Drift barriers for the proofreading of genes expressed at different levels

Xiong K[*], McEntee JP[†], Porfirio DJ[‡], Masel J[†]

[*]Department of Molecular & Cellular Biology, [†]Department of Ecology & Evolutionary Biology,

[‡]Department of Computer Science, University of Arizona, Tucson AZ 85721

7 Corresponding author: Joanna Masel 1041 E Lowell St, Tucson AZ 85721 USA

8 masel@email.arizona.edu Ph. +1 520-626-9888

9 Running title: Drift barriers for proofreading

10 Keywords: cryptic genetic variation, stop codon readthrough, robustness, evolvability,

11 transcriptional errors

ABSTRACT

Gene expression is imperfect, sometimes leading to toxic products. Solutions take two forms: globally reducing error rates, or ensuring that the consequences of erroneous expression are relatively harmless. The latter is optimal, but because it must evolve independently at so many loci, it is subject to a stringent "drift barrier" – a limit to how weak the effects of a deleterious mutation s can be, while still being effectively purged by selection, expressed in terms of the population size N of an idealized population such that purging requires $s < -1/N$. In previous work, only large populations evolved the optimal local solution, small populations instead evolved globally low error rates, and intermediate populations were bistable, with either solution possible. Here we take into consideration the fact that the effectiveness of purging varies among loci, because of variation in gene expression level and variation in the intrinsic vulnerabilities of different gene products to error. The previously found dichotomy between the two kinds of solution breaks down, replaced by a gradual transition as a function of population size. In the extreme case of a small enough population, selection fails to maintain even the global solution against deleterious mutations, explaining recent experimental findings of similarly high transcriptional error rates in large-N_e and small-N_e species. As expected from previous work, the evolvability of a population tracks the number of loci at which erroneous expression is tolerated.

INTRODUCTION

In classical population genetic models of idealized populations, the probability of fixation of a new mutant depends sharply on the product of the selection coefficient *s* and the population size *N*. As *s* falls below -1/*N*, fixation probabilities drop exponentially, corresponding to efficient selective purging of deleterious mutations. For $s > -1/N$, random genetic drift makes the fate of new mutants less certain. This nonlinear dependence of fixation probability on *sN* has given rise to the "drift barrier" hypothesis (Lynch 2007), which holds that populations are characterized by a threshold or "barrier" value of the selection coefficient *s*, corresponding to the tipping point at which the removal of deleterious mutations switches between effective and ineffective. In the idealized populations described by Wright-Fisher or Moran models, the drift barrier is positioned at $s = \sim -1/N$. Drift barriers also exist, albeit sometimes with less abrupt threshold behavior, in more complex models of evolution in which some assumptions of an idealized population are relaxed (Good and Desai 2014).

The drift barrier theory argues that variation among species in their characteristic threshold values for *s*, thresholds that are equal by definition to the inverse of the selection effective population size N_e, can explain why different species have different characteristics, e.g. streamlined versus bloated genomes (Lynch 2007). The simplest interpretation of the drift barrier would seem to imply that large-N_e species show higher levels of fidelity over all biological processes, e.g. DNA replication, transcription, and translation, than small-N_e species, because molecular defects that reduce fidelity are less effectively purged in the latter (Lynch 2010; Traverse and Ochman 2016).

However, as the fitness burden accumulates from the slightly deleterious mutations that a small-N_e species cannot purge, some forms of fidelity may evolve as a second line of defense. The ideal solution is to purge deleterious mutations; when this first line of defense fails, the second line of defense is to ameliorate the phenotypic consequences of deleterious mutations (Frank 2007; Rajon and Masel 2011; Warnecke and Hurst 2011; Lynch 2012; Wu and Hurst 2015). In some circumstances, as described further below, high fidelity can act as such an amelioration strategy (Rajon and Masel 2011). The existence of two distinct lines of defense complicates the naive drift barrier logic that large-N_e species should generally exhibit higher fidelity in all molecular processes. The superior performance of large-N_e species in the primary line of defense may lead to low fidelity in the secondary line of defense. This creates a seemingly counter-intuitive pattern in the secondary line of defense, in which small-N_e species can evolve more faithful processes than large-N_e species.

