DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Mutation Induced Extinction in Finite Populations: Lethal Mutagenesis and Lethal Isolation

Citation	Wylie, C. Scott, and Eugene I. Shakhnovich. 2012. Mutation induced extinction in finite populations: lethal mutagenesis and lethal isolation. PLoS Computational Biology 8(8): e1002609.
Published Version	doi:10.1371/journal.pcbi.1002609
Accessed	February 19, 2015 10:48:38 AM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:10636302
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)

Mutation Induced Extinction in Finite Populations: Lethal Mutagenesis and Lethal Isolation

C. Scott Wylie¤*, Eugene I. Shakhnovich

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts, United States of America

Abstract

Reproduction is inherently risky, in part because genomic replication can introduce new mutations that are usually deleterious toward fitness. This risk is especially severe for organisms whose genomes replicate "semi-conservatively," e.g. viruses and bacteria, where no master copy of the genome is preserved. Lethal mutagenesis refers to extinction of populations due to an unbearably high mutation rate (U), and is important both theoretically and clinically, where drugs can extinguish pathogens by increasing their mutation rate. Previous theoretical models of lethal mutagenesis assume infinite population size (N). However, in addition to high U, small N can accelerate extinction by strengthening genetic drift and relaxing selection. Here, we examine how the time until extinction depends jointly on N and U. We first analytically compute the mean time until extinction (τ) in a simplistic model where all mutations are either lethal or neutral. The solution motivates the definition of two distinct regimes: a survival phase and an extinction phase, which differ dramatically in both how τ scales with N and in the coefficient of variation in time until extinction. Next, we perform stochastic population-genetics simulations on a realistic fitness landscape that both (i) features an epistatic distribution of fitness effects that agrees with experimental data on viruses and (ii) is based on the biophysics of protein folding. More specifically, we assume that mutations inflict fitness penalties proportional to the extent that they unfold proteins. We find that decreasing N can cause phase transition-like behavior from survival to extinction, which motivates the concept of "lethal isolation." Furthermore, we find that lethal mutagenesis and lethal isolation interact synergistically, which may have clinical implications for treating infections. Broadly, we conclude that stably folded proteins are only possible in ecological settings that support sufficiently large populations.

Introduction

On average, mutations hurt organismal fitness, e.g. by destabilizing proteins. Thus, left unchecked, new mutations tend to erode fitness and endanger the long-term survival of any species. Fortunately, natural selection usually balances against mutational genetic decay by rewarding the fit and weeding out the unfit. However, when the genomic mutation rate (i.e. the expected number of mutations per genome duplication) exceeds a critical value (U_{crit}), mutation outpaces selection, causing population extinction in a process known as "lethal mutagenesis" [1]. Lethal mutagenesis is important both theoretically [2–7] and clinically, where drugs (e.g. Ribavirin) can extinguish pathogens, especially RNA viruses, by elevating the mutation rate beyond U_{crit} [1,8–14].

Aside from mutation rate, population size (N) also plays an important role in extinction. All existing estimates of U_{crit} assume that $N = \infty$ [2–4], so that extinction can be modeled with relatively simple deterministic equations. In contrast, every real population has only finitely many members and is consequently subjected to "random genetic drift," i.e. stochastic fluctuations in birth-death events. More precisely, every real population of size N is guaranteed to experience fluctuations of order $\sim 1/N$ reminiscent of "shot noise," since births and deaths occur as discrete events.

Upon first thought, it may seem that genetic drift merely represents a small correction to the deterministic dynamics. However, the actual behavior is dramatically more interesting: Since drift continually obfuscates fitness differences among individuals, it weakens selection and implicitly tilts the mutation-selection balance in favor of mutation [15,16]. By this mechanism, known as "Muller's ratchet," [17] *unbiased* birth-death fluctuations end up downwardly *biasing* mean fitness within a population. Muller's ratchet has long been studied theoretically [18–24] and routinely exploited experimentally to prepare low fitness lines of organisms [25]. However, the extent to which high mutation rates exacerbate Muller's ratchet en route to extinction is neither qualitatively nor quantitatively well understood. We revisit this issue and review the literature on Muller's ratchet in Discussion.

In principle, mutations can cause extinction by two distinct, though non mutually exclusive mechanisms. First, deleterious mutations might decrease the absolute birth rate of a population to such a great extent that individuals are killed by natural forces (e.g. old age, environmental stresses, etc) faster than they reproduce. Most previous studies, e.g. refs. [2,19,20,22,23], have analyzed this first scenario, which represents a struggle between a population and its environment. A second, qualitatively distinct scenario is

Author Summary

Most spontaneous mutations hurt organismal fitness, e.g. by destabilizing proteins. In many species, the normal mutation rate is strikingly high: on the order of one per genome per replication. In the face of these mutations, how can proteins maintain their native structure, and how can populations of organisms avoid extinction? Are there physics-based limits on how large the mutation rate of any species can be before the onslaught of mutations outpaces natural selection and melts-down proteins? Here, we address these questions with a computational model that combines protein folding thermodynamics with individual-based population genetics simulations. We calculate a theoretical "speed limit" equal to a few mutations per genome per replication—near the mutation rate of RNA viruses. Additionally, we find that the speed limit can be much lower in small populations where "random genetic drift" is strong. Thus, we conclude that stably folded proteins are only possible in ecological settings that support sufficiently large populations. These findings may have clinical implications for treating viral infections with drugs that elevate the viral mutation rate.

possible for organisms that reproduce "semi-conservatively," including all viruses and unicellular species (see Results for elaboration): Every birth event risks ruining the "original" genome with new lethal mutations, thereby reducing the census size and risking extinction [4,26]. These dynamics represent the struggle of a population against itself. In this paper, we focus primarily on this second mode of extinction.

A sticking point for all lethal mutagenesis models is the relationship between genotype and fitness, i.e. the fitness landscape (FL). The distribution of fitness effects (DFE) among new single mutations furnishes the first order description of the FL. The second order description specifies the form of epistasis, i.e. how pairs of mutations interact to impact fitness. With few exceptions [3,5,7], previous studies generally assume that the DFE conforms to a simple mathematical function and make drastic simplifying assumptions regarding epistasis. In particular, most previous studies assume either no epistasis, or that mutations interact either all synergistically or all antagonistically, leaving little room for phenomena such as compensatory mutations. The motivation behind those assumptions is in part due to the large number of (unknown) parameters necessary to even write down a reasonably complex, epistatic FL.

Here, we circumvent that impasse by utilizing a previously developed approach [27] that is virtually parameter-free. We do not explicitly *impose* a DFE or a model of epistasis. Instead, those features emerge as the *output* from a biophysics-based protein folding requirement: We assume that mutations inflict fitness penalties proportional to the extent that they unfold proteins. Remarkably, this minimal assumption roughly accounts for the DFE observed in site-specific mutagenesis experiments on several viral species [27].

Here, we combine our biophysics-based FL with individual-based population-genetics simulations where extinction can only result from lethal mutations. We measured the time until extinction (τ) as a function of population size (N) and mutation rate (U). In accord with previous studies, we observe an ultimate mutation rate ($U_{crit}^* \approx 2.5$ mutations per genome) beyond which even infinitely large populations go extinct "almost immediately" ($\tau \sim log(N)$ generations). However, when $U < U_{crit}^*$, we find that τ depends dramatically on N: Small populations go extinct in

$\tau \sim log(N)$ generations, whereas large populations survive "almost indefinitely" ($\tau \sim e^N/N$). The boundary between "large" and "small" populations depends on U and is reminiscent of a "phase transition" between survival and extinction. In addition to τ, the coefficient of variation (i.e. standard deviation divided by mean) also undergoes a transition from values near zero in the extinction phase to values near one in the survival phase. These results contradict the simplistic intuition that "small populations are more stochastic than large populations."

For comparison, we also analytically solve for τ in a very simple model in which all mutations are either lethal or neutral. Solutions to this model clarify the meaning of extinction in finite populations and motivate our definition of survival vs. extinction phases.

Results

Semi-conservative birth-death-mutation model

Birth. This paper concerns asexual populations of replicating entities, henceforth called "cells." "Births" occur when a mother cell gives rise to exactly two daughters and the mother simultaneously dies. In continuous time, individual cells are chosen to give birth with probability proportional to their fitness (W), i.e. their birth rate. See Methods for further details.

