Improve it or lose it: Evolvability cost of competition for expression

Jacob Moran, Devon Finlay, and Mikhail Tikhonov*
Department of Physics, Washington University in St. Louis, St. Louis, Missouri 63130, USA

(Received 8 February 2021; revised 6 May 2021; accepted 24 May 2021; published 4 June 2021)

Expression level is known to be a strong determinant of a protein's rate of evolution. But the converse can also be true: evolutionary dynamics can affect expression levels of proteins. Having implications in both directions fosters the possibility of an "improve it or lose it" feedback loop, where higher expressed systems are more likely to improve and be expressed even higher, while those that are expressed less are eventually lost to drift. Using a minimal model to study this in the context of a changing environment, we demonstrate that one unexpected consequence of such a feedback loop is that a slow switch to a new environment can allow genotypes to reach higher fitness sooner than a direct exposure to it.

DOI: 10.1103/PhysRevE.103.062402

I. INTRODUCTION

The rates of protein evolution are affected by a multitude of factors, including protein-protein interactions, stability-based constraints, or dispensability [1–10]. However, the strongest single determinate appears to be expression level [11,12]. For instance, substantial evidence suggests that lower-expressed proteins are less protected from drift, whereas highly expressed proteins are under stronger purifying selection [11–14].

Conversely, the expression level can itself be affected by evolution, especially for proteins or pathways that are dispensable or partially redundant. For example, a protein that is disabled by a deleterious mutation becomes a metabolic burden (or may be directly toxic), favoring a reduction in expression.

Since partial redundancy is believed to be widespread [15], this creates a theoretical possibility of a feedback loop. Consider an organism with several partially substitutable systems or pathways fulfilling a similar function; for example, several metabolic pathways to satisfy its requirement for carbon, or several sensing modalities to respond to environmental cues. In these circumstances, it seems plausible that the systems used more, being under a stronger selection pressure, would be more likely to improve and be used even more. In contrast, the lesser expressed systems could be more likely to deteriorate and be used even less (Fig. 1).

This process, effectively a "competition for expression", could be viewed as an extension of Savageau's "use it or lose it" principle and is conceptually similar to the generalist-to-specialist transition of ecological specialization [16,17], but is rarely discussed in an evolutionary context. One reason, perhaps, is that this intuition appears to predict that highly expressed proteins should evolve faster, the opposite of what is observed empirically [11]. However, as we will show, the consequences of such a feedback interaction are more nuanced.

To do so, we illustrate the feedback loop of Fig. 1 in a simple minimal model. For highly adapted systems depleted for beneficial mutations, we find that the highest-expressed proteins are still expected to evolve slowest, in full agreement with the empirical observations. In contrast, for an evolutionary process driven by strong adaptive mutations, e.g., following a strong environmental change, the sign of the correlation between expression and evolutionary rate is predicted to transiently invert. Moreover, at least in our model, the consequences of the "improve it or lose it" feedback include interesting qualitative effects, such as a loss of evolvability caused by an environmental perturbation that is too strong. As an example, we demonstrate how a gradual change to a new environment can lead to a higher rate of fitness gain than direct exposure.

II. MODEL AND CONTEXT

To study the "improve it or lose it" feedback loop, we need an evolutionary model that explicitly includes a notion of usage or expression. For this reason, we adopt the toolbox model from Ref. [18], summarized in Fig. 2(a).

Briefly, we think of a genotype as encoding a set of K systems that can be used at different levels to optimize the fitness of the organism in a given environment. Mathematically, we represent the K systems as basis vectors $\{\vec{g}_\mu\}$ ($\mu = 1, \ldots, K$) and the environment as a target vector $\vec{E}$ in an abstract L-dimensional space (which can be interpreted as the phenotype space [18]). The fitting problem can be written as

$$\{a_\mu\} = \underset{a_\mu \geqslant 0}{\operatorname{argmin}} \left\| \vec{E} - \sum_\mu a_\mu \vec{g}_\mu \right\|, \tag{1}$$

where the environment-dependent coefficients $\{a_\mu\}$ can be interpreted as the extent to which the organism relies on a given system $\vec{g}_\mu$ in $\vec{E}$. The quality of fit, which these $\{a_\mu\}$ optimize, can then be interpreted as the fitness of the genotype $G = \{\vec{g}_\mu\}$ in environment $\vec{E}$:

$$F(\mathrm{G}, \vec{E}) = -\min_{a_\mu \geqslant 0} \left\| \vec{E} - \sum_\mu a_\mu \vec{g}_\mu \right\|. \tag{2}$$