We are not aware of any definitive observations of this counter-intuitive phenomenon, but large-N_e *E. coli* have similarly high (although not, as predicted, even higher) transcriptional error rates as the small-N_e endosymbiont *Buchnera* (McCandlish and Plotkin 2016; Traverse and Ochman 2016). Unlike *Buchnera*, *E. coli* shows signs of having evolved a first line of defense in the form of a decreased frequency with which observed transcriptional errors translate into non-synonymous changes, relative to randomly sampled transcriptional errors (Traverse and Ochman 2016).

74 The existence of two substantively different lines of defense was first proposed by Krakauer
75 and Plotkin (2002), who contrasted the "redundancy" of robustness to the consequences of
76 mutational errors with the "antiredundancy" of hypersensitivity to mutations. By positing that
77 redundancy had a cost, they showed that the superior cost-free solution of antiredundancy was
78 available only with large N_e, giving small-N_e species higher levels of "redundancy".
79
80 A related argument was made by Rajon and Masel (2011) in the context of mitigating the harms
81 threatened by errors in molecular processes such as translation. Rajon and Masel (2011)
82 distinguished between "local" solutions, where a separate solution is required at each locus,
83 and "global" solutions that can deal with problems at many loci simultaneously. The evolution
84 of extensive proofreading mechanisms was deemed a global solution because it can
85 simultaneously prevent gene expression errors at many loci. Global fidelity via proofreading
86 should come, however, with a cost in time or energy. The alternative, local solution is to have a
87 benign rather than a strongly deleterious "cryptic genetic sequence" at each locus at which
88 expression errors might occur, making the consequence of an error at that locus relatively
89 harmless. In contrast to the global solution, these local solutions bear no direct fitness cost, but
90 because selection at any one locus is weak, mutations at any one locus pass more easily
91 through the drift barrier, making them more difficult to maintain than global solutions.
92
93 Both the proofreading of Rajon and Masel (2011) and the "redundancy" of Krakauer and Plotkin
94 (2002) to the consequences of mutations are global across loci, and also costly. Meantime, both
95 the "local" solutions of Rajon and Masel (2011) and the "antiredundancy" of Krakauer and

96 Plotkin (2002) carry no true fitness cost but instead require a large-N_e drift barrier and/or face a
97 "cost of selection" (Haldane 1957) as limits to their adaptation. A mutation disrupting a solution
98 specific to a single locus requires a large value of N_e for its purging, whereas a mutation
99 disrupting a global proofreading solution will have large fitness consequences and so be easier
100 to purge. The higher-fitness solution is the local one, but it is evolutionarily achievable only with
101 large N_e. With small N_e, we instead expect global solutions such as costly high-fidelity
102 proofreading.
103
104 Selection to purge deleterious cryptic sequences in favor of benign cryptic genetic sequences,
105 i.e. achieving the local solution, may be difficult and hence restricted to high-N_e populations,
106 but there are reasons to believe that it is not impossible. For example, when the error in
107 question is reading through a stop codon, the local cryptic genetic sequence is the 3'UTR, which
108 will now be read by the ribosome. One option for a more benign form of this cryptic sequence
109 is the presence of a "backup" stop codon that provides the ribosome with a second and
110 relatively early chance to terminate translation. Such backup stops are common at the first
111 position past the stop in prokaryotes (Nichols 1970). In *Saccharomyces cerevisiae*, there is also
112 an abundance of stop codons at the third codon position past the stop (Williams *et al.* 2004).
113 Moreover, conservation at this position depends strongly on whether or not the codon is a
114 stop, and the overrepresentation of stops at this position is greater in more highly expressed
115 genes (Liang *et al.* 2005). In some ciliates, where the genetic code has been reassigned so that
116 UAA and UAG correspond to glutamine, this overrepresentation is much more pronounced
117 (Adachi and Cavalcanti 2009). As with the consequences of erroneous readthrough, selective