Death. Besides the death of mothers upon birth, death also occurs by three additional mechanisms. First, if a birth event "tries to" increase the number of cells ($n(t)$) beyond a maximum number (N), then a random cell is removed from the population (similar to Moran's model [28]). Note that, strictly speaking, N is not the population's size but rather its capacity; density-dependent mortality kicks in abruptly when $n = N$. Secondly, as described below, some cells inherit new mutations, and if any of these mutations are lethal, then that daughter cell is killed immediately. Thirdly, cells can die (with rate δ) of "natural" causes, e.g. old age, washout, clearance, etc, independent of replication events or lethal mutations. Extinction in real populations is likely caused by a combination of these second and third mechanisms of death. For the bulk of this paper, we focus our attention on the second source, which represents an ultimate limit to population survival. A crucial feature of this limiting regime is that extinction cannot result from low fitness per se, which merely increases the generation time. Rather, lethal mutations are the only mechanism that can cause extinction in this regime (fig. 1).

Mutation. Each of the two daughter cells independently can acquire (nonsynonymous) mutations during their birth, i.e. replication is semi-conservative; see below for elaboration on this crucial assumption. In particular, if both daughters inherit lethal mutations, then the census size decreases by one. Note that if a master copy of the replicating genome was preserved (i.e. replication was conservative) and the natural death rate is zero, the population could not decrease in size and extinction would be impossible. We assume that the number of new nonsynonymous mutations per cell per birth event is Poisson distributed with mean U, i.e. U is the total genomic nonsynonymous mutation rate. In general, mutations can either be lethal or else merely perturb fitness (W), i.e. alter the doubling time. Our hypothesis that the onset of extinction is marked by excessive lethal mutations has experimental support [29] in viruses (see Discussion).

Biological interpretations. Our use of the term "semi-conservative" is based on the mechanism of DNA replication during cellular binary fission: Each daughter cell inherits one of the parent's two DNA strands, which then acts as a template for (potentially erroneous) synthesis of the remaining strand. Semi-conservative replication applies to all unicellular species. Additionally, our semi-conservative model can be interpreted in terms
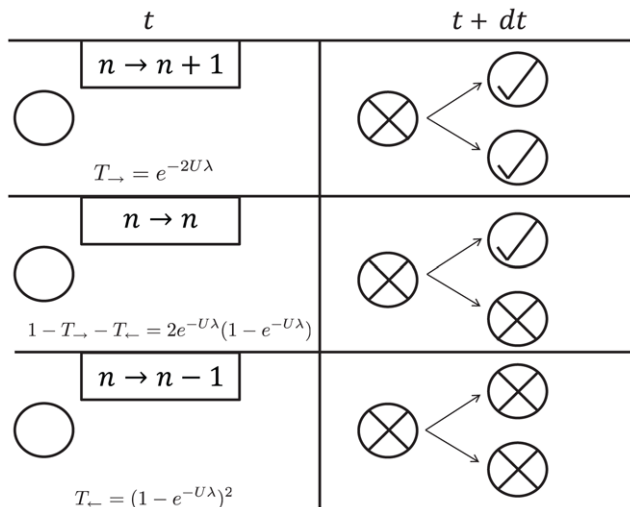
Figure 1. Mechanism by which the number of cells (*n(t)*) increases or decreases in the absence of natural death (*δ = 0*). The situation before and after a birth event is shown in the left and right columns, respectively. A parent cell immediately dies after giving rise to exactly two daughters. Replication is semi-conservative: Each daughter independently acquires a Poisson distributed number of lethal mutations, with mean *Uλ*. If both daughters are free of lethal mutations (check marks), then *n(t)* increases by one (top row). If both daughters acquire lethal mutations and die ("x" marks), then *n(t)* decreases by one (bottom row). The probability of each transition is also shown in the left column. In addition to these mechanisms, cells may experience natural death with a fixed rate (*δ*).
doi:10.1371/journal.pcbi.1002609.g001

of cells infected by viruses, even though real viruses often have single stranded genomes that, on a molecular level, are conserved during replication. In particular, viral reproduction is *effectively* semi-conservative if the following assumptions apply:

1. Infected cells lyse (i.e. the "mother" cell dies) when viruses emerge from them.
2. The number of virus particles that emerge upon lysis is very large.
3. Of the large number of released virus particles, only a small number (R_o) go on to infect other cells, independent of viral fitness. For simplicity, we assume that $R_o = 2$ throughout this paper.

Assumptions 2 and 3 together insure that the same exact genome molecule that infects a cell does not initiate subsequent infections (i.e. that the process is essentially semi-conservative). Note that, when considering viruses, we keep track of the number of infected cells as opposed to free viruses, which cannot autonomously replicate. We assume that only a single genome infects a particular cell, i.e. low multiplicity of infection; without this assumption, a virus's fitness would depend not only on its own genotype, but also on that of co-infecting viruses [30]. Also note that our model does not explicitly consider infected versus uninfected cells [31]; for a treatment of lethal mutagenesis with such a model, see ref. [7].

Flat, non-epistatic fitness landscape: survival phase versus extinction phase

The goal of this paper is to calculate which values of the population capacity (*N*) and mutation rate (*U*) support survival and which lead to extinction. In a sense, the answer is trivial: extinction

is certain if $N<\infty$ and $U>0$ since the population only has a finite number of configurations and all of them, including extinction, will be visited *eventually*. Nevertheless, the question remains as to which values of *N* and *U* enable populations to live a "long time" versus a "short time." However, it is not clear *a priori* even whether there exists a sharp, qualitative distinction between "long" and "short" or whether those concepts continuously blur together. Obviously, a crucial prerequisite for understanding extinction in finite populations is to define exactly what is meant by "long" and "short," i.e. "survival" vs. "extinction." To this end, we first consider a simple, analytically solvable fitness landscape whose solutions clarify these crucial preliminary issues. Later, we consider a more realistic FL based on protein biophysics.

By "fitness landscape" (FL) we mean a mathematical function relating genotype to fitness. We first consider a very simple FL in which the distribution of fitness effects (DFE) among new mutations is always the same, independent of genotype and/or fitness; this FL is non-epistatic, by definition. To further simplify our analysis, we assume that this preliminary landscape is "flat," insofar as all mutations are either lethal or completely neutral.

Given said assumptions, all relevant aspects of the population are completely described by the number of living cells (*n(t)*). *n(t)* thus undergoes a biased random walk, with a natural absorbing boundary at *n = 0* and a reflecting boundary at *n = N*. Since the number of lethal mutations per offspring is Poisson distributed with mean *Uλ*, the transition probabilities per unit time for increasing and decreasing *n* by a single individual are, respectively:

$$T_{\rightarrow}(n) = W^* n e^{-2U\lambda}, \ (n<N)$$
$$T_{\rightarrow}(n) = 0, \ (n=N) \tag{1a}$$

$$T_{\leftarrow}(n) = W^* n (1-e^{-U\lambda})^2 + \delta \tag{1b}$$

where W^* is the fitness of all viable cells and δ is the natural death rate. W^* plays no essential role, and is often set to one for convenience. Eqs.1 are plotted in fig.S1. The exponentials in eqs.1 express the probability that none or both of the offspring carry lethal mutations. Note the factors of *n* in eqs.1, which are not present in the simplest "text book" random walk. This model could be extended to handle bursts of *R* offspring by replacing factors of 2 with *R* and considering larger jumps.

We first investigate the average behavior eqs.1, and then perform a stochastic analysis. According to eqs.1 (see also Text S1), the expected change in the census *n* during the time interval *dt* ($\langle d n \rangle$) obeys the following equation:

$$\langle dn \rangle = \langle T_{\rightarrow}(n) - T_{\leftarrow}(n) \rangle \tag{2a}$$

$$= [W^*(2e^{-U\lambda} - 1) - \delta]\langle n(t) \rangle \tag{2b}$$

$$\equiv [W_{net} - \delta]\langle n(t) \rangle \quad (n<N). \tag{2c}$$

We call W_{net} the "net fitness," since it has a component related to fitness of living cells (W^*) discounted by a component that depends on the production rate (*Uλ*) of mutants. The most dramatic distinction between mean fitness and W_{net} is that, while the former must be ≥ 0 (birth rates obviously can't be negative), the latter becomes negative when $U\lambda > ln(2)$ because of lethal

mutations and semi-conservative replication. Note that for realistic values of U and λ, W_{net} is substantially less than W^*: e.g. if $U=1$ and $\lambda=0.3$ (see fig. 2c), $W_{net}=0.48W^*$ (i.e. mutation reduces fitness by 52%), which underscores the impact of lethal mutations and semi-conservative replication in limiting the growth of $\langle n \rangle$.