*Corresponding author: tikhonov@wustl.edu

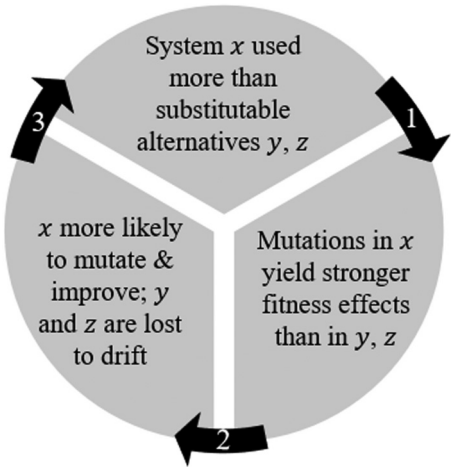

FIG. 1. *The "improve it or lose it" feedback loop.* In this schematic, x, y, and z are partially substitutable systems fulfilling a similar function (e.g., metabolic pathways for alternative sources of carbon). Adaptive mutations in the highest-used system x have stronger fitness effects than y, z (arrow 1). The stronger selection pressure makes system x more likely to mutate and improve (arrow 2). This improvement in x allows the organism to rely on it even more (arrow 3), completing the loop.

In Ref. [18], the coefficients $\{a_\mu\}$ are called "expression levels"; however, conceptually, they correspond more closely to the intuitive notion of "usage." Indeed, a larger a_μ in this model corresponds to a system whose deletion would have a stronger fitness effect, rather than one present in a larger copy number (although in practice, the two properties are, of course, correlated [8]). Throughout this work, we refer to $\{a_\mu\}$ as usage coefficients.

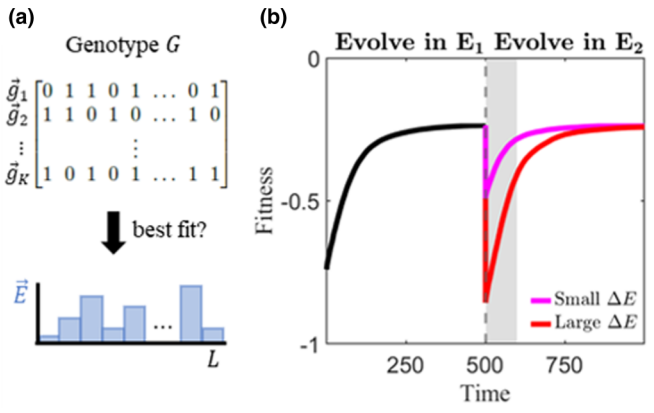

FIG. 2. *A context to study the "improve it or lose it" feedback loop.* (a) In the toolbox model, a genotype is a matrix representing the available "systems" an organism can (linearly) combine to approximate the optimal phenotype required by the environment, $\vec{E}$. The coefficients of the best approximation are interpreted as usage levels a_μ, serving as a proxy for expression. Matrix elements are chosen to be binary (0 or 1) so that mutations in the evolutionary process can be implemented as bit flips. (b) Fitness trajectories of initially random genotypes evolving under $\vec{E}_1$ before switching to $\vec{E}_2$ a distance ΔE away. We choose to study the feedback loop and its consequences during the early-time dynamics after switching (gray region).

For simplicity in simulating evolution within this model, we assume that mutations are rare and selection is strong, so that we need only track the evolutionary trajectory of a single genotype [19]. In each simulation step, we enumerate all beneficial point mutations of the current genotype by performing all single bit flips of the genotype matrix. We then pick one of these mutations as the first to rise to fixation; in this parameter regime, selection considers only beneficial mutations, and fixation probability is proportional to a mutation's fitness effect. We note that, when evaluating the fitness of mutants, the usage coefficients are optimized for the mutated genotype, and thus typically differ from those of the parent. This corresponds to the assumption that the evolution of $\{a_\mu\}$ occurs on a much faster timescale than evolution of system vectors $\{\vec{g}_\mu\}$ (a separation of timescales; see the Supplemental Material [20] for more discussion).

Figure 2(b) shows an example of fitness dynamics of random initial genotypes first exposed to a random environment $\vec{E}_1$ and then to a different random environment $\vec{E}_2$. The feedback loop we will describe is already present during the early-time dynamics of evolution in $\vec{E}_1$; however, we choose to focus on the time period that follows the environment switch (shaded gray region). This will allow us to use the difference between the two environments, $\Delta E = \|\vec{E}_2 - \vec{E}_1\|$ as a natural control parameter [see the Supplemental Material [20] for parametrization of environment pairs $(\vec{E}_1, \vec{E}_2)$].