118 pressure on erroneous amino acid misincorporation and/or misfolding (Drummond and Wilke
119 2008), and on erroneous protein-protein interactions (Brettner and Masel 2012) are also strong
120 enough to shape protein expression and interaction patterns.
121
122 Rajon and Masel (2011) also found that for intermediate values of N_e that correspond strikingly
123 well to many multicellular species of interest, the evolutionary dynamics of the system were
124 bistable, with either the global or the local solution possible. This is a natural consequence of a
125 positive feedback loop; in the presence of a global proofreading solution, specialized solutions
126 at particular loci are unnecessary and mutations destroying them pass through the drift barrier
127 (we use the expression "pass through the drift barrier" to mean that $0 > s > -1/N$), with their
128 subsequent absence increasing the demand for proofreading. Similarly, when specialized
129 solutions predominate, the advantage to proofreading is lessened, and resulting higher error
130 rates further increase selection for many locally specialized solutions. If true, this bistability
131 suggests that historical contingency, rather than current values of N_e, determine which
132 processes are error-prone vs. high-fidelity.
133
134 In the current work, we note that the model of Rajon and Masel (2011) contained an unrealistic
135 symmetry, namely that the fitness consequence of a molecular error at one locus was exactly
136 equal to that at any other loci. Here we find that with reasonable amounts of variation among
137 loci (e.g. in their expression level or the per-molecule damage from their misfolded form), the
138 bistability disappears. Intermediate solutions evolve instead, where cryptic deleterious
139 sequences are purged only in more highly expressed genes, and proofreading evolves to

intermediate levels. The findings of Traverse and Ochman (2016) can be explained by our model if the drift barrier in *Buchnera* allows not only deleterious mutations to local cryptic genetic sequences to pass through, but also increases in the global transcriptional error rate. Throughout our work, evolvability continues to track the proportion of loci that contain a benign rather than a deleterious cryptic sequence.

METHODS

In the following sections, we describe the computational model used to simulate the evolution of different solutions to errors in gene expression. All simulations were run with Matlab (R2014a). Scripts of the simulations are available at https://github.com/MaselLab/Xiong-et-al-Error-rate-evolution-with-variation-in-gene-expression.

Fitness

We follow the additive model of Rajon and Masel (2011), as outlined below, and with a few important modifications to accommodate variation in gene expression levels. The model's canonical example is the risk that a ribosome reads through a stop codon during translation. The global mitigation strategy is to increase proofreading of this gene expression subprocess. Geometric reductions in error rate come at a direct additive cost in the time or energy involved in the gene expression subprocess, so that the fitness component associated with reducing the readthrough error rate to ρ is given by

$$w_{proofread\_cost} = \frac{1}{1-\delta \ln \rho} \quad (1)$$

The parameter δ indicates the relative burden of checking stop codons relative to other molecular processes. Following Rajon and Masel (2011), we set $\delta = 10^{-2.5}$, such that reducing ρ from 10^{-2} to 10^{-3} corresponds to a 0.7% reduction in fitness.

When readthrough happens, with frequency ρ, the consequences for fitness depend on the nature of the "cryptic sequence" that lies beyond the stop codon in the 3'UTR. A striking finding from biology is that the consequences of mistakes, mutational or otherwise, have a bimodal distribution, being either strongly deleterious (often lethal), or relatively benign, but rarely in between (Fudala and Korona 2009). For example, a strongly deleterious variant of a protein might misfold in a dangerous manner, while a benign variant might fold correctly, although with reduced activity. We assume that alternative alleles of cryptic genetic sequences can be categorized according to a benign/deleterious dichotomy.