Eq.2 implies that the expected value $\langle n \rangle$ for the census size either grows (until $n=\mathcal{N}$) or decays exponentially with rate W_{net}-δ. We denote these two opposing regimes as the "survival phase" and "extinction phase." The boundary between survival and extinction occurs when the natural death rate balances net fitness:

$$(U\lambda)_{crit} = \ln(2) + \ln\left(\frac{1}{1+\delta/W^*}\right) \tag{3a}$$

$$= \ln(2) - \frac{\delta}{W^*} + O\left(\frac{\delta}{W^*}\right)^2 \tag{3b}$$

Note that, if $U\lambda > ln(2) \approx 0.7$, populations will be in the extinction phase even if no natural death occurs ($\delta=0$). To a close approximation, the effect of natural death is merely to decrease $(U\lambda)_{crit}$ by an amount δ/W^*, i.e. the number of natural deaths per generation. Also not that the value of $\mathcal{N}$ is irrelevant to whether populations are in the survival or extinction phase on this non-epistatic fitness landscape (FL); the picture will be radically different later, when we consider a more realistic FL.

We now turn to the stochastic features of this model. Given the initial condition that there are $\mathcal{N}$ cells at $t=0$, all populations go extinct with probability one, but we can calculate the statistics of how long the population survives before going extinct (i.e. hitting the absorbing state at $n=0$). In Text S1, we derive a general analytic formula for the mean time until extinction (τ), i.e. the mean "first passage time," by approximating n as a continuous variable and solving differential equations. The continuity assumption is valid for $|v/D| \ll 1$ (see Text S1). The asymptotic behavior of the general solution $\tau(U,\mathcal{N})$, valid for large but finite $\mathcal{N}$ and $U\lambda \neq ln(2)$, is given by

$$\tau(U,N) \sim -\frac{1}{v}\ln\left|\frac{Nv}{D}\right|, \; \frac{Nv}{D} < < -1 \tag{4a}$$

$$\tau(U,N) \sim \frac{D}{Nv^2}e^{Nv/D}, \; \frac{Nv}{D} > > 1, \tag{4b}$$

where $D(n,U) \equiv \frac{1}{2n}(T_{\rightarrow}+T_{\leftarrow})$ and $v(n,U) \equiv \frac{1}{n}(T_{\rightarrow}-T_{\leftarrow})$, plotted in fig.S1, can be interpreted as the diffusion coefficient and convection velocity from diffusion theory, respectively. Eqs.4, along with the exact analytic expressions, are plotted in fig.S2. Note that, since v equals the term in brackets in eq.2c, eqs.4a,b corresponds to the extinction and survival regimes, respectively.

The rough functional dependence of τ on $\mathcal{N}$ in eqs.4 might be anticipated intuitively. Eq.4a applies when populations are biased toward extinction. In that case one expects that $n(t)$ decays exponentially from $n=\mathcal{N}$ down to $n=0$: $n \sim Ne^{-vt}$, which implies that $\tau \sim \frac{1}{v}\ln(N)$, similar to eq.4A. On the other hand, eq.4b applies when populations are biased toward survival. In that case, one expects that extinction requires an extraordinary run of $\sim \mathcal{N}$ lethal mutations, which should occur with probability on order of $(U\lambda)^{-\mathcal{N}}$. The time until extinction in the survival regime thus might be expected to scale as $\tau \sim (U\lambda)^N = e^{N\ln(U\lambda)}$, similar to eq.4b which is also dominated by $\mathcal{N}$ in the exponent. In Text S1, we also analyze how τ scales with U. We find that in the survival phase, but not too far from the transition at $U\lambda=ln(2)$, $\tau \sim \frac{1}{4N(U\lambda+\delta-\ln(2))^2}e^{-4N(U\lambda+\delta-\ln(2))}$. In the extinction phase we find $\tau \sim \frac{1}{U\lambda+\delta-\ln(2)}\ln[4N(U\lambda+\delta-\ln(2))]$. Thus, τ depends sharply on both $\mathcal{N}$ and $U\lambda$ in the survival phase, yet depends only weakly on these variables in the extinction phase. These approximations are plotted in fig.S2. In Text S1 and fig.S3, we also consider the variance in extinction time.
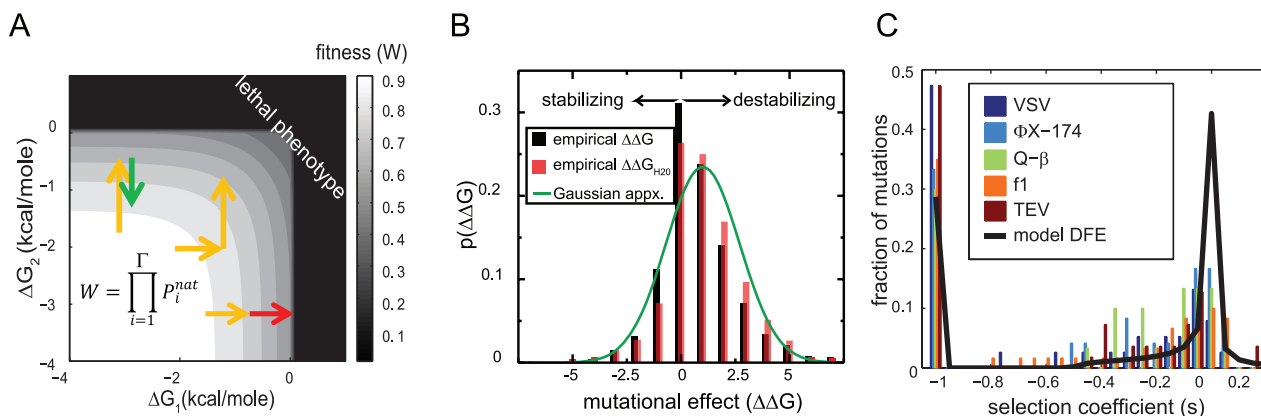

Figure 2. Biophysics-based fitness landscape (FL). A: A two dimensional slice of the Γ dimensional FL. Fitness values are shown in grayscale and pairs of mutations are represented by colored arrows. Deleterious mutations interact synergistically (yellow/red pair). Compensatory and non-epistatic mutations are also possible (yellow/green and yellow/yellow pairs, respectively). Mutations that push $\Delta G>0$ cause lethality. **B:** Distribution of mutational thermodynamic effects $p(\Delta\Delta G)$. Our approximation for $p(\Delta\Delta G)$ agrees with experimental values obtained via thermal (black bars) and/or solute (red bars) denaturation. The ~4,000 experimental values were taken from the ProTherm database [56]. **C:** The distribution of fitness effects (DFE) among new random mutations from our model (black curve) and several viral species (colored bars). The horizontal axis is the selection coefficient, which depends on fitness before and after the mutation: $s \equiv W_{after}/W_{before}-1$. The DFE from this model depends on N and U [27]; here $N=10240$, $U=2$ (chosen so that the population was near the extinction threshold). See Methods and ref. [27] for procedures used to obtain the DFE.
doi:10.1371/journal.pcbi.1002609.g002

Extinction on epistatic biophysical fitness landscape

We next consider a recently developed biophysics-based fitness landscape (FL), which features a continuous distribution of fitness effects (DFE) among random new mutations (fig. 2). The basic assumption of this approach is that mutations inflict a fitness penalty proportional to the extent that they unfold proteins by perturbing thermodynamic stability (ΔG). Below, and in fig. 2, we describe some important features of this model; see Methods and/or ref. [27] for details.

1. This FL is epistatic: a given mutation unfolds barely-stable proteins more so than very stable proteins.
2. Fitness increases (though usually very weakly) with increasing stability (i.e. decreasing ΔG).
3. Approximately 30% of all mutations are compensatory, although most increase fitness by only a negligibly small amount ($\ll 1/\mathcal{N}$).
4. Each cell has a fixed number (Γ) of proteins, or, more precisely, Γ protein folding domains. For convenience, we assume $\Gamma = 20$ in simulations.