In what follows, we use $\vec{E}$ vectors of unit length so that fitness is constrained to $-1 \leqslant F \leqslant 0$. We fix $L = 40$ and vary K, and consider genotype matrices with binary values, 0 or 1, initialized randomly with probability $p = 0.5$ of being 1. Since environments are represented by unit vectors with positive components, ΔE is confined to the range $\Delta E \in [0, \sqrt{2}]$. We will show that ΔE controls the strength of the feedback loop, with stronger changes in environment (large ΔE) inducing stronger feedback.

III. THE TOOLBOX MODEL EXHIBITS THE "IMPROVE IT OR LOSE IT" FEEDBACK

Figure 3(a) depicts a representative trajectory of the "improve it or lose it" feedback realized in the toolbox model. The panel shows the dynamics of usage coefficients after a genotype with $K = 5$ systems, preadapted to some environment $\vec{E}_1$, was switched to a different environment $\vec{E}_2$, with $\Delta E = 1$ (random environment pairs with a given ΔE were generated as described in the Supplemental Material [20]). Note that, after each mutation, the usage coefficients are reoptimized according to Eq. (1) and thus change discontinuously (see the Supplemental Material [20] and Ref. [18]); however, these steps are typically small, creating an illusion of smooth dynamics. We see that strong adaptive mutations initially concentrate in the two systems with highest usage (frequent redder dots). As they mutate, they also rise in usage, a_μ. In contrast, the lower-used systems decrease in usage and mutate only rarely, with relatively weak fitness effects (bluer dots).

Although the details of these dynamics are shaped by Eq. (2) and are of course model-dependent, on a qualitative level the instability driving a subset of usage coefficients up

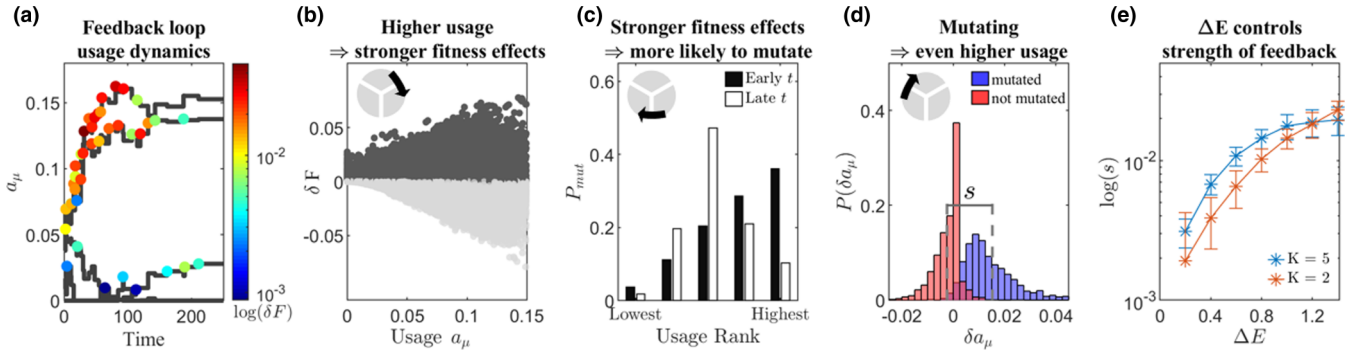

FIG. 3. *The toolbox model exhibits the feedback loop.* (a) An example of evolutionary dynamics of usage coefficients after a genotype adapted to a random environment E_1 is switched to another random environment E_2 with $\Delta E \equiv |\vec{E}_1 - \vec{E}_2| = 1$. Despite similar usage initially, by $t = 100$ only two of $K = 5$ systems remain in use. Dots mark the systems in which a beneficial mutation arose, color indicates fitness effect (red is strongest). In panels (b)–(d), we examine the statistics of usage dynamics and mutation effects within the first three time steps of 20 trajectories in 15 random environment pairs with the same $\Delta E = 1$. Inset pictograms refer to feedback steps as shown in Fig. 1. (b) Fitness effects of all available mutations in each system versus system usage. Dark and light gray points are beneficial and neutral or deleterious mutations, respectively. Higher-used systems possess stronger fitness effects. (c) Probability of a system to mutate, plotted against its usage rank (ascending order). At early times (black bars), higher used systems are more likely to mutate. As the strong beneficial mutations in highest-used systems are depleted, the probability of mutating shifts towards lower used systems (white bars). (d) Distribution of change in usage of a system that just mutated (blue) or a system that failed to mutate (red) at a particular simulation step. The difference in means of these conditional probability distributions, s, quantifies the strength of the feedback loop. (e) The strength of the feedback loop s is controlled by the magnitude of environmental change ΔE. Error bars represent one standard deviation (SD) over 300 replicates.

at the expense of others can be directly traced to the feedback loop summarized in Fig. 1, as we will now show.