The local mitigation strategy, the alternative to global proofreading, is thus for each cryptic sequence to evolve away from "deleterious" options and toward "benign" options. The local strategy of benign cryptic sequences has no direct fitness cost, but it is nevertheless difficult to evolve at so many loci at once. In contrast, expressing deleterious cryptic sequences has an appreciable cost. This cost scales both with the base rate of expression of the gene, and the proportion ρ of gene products that include the cryptic sequence.

Let the expression of gene *i* be E_i. We assign the concentration E_i of protein molecules of type *i* by sampling values of E_i from a $\log_2$-normal distribution with standard deviation σ_E. We define *D* to be the total frequency of protein expression that would be highly deleterious if expressed in error:

$$D = \frac{\sum_{i\_deleterious} E_i}{\sum_i E_i} \quad (2)$$

where the numerator sums only over loci that are deleterious and the denominator sums over all loci. This normalization makes the mean value of E_i irrelevant. We assume the costs of deleterious readthrough are additive across genes, based on the concept that misfolded proteins (Thomas *et al.* 1995) may aggregate in a non-specific and harmful manner with other proteins and/or membranes (Kourie and Henry 2002), or may simply be expensive to dispose of (Goldberg 2003). After the stop codon is read through, translation will usually end at a backup stop codon within the 3'UTR. Under the assumption of additivity, single readthrough events will reduce fitness by $c\rho D$, where *c* represents the strength of selection against misfolded proteins. Geiler-Samerotte et al. (2011) found that an increase in misfolded proteins of approximately 0.1% of total cellular protein molecules per cell imposed a cost of about 2% to relative growth rate. This gives an estimate of $c = 0.02/0.1\% = 20$.

Readthrough involving benign cryptic sequences does not incur this cost. However, we also consider double readthrough events, where the backup stop codon is also skipped. This ρ^2 term could also be interpreted in other ways, involving any other weak or unlikely event that causes

even benign cryptic sequences to incur some (smaller) cost. The primary purpose of the double readthrough cost term is to set a bound such that even when all cryptic sequences produced by single readthrough are benign, there is no danger that values of $\rho > 0.5$ will evolve, inadvertently switching the identities of "normal" expression with "erroneous" expression. We assume the second cryptic sequences are at neutral mutational equilibrium, with probability of being deleterious equal to $P_{del}/(P_{del}+P_{ben})$, where P_{del} is the rate of deleterious-to-benign mutations to cryptic sequences and P_{ben} the rate of the reverse mutations. Combining the fitness costs of single and double readthrough, the fitness component representing the cost of aberrant expression is given by

$$w_{misfolding} = \max\left(0, 1 - c\rho D - c\rho^2(1-D)\frac{P_{del}}{P_{del}+P_{ben}}\right) \quad (3)$$

The equation for $w_{misfolding}$ is the primary modification that we make to the additive model of Rajon and Masel (2011), allowing us to study variation in the degree of importance of cryptic loci, subsuming a model similar to the previous one as a special case of zero variation. Where previous work referred to the number L_{del} of loci having the deleterious rather than benign form, we now distinguish between two measures, L_{del} and D.

To study evolvability, let a subset of K (typically 50) out of the L (typically 600 or more) loci affect a quantitative trait x, selection on which creates a third fitness component. Error-free expression of locus k, occurring with frequency 1-ρ, has quantitative effect α_k, while expression that involves a benign version of the cryptic sequence has quantitative effect $\alpha_k + \beta_k$.

Expression that involves a deleterious version of the cryptic sequence is assumed to result in a misfolded protein that has no effect on the quantitative trait. We assume that expression level E_k is constant and already factored into values of α_k and β_k. This gives

$$x = \sum_k^K ((1-\rho)\alpha_k + \rho B_k(\alpha_k + \beta_k)) \quad (4)$$

where B = 1 indicates a benign cryptic sequence and B = 0 a deleterious one. Following Rajon and Masel (2011), we impose Gaussian selection on x relative to an optimal value x_{opt}

$$w_{trait}(x) = e^{\frac{-(x-x_{opt})^2}{2\sigma_f^2}} \quad (5)$$

where σ_f = 0.5.