Mutations that completely unfold proteins or hit a small fraction of functionally critical residues (e.g. the active site) are considered lethal. Points 1–3 above are not explicit assumptions of our model; rather, they follow naturally and implicitly from a biophysics-based framework (see Methods and ref. [27]). Crucially, the DFE from this FL roughly agrees with experimental data, at least for viruses [27] (fig. 2C). Unlike the flat landscape that we considered previously, the biophysical FL features many mutations that only slightly decrease fitness. These mutations profoundly increase the importance of $\mathcal{N}$ (even when $\delta = 0$), because they can only be purged by sufficiently large populations ($|\mathcal{N}s| > 1$).

As with most multi-locus models, dynamics on our biophysics-based FL is too complex to solve analytically for finite $\mathcal{N}$. Thus, we resort to stochastic computer simulations, as described in Methods. Fig. 3a shows how the mean extinction time (τ) depends on population capacity ($\mathcal{N}$) for various mutation rates (U). On this log-log plot, upward bending curves increase faster than a power law

(exponential-like scaling, c.f. eq.4b), whereas downward bending curves increase slower than a power law (logarithmic-like scaling, c.g. eq.4a). As we observed in the non-epistatic analytic model, U strongly impacts τ. Above an ultimate extinction rate ($U_{crit}^* \approx 2.5$, roughly estimated by eyeing simulation results), τ scales approximately logarithmically with $\mathcal{N}$ (extinction phase), even when $\delta = 0$ and $\mathcal{N} \to \infty$. For very small U, τ scales approximately exponentially with $\mathcal{N}$ (survival phase) for all $\mathcal{N}$. Fig.S4 explicitly shows τ versus U. For reference, we note that real RNA viruses have mutation rates in the approximate range $0.1 < U < 5$, whereas DNA based microbes generally have $U \approx 0.003$ [32,33]. According to fig. 3a, our model predicts that mutation rates characteristic of (non-mutator) DNA based microbes will always reside squarely in the survival regime for virtually any $\mathcal{N}$, whereas RNA viruses lie near the extinction regime, and may be pushed into it, by modestly adjusting $\mathcal{N}$ and/or U. A few viral species have slightly higher mutation rates than the threshold U_{crit}^* (≈ 2.5) from our model (e.g. bacteriophage Qβ has $U \approx 5$ [33]); this is likely due to large burst size: see Discussion.

In stark contrast to the preliminary, non-epistatic FL, fig. 3a shows that $\mathcal{N}$, not just U, also determines whether a population is in the survival or extinction phase. This is most apparent in the cyan curve representing $U = 2$, whose curvature suddenly changes near at a critical value ($\mathcal{N}_{crit}$). For large $\mathcal{N}$, populations are in the survival phase, whereas below $\mathcal{N}_{crit}$ populations enter the extinction phase. $\mathcal{N}_{crit}$ becomes arbitrarily large as $U \to U_{crit}$. To get a quantitative sense of these values, given reasonable parameter values, consider $U = 2$: $\mathcal{N}_{crit}$ equals 100 or so when $\delta = 0$ and rises to $\sim 10^4$ for the modest value of $\delta = 0.1$. τ is only ~ 100 generations in the extinction phase, and rises quickly from this level in the survival phase.

The extinction vs. survival phases are even more clearly delineated by the coefficient of variation (CV_τ) of the extinction time (i.e. its standard deviation divided by its mean (τ)). CV_τ measures stochasticity in populations' longevity. Fig. 3b shows that deep in the extinction phase, $CV_\tau \to 0$, whereas $CV_\tau \to 1$ in the survival regime. These limits make intuitive sense: $CV_\tau = 0$ represents deterministic extinction, whereas $CV_\tau = 1$ is a hallmark
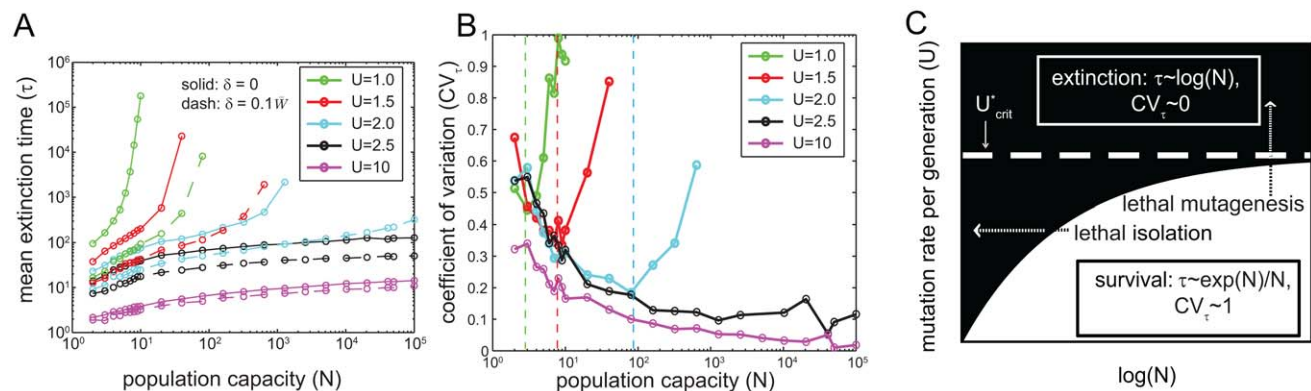

Figure 3. Extinction on biophysical fitness landscape. All finite populations eventually go extinct. The mean number of generations until extinction (τ) increases with population capacity (N) and decreases with mutation rate (U). **A:** τ versus N at various fixed m; notice the double logscale. Curves for $U < 2.5$ or so have an inflection point, signaling a qualitative transition from extinction to survival as N increases. Solid curves correspond to $\delta = 0$, while dashed curves correspond to $\delta/W^* = 0.1$. **B:** Coefficient of variation in time until extinction (CV_τ) for the same parameters as panel A. CV_τ increases towards one in the survival phase and decreases toward zero in the extinction phase, as N increases. Curves "peel off" toward $CV_\tau = 1$ at the critical population capacity (N_{crit}), shown approximately with dashed lines. When $U > U_{crit}^*$, both simulation results and general arguments (see main text) show that curves do not peel off, i.e. N_{crit} does not exist. **C:** Heuristic cartoon "phase diagram" summarizing the behavior from panels A,B. In panels A,B τ values are reported only in cases where extinction occurred within 10^5 generations in each replicate. See main text and fig.S5 for a quantitative sense of how N_{crit} depends on U_{crit}. See Methods for averaging procedures. $\Gamma = 20$ throughout this paper.
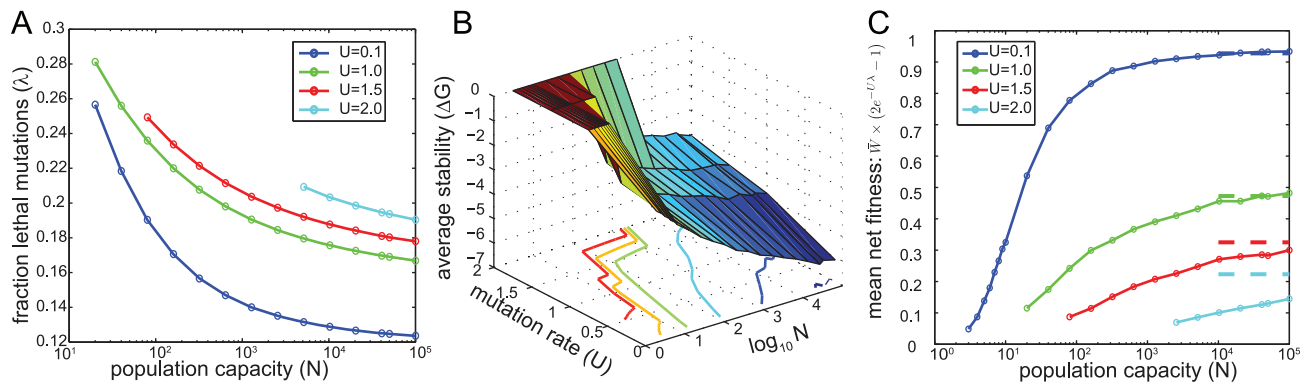doi:10.1371/journal.pcbi.1002609.g003