First, agreeing with the intuitive notion of $\{a_\mu\}$ as "usage," systems with higher a_μ tend to harbor stronger fitness effects. To see this in our model, we plot the fitness effects of all available mutations within the first few simulation steps against the usage coefficient of the system where they occur [Fig. 3(b)]. As expected, both beneficial (dark gray) and deleterious (light gray) mutations are stronger in systems that have a higher usage coefficient a_μ.

As a result, higher-used systems are more likely to mutate, because mutations with a larger fitness effect are more likely to escape drift and fix in the population [21]. The black bars in Fig. 3(c) show the early-time probability of each system to mutate, plotted against its usage rank.

Finally, Fig. 3(d) shows the distribution of usage changes, defined as the difference in usage δa_μ before and after a simulation step, over the same early time period as described above. Whenever a system mutates, its usage typically increases [Fig. 3(d), blue]. In contrast, the systems that did not mutate at that particular timestep typically drop in usage [Fig. 3(d), red]. In our model, this also is ultimately a consequence of Eq. (2), but it is not the model that justifies this behavior. Rather, it is this behavior that justifies using the model, making it appropriate for studying the feedback loop that this behavior induces. In summary, Figs. 3(b)–3(d) demonstrate all three arrows from Fig. 1 at play in our model.

Since a greater separation between the distributions of Fig. 3(d) would entail stronger feedback, we can use the difference in the mean of these conditional distributions, denoted as s, as a measure of the feedback strength. Figure 3(e) demonstrates that, as expected, the feedback becomes stronger (increasing $\log s$) as the change in environment becomes more severe.

The rapid evolution of highly used systems [Fig. 3(c)] may seem to be at odds with experimental work showing that highly expressed proteins evolve slowest [11]. However, the mechanism described here is fully compatible with the explanations previously proposed for this experimental result. The effect shown in Fig. 3(b) (higher used systems have stronger fitness effects) applies to both beneficial and deleterious mutations. For early stages of adaptation driven by beneficial mutations (as considered here), this means the most-used systems will evolve first. However, at later stages, as beneficial mutations are depleted, the same argument dictates that the most-used systems become the most protected and evolve slowest. We demonstrate the presence of this effect by replotting the per-system mutation probabilities at a later time [Fig. 3(c), white bars]; the probability of mutating begins to shift from higher used to lesser used systems. This result therefore predicts that the negative correlation between expression and evolution rate observed in [11] should transiently invert following a change in environment. If additional factors like interaction and stability constraints on evolutionary rates are considered, our framework predicts that the negative correlation would at least weaken, with the size of the effect controlled by the magnitude of the environmental change. Encouragingly, this transient weakening in negative correlation between expression and evolutionary rate is consistent with recent analysis of evolutionary rates in yeast [22].

IV. THE COST TO EVOLVABILITY

Intuitively, one might expect that the competition for usage mediated by the "improve it or lose it" feedback loop may be detrimental for the organism, since it effectively reduces the number of systems it has available. Implementing this effect in a simple model allows us to make this intuition precise. We will see that, at least in our model, the feedback loop exhibited

