agronomique grant "MEDIA network" and by the French National Research Agency (projects NONLOCAL, ANR-14-CE25-0013 and MeCC, ANR-13-ADAP-0006) and the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n.321186—ReaDi.

Literature Cited

Barton, N. H., 1998 The geometry of adaptation. Nature 395: 751–752.

Brunet, E., I. M. Rouzine, and C. O. Wilke, 2008 The stochastic edge in adaptive evolution. Genetics 179: 603–620.

Burger, R., 1991 Moments, cumulants, and polygenic dynamics. J. Math. Biol. 30: 199–213.

Burger, R., 1998 Mathematical properties of mutation-selection models. Genetica 103: 279–298.

Burger, R., 2000 *The Mathematical Theory of Selection*, *Mutation*, *Recombination*. John Wiley & Sons, Chichester, United Kingdom.

Burger, R., and J. Hofbauer, 1994 Mutation load and mutation-selection-balance in quantitative genetic-traits. J. Math. Biol. 32: 193–218.

Chou, H. H., H. C. Chiu, N. F. Delaney, D. Segre, and C. J. Marx, 2011 Diminishing returns epistasis among beneficial mutations decelerates adaptation. Science 332: 1190–1192.

Couce, A., and O. A. Tenaillon, 2015 The rule of declining adaptability in microbial evolution experiments. Front. Genet. 6: 99.

Desai, M. M., 2013 Statistical questions in experimental evolution. J. Stat. Mech. 2013: P01003.

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., and D. S. Fisher, 2011 The balance between mutators and nonmutators in asexual populations. Genetics 188: 997–1014.

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: 385–394.

Dwyer, J. P., 2012 The dynamics of adapting, unregulated populations and a modified fundamental theorem. J. R. Soc. Interface 10: 20120538.

Eigen, M., 1971 Self-organization of matter and evolution of biological macromolecules. Naturwissenschaften 58: 465–523.

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

Frank, S. A., 2014 Generative models *vs.* underlying symmetries to explain biological pattern. J. Evol. Biol. 27: 1172–1178.

Gerrish, P., 2001 The rhythm of microbial adaptation. Nature 413: 299–302.

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

Gerrish, P. J., and P. D. Sniegowski, 2012 Real time forecasting of near-future evolution. J. R. Soc. Interface 9: 2268–2278.

Good, B. H., and M. M. Desai, 2013 Fluctuations in fitness distributions and the effects of weak linked selection on sequence evolution. Theor. Popul. Biol. 85: 86–102.

Good, B. H., and M. M. Desai, 2014 Deleterious passengers in adapting populations. Genetics 198: 1183–1208.

Good, B. H., and M. M. Desai, 2015 The impact of macroscopic epistasis on long-term evolutionary dynamics. Genetics 199: 177–190.

Good, B. H., I. M. Rouzine, D. J. Balick, O. Hallatschek, and M. M. Desai, 2012 Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proc. Natl. Acad. Sci. USA 109: 4950–4955.

Good, B. H., A. M. Walczak, R. A. Neher, and M. M. Desai, 2014 Genetic diversity in the interference selection limit. PLoS Genet. 10: e1004222.