Figure 4. Small population capacity (N) and high mutation rate (U) cause depressed fitness, unstable proteins, and many lethal mutations. A: Fraction of mutations (λ) which were lethal during simulations on our biophysical fitness landscape. Fig.S6 shows a related plot of $U\lambda/ln(2)$ versus N. **B**: Protein stability (ΔG) averaged over both proteomes and populations. The accumulation of unstable proteins when N is small and/or U is large is the underlying cause of changes in λ observed in panel A. **C**: Mean net fitness, which takes into account both average birth rate and lethal mutations (eq.2c). Populations are not perfectly fit because of genetic drift (caused by small N) and mutation load (caused by large U). The classical expectation, which assumes $N = \infty$ and no beneficial/compensatory mutations, predicts that overall growth rate is given by e^{-U} (shown with dashed lines on the right of panel B). The classical expectation fares poorly at small N and large U. Data is shown only for (N,U) values for which at least one of the replicate populations survived until the end of the simulation (see Methods); otherwise, λ, W, and ΔG are not clearly defined, since a quasi-steady state does not exist.
doi:10.1371/journal.pcbi.1002609.g004

of an exponential distribution describing the waiting time for an extraordinary run of independent lethal mutations in all N cells. Thus, CV_τ behaves as an "order parameter" familiar from phase-transition theory: survival plays the role of the "ordered phase" ($CV_\tau = 1$) while extinction represents the disordered phase ($CV_\tau = 0$). In contrast, the preliminary, non-epistatic FL does not transition from the extinction to survival phase as N increases, since CV_τ in a monotonic function of N in that model (fig.S3).

A crucial lesson from fig. 3b is that N_{crit} depends on U_{crit}, and vice versa. Curves representing higher mutation rates "peel off" to $CV_\tau = 1$ at larger values of N than do curves representing lower U. However, for $U > U_{crit}^*$, the curves cannot transition to $CV_\tau = 1$ for any value of N. This assertion is clear from the fact that a finite percentage (10% here, see Methods) of mutations are uncondi-tionally lethal, and at sufficiently large U, nearly all progeny will acquire these (if no other) lethal mutations, resulting in extinction. Following the logic of eq.2, an upper bound for U_{crit}^* given 10% unconditional lethals, is $U_{crit}^* < \dfrac{\ln(2)}{0.1} \approx 7$ nonsynonymous muta-tions per genome replication. Overall, the transition from extinction to survival is summarized by the "phase diagram" in fig. 3c. The non-rectangular phase boundary in fig. 3c emphasizes the interrelatedness of N_{crit} and U_{crit}. In particular, the boundary can be crossed by either increasing U (i.e. lethal mutagenesis) or decreasing N, which we refer to as "lethal isolation." Fig.S5 shows a quantitative version of fig. 3c.

Why does the biophysics-based FL enable "lethal isolation" while the non-epistatic FL does not? The answer is that mean net fitness ($\bar{W}_{net}$) increases with N (fig. 4c) on the biophysics-based landscape; this increase derives from two distinct sources. First, larger populations more effectively purge weakly deleterious mutations having $Ns < 1$, thereby driving up the mean birth rate ($\bar{W}$) among living members of the population. Consequently, large populations can grow fast enough to outpace natural death (δ). Secondly, and more profoundly, small populations produce a larger fraction (λ) of lethal mutations (fig. 4a) on the epistatic, biophysics-based FL; these additional lethal mutations can decrease W_{net} below zero and cause extinction even when $\delta = 0$ (figs. 3,4). The biophysical basis for this effect is that, when

proteins are only barely stable (as predicted to be the case in small populations, see fig. 4b), more mutations are within striking range of the unfolding transition at $\Delta G = 0$ (fig. 2a) and a corresponding lethal phenotype.

Discussion

A prerequisite for understanding extinction in finite populations is a coherent method for classifying extinction versus survival: Although all finite populations eventually go extinct, our analytic and simulation models show that "not all extinction is the same." That observation led us to define two opposing dynamical phases for evolving populations: The extinction phase is characterized by rapid, nearly deterministic ($CV_\tau \approx 0$) decline whereas the survival phase is characterized by long yet uncertain extinction times ($CV_\tau \approx 1$). Intuitively, one usually thinks of small populations as being more stochastic than large populations. However, results from our biophysical fitness landscape (FL) show that that intuition needs refinement: the increased *stochasticity caused by small N actually makes extinction more deterministic* (e.g. CV_τ can decrease with N in fig. 3b). The situation is analogous to a building experiencing an earthquake: if the strength of noisy seismic vibrations (i.e. genetic drift) crosses a threshold, gravity (i.e. deleterious mutations) deterministically destroys the building (i.e. population goes extinct).

Previous studies of mutation-induced extinction focused exclu-sively on either the role of high mutation rate (U) or small population capacity (N). Those that focused on high U neglected the role of genetic drift by assuming that $N = \infty$ [2–4]. Likewise, those studies that included genetic drift generally neglected the role of high U [7,19–21,23,34]. This paper bridges those previous two approaches by exploring how interplay between small N and large U accelerates extinction. We found that rapid extinction occurs on the biophysical FL whenever N is too small ($N < N_{crit}$) or U is too large ($U > U_{crit}$). Furthermore, we found that N_{crit} depends on U_{crit} and vice versa, i.e. the phase boundary in fig. 3c is not rectangular. In particular, small population capacity reduces U_{crit}. This reduction is generally modest: e.g. we see U_{crit} shift from about two to about one in figs. 3 and S5, as N varies from 10 to 10^5. Although U_{crit} depends only weakly on N, fig. 3 shows that τ,

as well as the qualitative behavior of the population, can depend dramatically on N.

Semi-conservative reproduction is a key feature of our model that distinguishes it from most previous studies, e.g. refs. [2,19,20,22,23]. This distinction is sharpest in the regime that $\delta = 0$, where low fitness (i.e. long generation time) contributes to extinction only insofar as it increases the fraction of lethal mutations (λ). In other words, the population is its own (and only) enemy in the $\delta = 0$ regime. In reality, populations must survive not only in spite of themselves, but also in spite of death imposed by the environment. Fig. 3a shows that, in the plausible scenario where 10% of the population dies from natural causes each generation, τ is further shortened significantly. We note that obtaining parameter values for δ is not always straightforward because it requires discriminating between natural death and density-dependent death (i.e. death due to fixed N). For example, cells infected with HIV turn over approximately once per day [35], but it is unclear what fraction of the turnover is due to density dependence versus other "natural" causes.

For the sake of simplicity, we assumed that only two offspring result from a birth event. In the case of viruses, two should be replaced by an "effective burst size" that takes into account the number of virions released during the infected cell's lifetime as well as the fraction of those virions that go on to infect future cells. A simple deterministic analysis [2] suggests that U_{crit} increases as the logarithm of effective burst size, but otherwise does not change the qualitative picture. In the context of finite populations, large burst size may also non-trivially reduce the effective population size [36], since all members of a burst are closely related; we leave investigation of this topic to future work.

Apart from large burst size, Martin and Gandon [7] recently pointed out another mechanism that may partially buffer viral populations against extinction. Using an explicit viral dynamics model that includes both susceptible and infected cells, those authors point out that as viral load declines under elevated mutation rates, the number of susceptible cells is predicted to correspondingly increase. This effect may tend to offset and/or halt the decline in growth rate caused by elevated mutation rates. While this mechanism may be important, our model predicts that its extinction-buffering potential may be limited. In particular, we predict that the fraction of lethal mutations (λ) increases as fitness decreases (fig. 4a); in our model, any increased growth rate, from whatever origin, will be countered by a correspondingly elevated death rate from lethal mutations.

Previous calculations of U_{crit} (deterministic studies)

Deterministic models remove N from consideration by assuming that $N = \infty$, which enables a comparatively straightforward calculation of the ultimate mutation rate (U_{crit}^*) beyond which even infinitely large populations go extinct. Our biophysical model also features an ultimate mutation rate (U_{crit}^*) (see horizontal asymptote in fig. 3c), and additionally it predicts that when N is finite, $U_{crit} < U_{crit}^*$. While it is unsurprising that τ decreases as N does, it is rather surprising that decreasing N can fundamentally change the dynamical regime of the population from survival to extinction.