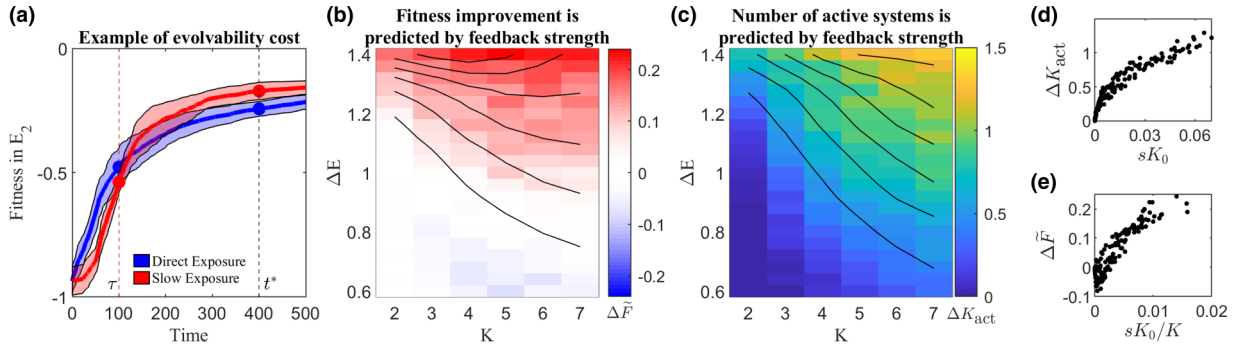

FIG. 4. *Higher evolvability from slow exposure than direct exposure.* (a) Fitness, evaluated in environment $\vec{E}_2$, of genotypes that are either directly exposed (DE) to $\vec{E}_2$ at $t = 0$ (blue trace) or slowly exposed (SE) to $\vec{E}_2$ over a time $\tau = 100$ according to the protocol defined in Eq. (3) (red trace). Each trace shows mean ± 1 SD (shading) of 20 replicate trajectories of genotypes with $K = 4$ systems in a random environment pair with $\Delta E = 1.4$. Colored dots highlight that slow exposure leads to higher long-term fitness, despite slower fitness gain initially. The relative improvement in fitness, $\Delta\widetilde{F}$, is measured at an arbitrarily late time point $t^* = 400$ (see the Supplemental Material [20] Fig. S3C for later t^*). (b) Heatmap of the long-term relative fitness improvement, $\Delta\widetilde{F}$. Contour lines show $\Delta\widetilde{F}$ can be predicted by the feedback loop strength s and the number of initially inactive systems K_0 (see panel e). Here and in the remaining panels, results are averages over 20 trajectories in 15 random environment pairs with varying ΔE. (c) Heatmap of $\Delta K_{\rm act}$, the average difference in number of active systems ($a_\mu > 10^{-3}$) between the SE and DE protocols at $t = \tau$. Contour lines show it is predicted by the product sK_0; see panel (d). (d) $\Delta K_{\rm act}$, the increased number of active systems at $t = \tau$, is predicted by sK_0, measured at trajectory start. (e) The long-term fitness improvement $\Delta\widetilde{F}$ at $t = t^*$ is predicted by sK_0/K, measured at trajectory start.

above reduces the adaptive potential of the genotype, and mitigating its effects can allow for faster adaptation.

For this, we compare the fitness trajectories of genotypes evolving in conditions that exacerbate the feedback and those that weaken it. Specifically, starting from a genotype preadapted to $\vec{E}_1$, we compare two ways of adapting it to a new, strongly different environment $\vec{E}_2$: either by exposing it to $\vec{E}_2$ directly (as discussed above), or by changing the environment from $\vec{E}_1$ to $\vec{E}_2$ slowly (on a timescale that is slow compared to mutation fixation). By avoiding large environment jumps, we expect the gradual switch to weaken the feedback loop. The question we ask is which exposure protocol will ultimately lead to higher fitness in the environment of interest, $\vec{E}_2$.

An example of this comparison is shown in Fig. 4(a). The red curve shows fitness (in the environment of interest $\vec{E}_2$) for genotypes evolving under the slow-exposure protocol, implemented by linearly relaxing the environment vector from $\vec{E}_1$ to $\vec{E}_2$ over a time τ:

$$\vec{E}(t) = \begin{cases} \text{normalize}\left[\vec{E}_2 + \dfrac{\tau - t}{\tau}(\vec{E}_1 - \vec{E}_2)\right] & \text{if } t < \tau \\ \vec{E}_2 & \text{if } t \geqslant \tau \end{cases} \tag{3}$$

(the environment vector in our model is always normalized to unit length). The τ we use is large relative to the typical time between mutations [$\tau = 100$; compare to Fig. 3(a)]. The red curve $F_{\rm SE}(t)$ (slow exposure) is to be compared to the blue curve $F_{\rm DE}(t)$ (direct exposure), showing fitness of the same initial genotypes evolving directly in $\vec{E}_2$.

The vertical dashed line at $t = \tau$ marks the timepoint where the "red genotypes" evolving under the slow-switching protocol are finally exposed to $\vec{E}_2$ for the first time. It is therefore not surprising that they are less fit than the "blue genotypes," who have been evolving in $\vec{E}_2$ from the start ($F_{\rm SE}(\tau) < F_{\rm DE}(\tau)$; red curve below the blue). However, while more fit, the blue genotypes are manifestly less evolvable: From $t = \tau$ onwards, both red and blue curves document evolution in the same environment $\vec{E}_2$, but the red curve gains fitness much faster, and overtakes the blue.