Gordo, I., and P. R. A. Campos, 2012 Evolution of clonal populations approaching a fitness peak. Biol. Lett. 9: 20120239
Hallatschek, O., 2011 The noisy edge of traveling waves. Proc. Natl. Acad. Sci. USA 108: 1783–1787.
Hansen, T. F., 1992 Selection in asexual populations - an extension of the fundamental theorem. J. Theor. Biol. 155: 537–544.
Hietpas, R. T., C. Bank, J. D. Jensen, and D. N. A. Bolon, 2013 Shifting fitness landscapes in response to altered environments. Evolution 67: 3512–3522.
Johnson, T., 1999 The approach to mutation-selection balance in an infinite asexual population, and the evolution of mutation rates. Proc. R. Soc. Lond. B Biol. Sci. 266: 2389–2397.
Keightley, P. D., and A. Eyre-Walker, 1999 Terumi Mukai and the riddle of deleterious mutation rates. Genetics 153: 515–523.
Khan, A. I., D. M. Dinh, D. Schneider, R. E. Lenski, and T. F. Cooper, 2011 Negative epistasis between beneficial mutations in an evolving bacterial population. Science 332: 1193–1196.
Kimura, M., 1965 A stochastic model concerning maintenance of genetic variability in quantitative characters. Proc. Natl. Acad. Sci. USA 54: 731–736.
Kingman, J. F. C., 1978 Simple-model for balance between selection and mutation. J. Appl. Probab. 15: 1–12.
Kryazhimskiy, S., G. Tkacik, and J. B. Plotkin, 2009 The dynamics of adaptation on correlated fitness landscapes. Proc. Natl. Acad. Sci. USA 106: 18638–18643.
Lande, R., 1979 Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33: 402–416.
Lande, R., 1980 The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94: 203–215.
Manna, F., G. Martin, and T. Lenormand, 2011 Fitness landscapes: an alternative theory for the dominance of mutation. Genetics 189: 923–937.
Manna, F., R. Gallet, G. Martin, and T. Lenormand, 2012 The high-throughput yeast deletion fitness data and the theories of dominance. J. Evol. Biol. 25: 892–903.
Martin, G., 2014 Fisher's geometrical model emerges as a property of large integrated phenotypic networks. Genetics 197: 237–255.
Martin, G., and S. Gandon, 2010 Lethal mutagenesis and evolutionary epidemiology. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365: 1953–1963.
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: 893–907.
Martin, G., S. F. Elena, and T. Lenormand, 2007 Distributions of epistasis in microbes fit predictions from a fitness landscape model. Nat. Genet. 39: 555–560.
McCandlish, D. M., C. L. Epstein, and J. B. Plotkin, 2014 The inevitability of unconditionally deleterious substitutions during adaptation. Evolution 68: 1351–1364.
Miralles, R., A. Moya, and S. F. Elena, 2000 Diminishing returns of population size in the rate of RNA virus adaptation. J. Virol. 74: 3566–3571.
Muller, H. J., 1932 Some genetic aspects of sex. Am. Nat. 66: 118–138.
Neher, R. A., and O. Hallatschek, 2013 Genealogies of rapidly adapting populations. Proc. Natl. Acad. Sci. USA 110: 437–442.
Orr, H. A., 2000 Adaptation and the cost of complexity. Evolution 54: 13–20.
Orr, H. A., 2005 The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6: 119–127.
Otto, S., and N. Barton, 2001 Selection for recombination in small populations. Evolution 55: 1921–1931.
Poon, A., and S. P. Otto, 2000 Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54: 1467–1479.
Rattray, M., and J. L. Shapiro, 2001 Cumulant dynamics of a population under multiplicative selection, mutation, and drift. Theor. Popul. Biol. 60: 17–31.
Rouzine, I. M., J. Wakeley, and J. M. Coffin, 2003 The solitary wave of asexual evolution. Proc. Natl. Acad. Sci. USA 100: 587–592.
Roze, D., and A. Blanckaert, 2014 Epistasis, pleiotropy, and the mutation load in sexual and asexual populations. Evolution 68: 137–149.
Sniegowski, P. D., and P. J. Gerrish, 2010 Beneficial mutations and the dynamics of adaptation in asexual populations. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365: 1255–1263.
Sousa, A., S. Magalhaes, and I. Gordo, 2011 Cost of antibiotic resistance and the geometry of adaptation. Mol. Biol. Evol. 29: 1417–1428.
Tenaillon, O., 2014 The Utility of Fisher's Geometric Model in Evolutionary Genetics. Annu. Rev. Ecol. Evol. Syst. 45: 179–201.
Tenaillon, O., O. K. Silander, J. Uzan, and L. Chao, 2007 Quantifying organismal complexity using a population genetic approach. Plos One 2: e217.
Trindade, S., L. Perfeito, and I. Gordo, 2010 Rate and effects of spontaneous mutations that affect fitness in mutator Escherichia coli. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365: 1177–1186.
Trindade, S., A. Sousa, and I. Gordo, 2012 Antibiotic resistance and stress in the light of Fisher's model. Evolution 66: 3815–3824.
Tsimring, L. S., H. Levine, and D. A. Kessler, 1996 RNA virus evolution via a fitness-space model. Phys. Rev. Lett. 76: 4440–4443.
Turelli, M., 1984 Heritable genetic-variation via mutation selection balance - lerch zeta meets the abdominal bristle. Theor. Popul. Biol. 25: 138–193.
Waxman, D., and J. R. Peck, 1998 Pleiotropy and the preservation of perfection. Science 279: 1210–1213.
Waxman, D., and J. R. Peck, 2006 The frequency of the perfect genotype in a population subject to pleiotropic mutation. Theor. Popul. Biol. 69: 409–418.
Wilke, C. O., 2005 Quasispecies theory in the context of population genetics. BMC Evol. Biol. 5: 44.

Communicating editor: J. Hermisson

File S1 All appendices with analytical derivations and proofs. (.pdf, 2.05 MB)

Available for download as a .pdf file at
www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.187385/-/DC1/FileS1.pdf

File S2 .rar archive containing the MATLAB code for numerically solving and simulating the FGM. (.rar, 20 KB)

Available for download as a .rar file at
www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.187385/-/DC1/FileS2.rar

File S3 Mathematica(R) notebook illustrating each sub-case in Table 2: analytical trajectories, numerical solver for Eq. (7), simulation code and comparison of simulation vs. theory. (.zip, 591 KB)

Available for download as a .zip file at
www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.187385/-/DC1/FileS3.zip

File S4 Movies showing the dynamics of fitness (and trait) distributions over time, in context-independent models (movies 1A and 1B, fitness distributions) or FGM: movies 2 A & B (starting from a clonal population) and movies 3 A & B (starting from an equilibrium population). (.rar, 1.63 MB)

Available for download as a .rar file at
www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.187385/-/DC1/FileS4.rar