Zeldovich et al. [3] utilized a biophysical fitness landscape similar to the one presented here. Apart from their assumption that $N = \infty$, the main difference with our approach is that their fitness landscape had a strictly flat "mesa," i.e. they approximated eq.6 (Methods) as a true step function. By contrast, our model features nearly-neutral mutations (fig.S4) which enhance the role of population capacity (N), since mutations with $Ns < 1$ are invisible to natural selection [27,37].

The deterministic theory of Bull and Wilke, first laid out in ref. [2] and subsequently elaborated upon in ref. [4], is another important benchmark for comparison. Using a simple, classical equation, those authors calculated equilibrium mean fitness and compared this to the rate of natural death. Neglecting beneficial/compensatory mutations, they calculated a maximum allowable "deleterious mutation rate" of $ln(2) \approx 0.7$, which is the same value we calculated for the *lethal* mutation rate ($U\lambda$) in the preliminary, non-epistatic FL. By contrast, on our biophysical FL, we predict an *overall* maximum nonsynonymous mutation rate of $U_{crit}^* \approx 2.5$ (fig. 3). The discrepancy between 0.7 and 2.5 derives from many factors, including compensatory mutations in our model and ambiguity in what those authors mean by "deleterious," i.e. which mutations they would define as deleterious as opposed to neutral.

Previous calculations of N_{crit} (stochastic studies)

Random drift is the paramount concern of a separate line of previous studies that describe extinction in terms of Muller's ratchet [19–22,38]. However, those studies minimize the importance of mutation rate. For example, neglecting beneficial mutations and using an approach based on fixation probabilities, Lande [21] calculated that $\tau \sim 1/U$; i.e. his result is that U merely sets the time units but is irrelevant to the essential behavior. As another example, Whitlock [23] included beneficial mutations and calculated that $N_{crit} \sim (U_{deleterious}/U_{beneficial})^{1/3}$, which depends only on the balance of beneficial to deleterious mutations and *not* on the mutation rate itself. Both of those examples contradict our results, which show that N_{crit} and τ depend dramatically on $|U|$. The dominant reason for the discrepancy is that those authors assumed that deleterious mutations occur "one at a time," which is not true when the rate that mutations are introduced (U) exceeds the rate at which selection removes them ($\sim 1/s$). When $U/s \gg 1$, the population experiences "Hill-Robertson interference" [39], which both accelerates extinction and also makes analytic solutions intractable.

A separate, very serious concern about many previous studies (e.g. refs. [19,21]) is that, for all parameter values they explored, they always observed a small coefficient of variation in extinction time (CV_τ). Based on our results (fig. 3), this suggests that *those authors only probed the extinction regime*. In other words, their models were constructed such that extinction occurred nearly deterministically. By contrast, it seems likely that most, if not all, natural populations are in the survival regime as long as their population size and/or mutation rate are not interfered with externally (e.g. via mutagens or habitat destruction). A related issue concerns the initial conditions of those models. They assumed that populations were extremely fit initially, such that each individual leaves a large number (R_o) of descendants ($R_o \gg 1$). Extinction occurs in those models when R_o semi-deterministically drops to just below one, after several deleterious mutations achieve fixation. By contrast, our simulations begin in a natural condition (see Methods): mutation-selection-drift equilibrium, which may not even exist in those previous models. The existence of a quasi-equilibrium state, i.e. the survival phase, is a major advantage of our approach. Indeed, the survival phase can be viewed as a stochastic analog of deterministic mutation-selection equilibrium.

Other sources of random genetic drift

In this paper we have focused on the subtly deleterious impact of unbiased fluctuations on allele frequency (i.e. genetic drift) caused by finite population capacity (N). In addition to finite N, several other factors can have a similar effect, including population bottlenecks, micro-environmental fluctuations, and stochasticity in gene expression [40]; these effects are sometimes summarized

collectively by an "effective population size" (N_e) [36]. Of particular relevance to extinction of RNA viruses are population bottlenecks that occur during transmission events. Indeed, it is estimated that most HIV-1 infections originate from a single infectious particle, which would greatly reduce N_e below the viral load.

Importance of fitness landscape

An important result from previous studies is that τ depends strongly on both the severity of deleterious mutations (i.e. the DFE [19,21,34]) and on epistasis [22,38,41]. The DFE has traditionally been represented by either a single selection coefficient (i.e. a Dirac delta function) or by a continuous function (e.g. Gamma distributions), which was assumed not to change with fitness and/or time. Thus, even relatively simple approaches were forced to make somewhat ad-hoc modeling choices and also introduce several parameters. The fact that τ depends strongly on the DFE thus presents a dilemma to researchers: they must either comb through a high-dimensional parameter space or else their results depend on myriad questionable assumptions. Our model circumvents this problem because the parameters (e.g. eqs. 6,7 in Methods) are not "adjustable"; rather, they are set by strictly biophysical considerations. In effect, we exchanged a poorly understood, high-level question (how mutations affect fitness) for a well-understood, microscopic question (how mutations affect protein folding thermodynamics). The validity of this exchange is commensurate with the extent to which our DFE matches experimental data (ref. [27] and fig.S4).

Apart from the advantages of our FL, we expect that the qualitative behavior in fig. 3 might also be observed in some traditional models. We anticipate that the essential requirements are (i) both beneficial and deleterious mutations (so that the population does not inevitably "slide downhill") and some upper bound on fitness (so that the population cannot forever "climb uphill").

Extinction versus "error catastrophe"

Decades ago, Eigen calculated that genomes can become "delocalized" in sequence space during an "error catastrophe" when the mutation rate exceeds a critical value [42]. As pointed out by previous authors [2], extinction and error catastrophe are distinct concepts: The former is a demographic process whereas the latter refers to loss of the single fittest genotype on a toy fitness landscape, usually in the $N = \infty$ limit. Nevertheless, some results from one study of error catastrophe in the context of finite N [43] hint at our results in fig. 3; e.g. they observed that the time taken for populations to experience delocalization decreases with N.

Connections with experiments and viral infection treatment

In reality, is extinction accompanied by excessive lethal mutations (as in our model) or merely by a slow generation time that is unable to keep pace with natural death (as in previous models, e.g. refs. [2,19,20,22,23]? These two scenarios are distinguishable in laboratory evolution experiments on viruses because viral load (nucleic acid molecules per mL) and infectivity (plaques formed per mL viral suspension) can be measured separately. Several experiments on at least three viral species [29,44–47] show that when U is elevated near/past U_{crit}, viral load transiently continues to increase, simultaneous with a decline in infectivity. Thus, noninfectious genomes (i.e. those carrying lethal mutations) signal extinction during experiments, in accord with our model's interpretation of extinction.

A clinically relevant prediction of our model is that changing N can radically alter population survival, especially when U is elevated by drugs. This phenomenon was observed experimentally [8] with foot-and-mouth disease virus, where merely 10-fold dilutions during viral passages dramatically accelerated extinction in the presence of mutagenic drugs. As remarked by those authors, this finding suggests that therapies combining both mutagenic drugs and traditional drugs (which reduce the number of viable viruses) could substantially increase efficacy. Indeed, our analytic results (eqs.4 and eq.S11a) imply that even in the survival phase, the expected time to extinction depends exponentially on (i.e. is very sensitive to) both U and N, suggesting that altering either of these parameters could dramatically impact the chances of population extinction during a fixed time interval.

Apart from extinction per se, our general biophysics-based approach also has substantial experimental support. Our basic assumption is that protein unfolding/misfolding accounts for the deleterious effects of most mutations. If this were true, species with high U and/or low N should have less stable proteins. Several experimental facts suggest that this is in fact the case. First, chaperone overexpression compensated for the fitness decline caused by single-cell bottlenecks (low N) in bacterial populations [48,49]. Thus, these populations likely contained unstable, unfolded proteins which caused the fitness decline. Secondly, Fernandez and Lynch [50] recently reported more structural defects and thermodynamic instability among monomeric protein subunits in small populations than in large populations. Along similar lines, another study calculated less stability among proteins in endosymbiotic bacteria (small N) than in orthologs from free living relatives (large N) [51]. Thirdly, proteins from RNA viruses (high U) have a lower density of van der Waals contacts than orthologs in DNA viruses (lower U), suggesting, though not proving that RNA viral proteins are less stable [52]. Indeed, we have gone even further and predicted the distribution of stabilities within proteomes from species with various U and N (see fig. 5 from ref. [27]).