To quantify the strength of this effect, we consider the relative improvement of fitness provided by the smooth protocol, compared to direct exposure:

$$\Delta\widetilde{F}(t^*) \equiv \frac{F_{\rm SE}(t^*) - F_{\rm DE}(t^*)}{|F_{\rm DE}(t^*)|}. \tag{4}$$

While initially negative, in the example of Fig. 4(a) this quantity becomes positive at a later time. To demonstrate the robustness of this observation, Fig. 4(b) shows $\Delta\widetilde{F}(t^*)$ for a range of K and ΔE, computed at an arbitrary late timepoint $t^* = 400$ (see the Supplemental Material [20] Fig. S3C for $\Delta\widetilde{F}$ at a later value of t^*). We see that, at large ΔE, the slow-switching protocol consistently outperforms direct exposure, and more so as K increases. While the scenario of an organism possessing $K = 7$ competing systems fulfilling a similar function is arguably unrealistic, we note that the effect is already present at $K = 2$. (For the purposes of illustration, the example in Fig. 4(a) used $K = 4$ and a dramatic environment change $\Delta E = 1.4$, when the effect is strongest.) Note that, for simplicity, in Fig. 4(b) our slow exposure protocol (3) used the same value of the relaxation time $\tau = 100$ for all K and ΔE; optimizing over this parameter and the observation timepoint t^* could of course render the effect stronger.

The origin of this effect is the "improve it or lose it" instability affecting the genotypes undergoing an abrupt environment switch, effectively leaving them with fewer systems. To confirm this, we record the average number $K_{\rm act}^{\rm DE}$, $K_{\rm act}^{\rm SE}$ of "active" systems (usage $a_\mu > 0.001$) observed at time $t = \tau$ under both protocols. As expected, a slow environment change leaves more systems active; the difference $\Delta K_{\rm act} \equiv K_{\rm act}^{\rm SE} - K_{\rm act}^{\rm DE}$ is shown in Fig. 4(c) and exhibits a trend similar to Fig. 4(b). Since unused systems harbor weak mutations only

[cf. Fig. 3(b)], a genotype with few active systems finds itself on a fitness plateau, and its rate of fitness gain is reduced.

Finally, we can quantitatively relate both effects to the strength of the feedback loop as defined above. To start, we focus on the increase in the number of active systems ΔK_{act} in Fig. 4(c). Denote K_0 the number of inactive systems at time $t = 0$ (immediately after the environment switch; usage $a_\mu < 0.001$). This is the number of systems that the slow-exposure protocol could conceivably "rescue." One expects ΔK_{act} to scale with K_0, and if our argument is correct, it should also scale with the strength of the feedback loop s. Indeed, we find ΔK_{act} to be predicted by the product sK_0 [Fig. 4(d)]. The availability of these additional systems translates into additional adaptive opportunities and ultimately a higher fitness. In a strongly epistatic model like ours, the exact relationship to the long-term fitness is hard to predict. Nevertheless, it is reasonable to expect the fractional effect on fitness $\Delta\widetilde{F}$ to at least correlate with the fractional effect on the number of active systems $\Delta K_{\text{act}}/K$. If so, then $\Delta\widetilde{F}$ should correlate with sK_0/K, an expectation confirmed in Fig. 4(e). Given the approximate nature of this argument, the correlation observed in Fig. 4(e) is in fact surprisingly good. For convenience, the same sK_0 and sK_0/K data, Gaussian smoothed for visualization purposes, are shown as contour lines superimposed on the heatmaps of Figs. 4(b) and 4(c). It is worth emphasizing that our definition of the feedback strength s is computed from the statistics of the first three mutations, which take only $t \sim 7 \pm 5$ to occur; and K_0 is similarly measured at the very start of the trajectory. Nevertheless, at least in our model, these early-time properties are predictive of the long-term evolutionary outcome at $t^* = 400$.

V. DISCUSSION

In this work, we used a minimal model to explore a possible feedback loop between the usage of a system and its rate of evolution. Within this model, we demonstrated that this feedback loop is particularly pronounced after strong shifts in the selecting environment and can negatively impact evolvability (future fitness gain). In particular, we described a mechanism by which a slow switch to a new environment can allow the genotypes to reach higher fitness sooner than a direct exposure to it. Interestingly, this effect is reminiscent of recent results from the Evolthon crowdsourcing effort, which found that when yeast and *E. coli* populations are slowly exposed to cold temperatures they attain higher fitness than those that undergo a direct exposure [23].

A situation where exposure to a different environment E' can help evolve better fitness in E than a direct exposure to E itself is not, in itself, novel. One well-established scenario for this to occur is the crossing of fitness valleys (or plateaus): much like an enzyme that catalyzes a reaction by stabilizing the reaction transition state, a transient exposure to E' can facilitate reaching a higher fitness peak by enabling prerequisite mutations that would otherwise be unfavorable (or neutral) [24,25]. However, the scenario described here is particularly interesting because the fitness plateau is not an idiosyncratic property of a particular landscape, but emerges through evolution itself. Fitness landscapes of evolved systems are themselves shaped by evolution [26,27], and at least in our model, the feedback mechanism we described generically induces a fitness plateau following a sufficiently strong environmental change.

To focus on this effect, our proof-of-principle model ignored many other factors contributing to rates of protein evolution. In any realistic scenario, the feedback interaction we described will only be a part of a larger picture. Nevertheless, our analysis predicts that the empirically observed negative correlation between expression and evolution rate would transiently weaken following a change in environment ΔE, and this weakening should be more pronounced for stronger ΔE. We expect this effect to be more evident if other constraints not included in our model are weakened, following, for example, a genome duplication event [22].

It is worth stressing that we considered beneficial mutations only. Clearly, if deleterious mutations were included, our feedback loop would become even stronger: in addition to the effect described, the lesser-used systems would also be less protected from drift [28–30]. This observation could then be seen as the traditional manifestation of the "use it or lose it" principle; in particular, the problem of maintaining redundancy in the face of drift has been extensively discussed [31]. Focusing on beneficial mutations only, and thus explicitly excluding any drift-dependent effects, allows us to highlight a novel aspect. Unlike the discussion of Ref. [31], here no system is ever fully redundant, and all remain under selection. Nevertheless, some are progressively lost even in the absence of deleterious mutations—simply because the beneficial mutations preferentially target the systems used more, and those that fail to improve become obsolete. This mechanism is clearly analogous to the Red Queen effect [32] (to remain useful, a system must keep improving), except here it applies to an effective competition for expression. In this way, the loss of evolvability described in Fig. 4 can be seen as a form of a conflict of levels of selection [33]: the competitive dynamics between lower-level entities (the K "systems" in our model) lead to negative consequences for the organism as a whole—a decline of phenotypic flexibility and evolvability due to a reduction of the effective K. On a related note, while our model considered K as a fixed parameter, it could easily be extended to allow for system loss and duplication events.

All simulations were performed in MATLAB (Mathworks, Inc.). The associated code, data, and scripts to reproduce all figures in this work are available at Mendeley Data [34].

ACKNOWLEDGMENTS

We thank C. Holmes, S. Kaplan, S. Kuehn, and S. Maslov for helpful discussions.

The authors declare no competing interests.

[1] K. Jordan, Y. I. Wolf, and E. V. Koonin, No simple dependence between protein evolution rate and the number of protein-protein interactions: Only the most prolific interactors tend to evolve slowly, BMC Evol. Biol. **3**, 1 (2003).