Methods

Biophysics-based fitness landscape

The approach here closely follows ref. [27]. Every cell contains a number (Γ) of *well-adapted* proteins, each of which exists in thermal equilibrium between its native, functional conformation and an ensemble of unfolded, nonfunctional conformations. The fraction of time in equilibrium that protein i spends in its native conformation is P_i^{nat}. Qualitatively, our assumption is that fitness is impaired when either the concentration of folded proteins decreases or, equivalently, the concentration of unfolded proteins increases. Quantitatively, we assume that

$$W = \prod_{i=1}^{\Gamma} P_i^{nat} \tag{5a}$$

$$\approx 1 - \sum_{i=1}^{\Gamma} \left(1 - P_i^{nat}\right)_i \quad \text{(if all } W = P_i^{nat} > 0.5) \tag{5b}$$

$$W = 0 \text{ (lethal if any } P_i^{nat} < 0.5), \tag{5c}$$

where P_i^{nat} is the fraction of time in equilibrium that protein i spends in its native conformation and W is fitness (i.e. birth rate). The approximation between eq.5a and eq.5b is valid for $P^{nat} \approx 1$, as

is the case for real proteins (see below). Eq.5a emphasizes the positive interpretation that each of the proteins is required in order for the organism to live and function. Likewise, eq.5b emphasizes the negative interpretation that misfolded/unfolded protein hurts the organism in proportion to their concentration in the cell.

In our model, P^{nat} is the master variable that connects proteins with fitness. However, it is simpler to work with a closely related quantity: the free energy difference (ΔG, also called "protein stability") between the folded conformation and the ensemble of nonfunctional conformations. We assume that proteins fold "two-state" [53], which implies the relationship

$$P^{nat} = \frac{1}{1+e^{\Delta G/k_b T}}, \qquad (6)$$

where k_b is Boltzman's constant and T is temperature.

Working with ΔG simplifies mutational effects because, (i) effects of mutations on free energy ($\Delta\Delta G$) are well characterized experimentally and (ii) $\Delta\Delta G$ (but not ΔP^{nat}) is additive when several mutations accumulate sequentially [54,55]. We approximate $p(\Delta\Delta G)$ by a Gaussian function with a mean of +1 kcal/mole and standard deviation 1.7 kcal/mole [3,56], which reasonably matches empirical data from the ProTherm database [56] (fig.S4):

$$p(\Delta\Delta G) = \frac{1}{\sqrt{2\pi(1.7)^2}} \exp\left[\frac{-(\Delta\Delta G-1)^2)}{2(1.7)^2}\right]. \qquad (7)$$

We also assume that $p(\Delta\Delta G)$ is independent of ΔG, which is consistent with empirical data [5,56], though only to a rough approximation.

While most mutations in our model only alter ΔG, a small fraction destroy protein function for non-thermodynamic reasons. For example, a few amino acid positions comprise the active catalytic site, and virtually all mutations there will abolish functional activity. Assuming that there are 3 catalytic residues, another 3 nearby critical sites and 100 total residues in the folding domain, these represent 6% of all random mutations. Besides the active site, some point mutations abolish activity by introducing premature STOP codons. Assuming random codon usage, premature STOP codons represent $\approx$4% of random mutations [27]. Thus, together, these categories comprise $\approx$10% of all nonsynonymous mutations, which we assume to unconditionally confer a lethal phenotype.

Eqs.6,7 al ong with said assumptions regarding lethal mutations, indirectly imply the distribution of fitness effects (DFE) in the biophysical model; detailed explanation of how this works is the subject of ref. [27]. Briefly, we first equilibrated populations for at least 10^5 generations, at which point populations had substantial diversity in fitness. Next, we measured the DFE among all single point mutations for each clone in the population. Finally, these DFE were averaged to obtain the overall DFE, e.g. in fig. 2C. This procedure essentially averages the DFE of each clone, weighting each in proportion to its probability of being randomly chosen as the starting point for mutagenesis experiments. Since our DFE describes nonsynonymous mutations only, synonymous mutations were removed from the experimental datasets in fig. 2c.

Simulation procedures

We iterated the birth-death-mutation process for 10^5 generations or until population extinction, whichever occurred first. Each birth event represents $1/n(t)$ generations. All populations were initialized with genomes (i.e. sets of ΔG values) sampled from a single, "burn-in" population that had previously achieved mutation-selection-drift equilibrium during 10^5 generations of evolution. The parameter values ($N=10^5$, $U=1$) of the burn-in population were chosen so as to lie clearly in the survival regime yet close to the regions of parameter space being probed throughout the paper. This choice minimizes the impact of (inherently somewhat artificial) initial conditions. The fraction of lethal mutations (λ) was estimated during each simulation run as the total number of lethal mutations divided by $2U\bar{n}$, where $\bar{n}$ is the time-averaged number of cells during the run.

Supporting Information

Figure S1 Transition probabilities, "convection velocity" and "diffusion coefficient" as functions of the lethal mutation rate ($U\lambda$), assuming that $\delta=0$. These are the per capita quantities (i.e. $n=1$). (EPS)

Figure S2 Mean time until extinction on flat, non-epistatic fitness landscape. Solid curves illustrate the exact solution, eq.S11. The dashed curves in panel A illustrate the approximations eqs.S12A,S12B. Dashed curves in panel B illustrate eq.S13A. The vertical dotted line marks the transition at $U\lambda = ln(2)$. $\delta=0$. Note that eq.S13A breaks down for very small $U\lambda$. As discussed in the text, the entire continuum approach breaks down in that regime. (EPS)

Figure S3 Coefficient of variation (standard deviation divided by mean) in time until extinction ($CV\tau$) when $\delta=0$. A: Supercritical ($U\lambda > ln(2)$) populations become more deterministic as N increases, while subcritical populations become more stochastic. B: $CV\tau$ decreases sharply at $U\lambda = ln(2)$, though the sharpness of the transition increases with N. (EPS)

Figure S4 The mean time until extinction declines rapidly as U increases when $U < Ucrit$. Data here is a subset of that shown in fig. 3a from the main text. (EPS)

Figure S5 Quantitative version of fig. 3c from the main text. As in fig. 3c, here we see a boundary between phases that increases up and to the right. However, fluctuations at very low N inevitably obscure the underlying phase boundary. Points in the survival regime were colored white (i.e. we assumed $CV\tau = 1$) if extinction never occurred during simulations during a feasible amount of time (10^5 generations). $\delta=0$. (EPS)

Figure S6 Lethal mutations during simulations on our bio-physical landscape when $\delta=0$. As in fig. 4 from the main text, this plot shows data only from (N,U) pairs such that populations survived for the duration of simulations (105 generations). However, the left-most terminus of each curve is near the extinction phase. In our non-epistatic, analytical model, the boundary between survival and extinction phases occurs at $U\lambda = ln(2)$. Here we see that the criterion that $U\lambda = ln(2)$ is unlikely to mark the transition to extinction on our (epistatic) biophysical model, since the left-most termini always have $U\lambda$ substantially below $ln(2)$. (EPS)

Text S1 Analytical derivations of mean and variance in extinction time on flat, non-epistatic fitness landscape. (PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: CSW EIS. Performed the experiments: CSW. Analyzed the data: CSW EIS. Contributed reagents/materials/analysis tools: CSW. Wrote the paper: CSW.