[2] T. Friedlander, R. Prizak, N. H. Barton, and G. Tkacik, Evolution of new regulatory functions on biophysically realistic fitness landscapes, Nat. Commun **8**, 216 (2017).
[3] G. Reddy and M. M. Desai, Global epistasis emerges from a generic model of a complex trait, eLife **10**, e64740 (2021).
[4] N. Tokuriki and D. S. Tawfik, Stability effects of mutations and protein evolvability, Curr. Opin. Struct. Biol. **19**, 596 (2009).
[5] J. D. Bloom, S. T. Labthavikul, C. R. Otey, and F. H. Arnold, Protein stability promotes evolvability, Proc. Natl. Acad. Sci. USA **103**, 5869 (2006).
[6] M. Goldsmith and D. S. Tawfik, Potential role of phenotypic mutations in the evolution of protein expression and stability, Proc. Natl. Acad. Sci. USA **106**, 6197 (2009).
[7] A. E. Hirsh and H. B. Fraser, Protein dispensability and rate of evolution, Nature (London) **411**, 1046 (2001).
[8] J. Zhang and X. He, Significant impact of protein dispensability on the instantaneous rate of protein evolution, Mol. Biol. Evol. **22**, 1147 (2005).
[9] W. P. Wall, A. E. Hirsh, H. B. Fraser, J. Kumm, G. Giaever, M. B. Eisen, and M. W. Feldman, Functional genomic analysis of the rates of protein evolution, Proc. Natl. Acad. Sci. USA **102**, 5483 (2005).
[10] B.-Y. Liao, N. M. Scott, and J. Zhang, Impacts of gene essentiality, expression pattern, and gene compactness on the evolutionary rate of mammalian proteins, Mol. Biol. Evol. **23**, 2072 (2006).
[11] D. A. Drummond, J. D. Bloom, C. Adami, C. O. Wilke, and F. H. Arnold, Why highly expressed proteins evolve slowly, Proc. Natl. Acad. Sci. USA **102**, 14338 (2005).
[12] J. Zhang and J. R. Yang, Determinants of the rate of protein sequence evolution, Nat. Rev. Genet. **16**, 409 (2015).
[13] M. A. Savageau, Design of molecular control mechanisms and the demand for gene expression, Proc. Natl. Acad. Sci. USA **74**, 5647 (1977).
[14] U. Gerland and T. Hwa, Evolutionary selection between alternative modes of gene regulation, Proc. Natl. Acad. Sci. USA **106**, 8841 (2009).
[15] R. Kafri, M. Springer, and Y. Pilpel, Genetic redundancy: New tricks for old genes, Cell **136**, 389 (2009).
[16] V. S. Cooper and R. E. Lenski, The population genetics of ecological specialization in evolving *Escherichia coli* populations, Nature (London) **407**, 736 (2000).
[17] T. J. Kawecki, N. H. Barton, and J. D. Fry, Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation, J. Evol. Biol. **10**, 407 (1997).
[18] M. Tikhonov, S. Kachru, and D. S. Fisher, A model for the interplay between plastic tradeoffs and evolution in changing environments, Proc. Natl. Acad. Sci. USA **117**, 8934 (2020).
[19] J. H. Gillespie, A simple stochastic gene substitution model, Theor. Popul. Biol. **23**, 202 (1983).
[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.103.062402 for a discussion of model parameters, a description of the evolutionary process and other technical details, which includes Ref. [35].
[21] M. Kimura, On the probability of fixation of mutant genes in a population, Genetics **47**, 713 (1962).
[22] D. Ascencio, S. Ochoa, L. Delaye, and A. DeLuna, Increased rates of protein evolution and asymmetric deceleration after the whole-genome duplication in yeasts, BMC Evo. Bio. **17** (2017).
[23] S. Kaminski Strauss, D. Schirman, G. Jona, A. N. Brooks, A. M. Kunjapur, A. N. Nguyen Ba, A. Flint, A. Solt, A. Mershin, A. Dixit *et al.*, Evolthon: A community endeavor to evolve lab evolution, PLoS Biol. **17**, e3000182 (2019).
[24] B. Steinberg and M. Ostermeier, Environmental changes bridge evolutionary valleys, Sci. Adv. **2**, e1500921 (2016).
[25] A. F. Bitbol and D. J. Schwab, Quantifying the role of population subdivision in evolution on rugged fitness landscapes, PLoS Comput. Biol. **10**, e1003778 (2014).
[26] D. P. Rice, B. H. Good, and M. M. Desai, The evolutionarily stable distribution of fitness effects, Genetics **200**, 321 (2015).
[27] H. A. Orr, The distribution of fitness effects among beneficial mutations, Genetics **163**, 1519 (2003).
[28] F. W. Allendorf, Protein polymorphism and the rate of loss of duplicate gene expression, Nature (London) **272**, 76 (1978).
[29] M. Kimura and J. L. King, Fixation of a deleterious allele at one of two "duplicate" loci by mutation pressure and random drift, Proc. Natl. Acad. Sci. USA **76**, 2858 (1979).
[30] T. Ohta, Time for spreading of compensatory mutations under gene duplication, Genetics **123**, 579 (1989).
[31] M. A. Nowak, M. C. Boerlijst, J. Cooke, and J. Smith, Evolution of genetic redundancy, Nature (London) **388**, 167 (1997).
[32] L. V. Valen, A new evolutionary law, Evol. Theory **1**, 1 (1973).
[33] S. Okasha, *Evolution and the Levels of Selection* (Oxford University Press, Oxford, 2006).
[34] J. Moran, Dataset for "Improve it or lose it: Evolvability cost of competition for expression", doi: 10.17632/zdsnttv2dt.2, Mendeley Data, 2021.
[35] G. Martin, Fisher's geometrical model emerges as a property of complex integrated phenotypic networks, Genetics **197**, 237 (2014).