References

1. Loeb LA, Essigmann JM, Kazazi F, Zhang J, Rose KD, et al. (1999) Lethal mutagenesis of HIV with mutagenic nucleoside analogs. Proc Natl Acad Sci U S A 96: 1492–1497.
2. Wilke CO, Bull JJ, Sanjuan R (2007) Theory of lethal mutagenesis for viruses. J Virol 81: 2930–2939.
3. Zeldovich KB, Chen P, Shakhnovich EI (2007) Protein stability imposes limits on organism complexity and speed of molecular evolution. Proc Natl Acad Sci U S A 104: 16152.
4. Wilke CO, Bull JJ (2008) Lethal Mutagenesis of Bacteria. Genetics 180: 1061–1070.
5. Chen P, Shakhnovich EI (2009) Lethal mutagenesis in viruses and bacteria. Genetics 183: 639–650.
6. O'dea EB, Keller TE, Wilke CO (2010) Does Mutational Robustness Inhibit Extinction by Lethal Mutagenesis in Viral Populations? PLoS Comput Biol 6: e1000811.
7. Martin G, Gandon S (2010) Lethal mutagenesis and evolutionary epidemiology. Philos Trans R Soc B Biol Sci 365: 1953–1963.
8. Domingo E, Sierra S, Davila M, Lowenstein PR (2000) Response of foot-and-mouth disease virus to increased mutagenesis: Influence of viral load and fitness in loss of infectivity. J Virol 74: 8316–8323.
9. Domingo E, Pariente N, Sierra S, Lowenstein PR (2001) Efficient virus extinction by combinations of a mutagen and antiviral inhibitors. J Virol 75: 9723–9730.
10. Lowenstein PR, Grande-Perez A, Sierra S, Castro MG, Domingo E (2002) Molecular indetermination in the transition to error catastrophe: Systematic elimination of lymphocytic choriomeningitis virus through mutagenesis does not correlate linearly with large increases in mutant spectrum complexity. Proc Natl Acad Sci U S A 99: 12938–12943.
11. Anderson JP, Daifuku R, Loeb LA (2004) Viral error catastrophe by mutagenic nucleosides. Annu Rev Microbiol 58: 183–205.
12. Freistadt MS, Meades GD, Cameron CE (2004) Lethal mutagens: broad-spectrum antivirals with limited potential for development of resistance? Drug Resist Updat 7: 19–24.
13. Cameron CE, Graci JD, Harki DA, Korneeva VS, Edathil JP, et al. (2007) Lethal mutagenesis of poliovirus mediated by a mutagenic pyrimidine analogue. J Virol 81: 11256–11266.
14. Cameron CE, Graci JD, Too K, Smidansky ED, Edathil JP, et al. (2008) Lethal mutagenesis of picornaviruses with n-6-modified purine nucleoside analogues. Antimicrob Agents Chemother 52: 971–979.
15. Lynch M (2007) The frailty of adaptive hypotheses for the origins of organismal complexity. Proc Natl Acad Sci U S A 104: 8597–8604.
16. Kimura M (1985) The neutral theory of molecular evolution. Cambridge University Press.
17. Felsenstein J (1974) The evolutionary advantage of recombination. Genetics 78: 737–756.
18. Haigh J (1978) The accumulation of deleterious genes in a population-Muller's Ratchet. Theor Popul Biol 14: 251–267.
19. Lynch M, Gabriel W (1990) Mutation load and the survival of small populations. Evolution 44: 1725–1737.
20. Lynch M, Burger R, Butcher D, Gabriel W (1993) The mutational meltdown in asexual populations. J Hered 84: 339–44.
21. Lande R (1994) Risk of Population Extinction from Fixation of New Deleterious Mutations. Evolution 48: 1460–1469.
22. Schultz ST, Lynch M (1997) Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. Evolution: 1363–1371.
23. Whitlock MC (2000) Fixation of new alleles and the extinction of small populations: Drift load, beneficial alleles, and sexual selection. Evolution 54: 1855–1861.
24. Gordo I, Charlesworth B (2000) The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics 154: 1379–1387.
25. Halligan DL, Keightley PD (2009) Spontaneous mutation accumulation studies in evolutionary genetics. Annu Rev Ecol Evol Syst 40: 151–172.
26. Tannenbaum E, Deeds EJ, Shakhnovich EI (2004) Semiconservative replication in the quasispecies model. Phys Rev E Stat Nonlin Soft Matter Phys 69: 061916.
27. Wylie CS, Shakhnovich EI (2011) A biophysical protein folding model accounts for most mutational fitness effects in viruses. Proc Natl Acad Sci U S A 108: 9916–9921.
28. Ewens WJ (2004) Mathematical Population Genetics. Springer.
29. Grande-Perez A, Lazaro E, Lowenstein P, Domingo E, Manrubia SC (2005) Suppression of viral infectivity through lethal defection. Proc Natl Acad Sci U S A 102: 4448–4452.
30. Elena SF, Carrasco P, Daros JA, Sanjuan R (2006) Mechanisms of genetic robustness in RNA viruses. EMBO Rep 7: 168–173.
31. Perelson AS (2002) Modelling viral and immune system dynamics. Nat Rev Immunol 2: 28–36.
32. Drake JW, Charlesworth B, Charlesworth D, Crow JF (1998) Rates of spontaneous mutation. Genetics 148: 1667–1686.
33. Sanjuan R, Nebot MR, Chirico N, Mansky LM, Belshaw R (2010) Viral mutation rates. J Virol 84: 9733–9748.
34. Gabriel W, Lynch M, Burger R (1993) Mullers Ratchet and Mutational Meltdowns. Evolution 47: 1744–1757.
35. Markowitz M, Louie M, Hurley A, Sun E, Di Mascio M (2003) A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-Cell decay in vivo. J Virol 77: 5037–5038.
36. Charlesworth B (2009) Fundamental Concepts in Genetics: Effective population size and patterns of molecular evolution and variation. Nat Rev Genet 10: 195–205.
37. Crow J, FaK M (1970) An introduction to population genetics theory. Harper and Row Publishers.
38. Poon A, Otto SP (2000) Compensating for our load of mutations: freezing the meltdown of small populations. Evolution: 1467–1479.
39. Hill W, Robertson A (1966) The effect of linkage on limits to artificial selection. Genet Res 8: 269–294.
40. Wang Z, Zhang J (2011) Impact of gene expression noise on organismal fitness and the efficacy of natural selection. Proc Natl Acad Sci U S A 108: E67–76.
41. Kondrashov A (1994) Muller's ratchet under epistatic selection. Genetics 136: 1469.
42. Eigen M (1971) Self organization of matter and the evolution of biological macromolecules. Naturwissenschaften 58: 465–523.
43. Nowak M, Schuster P (1989) Error thresholds of replication in finite populations mutation frequencies and the onset of Muller's ratchet. J Theor Biol 137: 375–395.
44. Crotty S, Cameron CE, Andino R (2001) RNA virus error catastrophe: direct molecular test by using ribavirin. Proc Natl Acad Sci U S A 98: 6895–6900.
45. Crotty S, Andino R (2002) Implications of high RNA virus mutation rates: lethal mutagenesis and the antiviral drug ribavirin. Microbes Infect 4: 1301–1307.
46. Crotty S, Maag D, Arnold JJ, Zhong W, Lau JY, et al. (2000) The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. Nat Med 6: 1375–1379.
47. Severson WE, Schmaljohn CS, Javadian A, Jonsson CB (2003) Ribavirin causes error catastrophe during Hantaan virus replication. J Virol 77: 481–488.
48. Fares MA, Ruiz-González MX, Moya A, Elena SF, Barrio E (2002) Endosymbiotic bacteria: groEL buffers against deleterious mutations. Nature 417: 398.
49. Maisnier-Patin S, Roth JR, Fredriksson A, Nystrom T, Berg OG, et al. (2005) Genomic buffering mitigates the effects of deleterious mutations in bacteria. Nat Genet 37: 1376–1379.
50. Fernandez A, Lynch M (2011) Non-adaptive origins of interactome complexity. Nature 474: 502–505.
51. van Ham RCHJ, Kamerbeek J, Palacios C, Rausell C, Abascal F, et al. (2003) Reductive genome evolution in Buchnera aphidicola. Proc Natl Acad Sci U S A 100: 581–586.
52. Tokuriki N, Oldfield CJ, Uversky VN, Berezovsky IN, Tawfik DS (2009) Do viral proteins possess unique biophysical features? Trends Biochem Sci 34: 53–59.
53. Privalov P (1979) Stability of proteins: small globular proteins. Adv Protein Chem 33: 167.
54. Serrano L, Day AG, Fersht AR (1993) A Procedure for Engineering Increased Stability of Proteins and an Experimental Analysis of the Evolution of Protein Stability. J Mol Biol 233: 305–312.
55. Tokuriki N, Stricher F, Schymkowitz J, Serrano L, Tawfik DS (2007) The stability effects of protein mutations appear to be universally distributed. J Mol Biol 369: 1318–1332.
56. Kumar M, Bava KA, Gromiha MM, Prabakaran P, Kitajima K, et al. (2006) ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. Nucleic Acids Res 34: 204–206.