Putting the three fitness components together, the relative fitness of a genotype is given by the product

$$w = w_{proofread\_cost} \times w_{misfolding} \times w_{trait}. \quad (6)$$

Variance in expression levels

We estimated the variance in expression σ_E from PaxDB (Wang *et al.* 2012; Wang *et al.* 2015), which is based on data released by the Global Proteome Machines (GMP) and other sources. We inferred σ_E equal to 2.24 (based on GMP 2012 release) or 3.31 (GMP 2014 release), for *S.*

249 *cerevisiae*, and 2.93 (GMP 2014 release) for *S. pombe*. Note that while our quantitative
250 estimate of σ_E comes from variation in the expression levels of different proteins, consideration
251 of variation along other lines might make a standard deviation of 2.25 into a conservative
252 underestimate of the extent of variation. See Fig. S2 for an exploration of this parameter value.
253
254 **Mutation**
255 There are six kinds of mutation: 1) conversion of a deleterious cryptic sequence to a benign
256 form, 2) conversion from benign to deleterious, 3) change to the error rate ρ, 4) change in the α
257 value of one of the *K* quantitative trait genes, 5) change in the β value of one of those *K* genes,
258 and 6) the co-option of a cryptic sequence to become constitutive, replacing the value of
259 replacing α_k with that of α_k+β_k and re-initializing B_k and β_k.
260
261 It is this sixth kind of mutation that is responsible for the evolvability advantage of the local
262 solution of benign cryptic sequences, providing more mutational raw material by which *x* might
263 approach x_{opt} (Rajon and Masel 2011; Rajon and Masel 2013). The mutational co-option of a
264 deleterious sequence (B = 0) is too strongly deleterious to be favored, even when replacing α_k
265 and β_k might be advantageous. Only benign cryptic sequences are available for mutational co-
266 option. We use the term co-option of a 3'UTR readthrough sequence to refer to the case when
267 a stop codon is lost by mutation, and not just read through by the ribosome (Giacomelli *et al.*
268 2007; Vakhrusheva *et al.* 2011; Andreatta *et al.* 2015). Mutational co-option for mimicking the
269 consequences of errors other than stop codon readthrough might involve mutations that

change expression timing to make a rare protein-protein interaction common, or switch a protein's affinity preference between two alternative partners.

Because we use an origin-fixation approach to simulate evolution (see below), only relative and not absolute mutation rates matter for our outcomes, with the absolute rates setting only the timescale. We use the same mutations rates as Rajon and Masel (2011), reduced ten-fold for convenience. Each locus with a benign cryptic sequence mutates to deleterious with probability $P_{del} = 2.4\times10^{-8}$ per generation, while deleterious loci mutate to benign less often, with probability $P_{ben} = 6\times10^{-9}$.The error rate ρ changes with probability 10^{-6} per generation, while the α and β values of each quantitative locus each change with probabilities 3×10^{-7} and 3×10^{-8}, respectively. Mutational co-option occurs with probability 2.56×10^{-9} per quantitative trait locus per generation.

Each mutation to ρ increases $\log_{10}\rho$ by an amount sampled from Normal(ρ_{bias}, 0.2^2). By default, we set $\rho_{bias} = 0$. To study extremely small populations with drift barriers to evolving even a global solution, we set $\rho_{bias} = 0.256$ and 0.465, corresponding to ratios of ρ-increasing mutations: ρ-decreasing mutations of 9:1 and 99:1, respectively.

A similar scheme for α and β might create, in the global solution case of relaxed selection, a probability distribution of β whose variance increases in an unbounded manner over time (Lande 1975; Lynch and Gabriel 1983). Following previous work (Rajon and Masel 2011; Rajon and Masel 2013), we therefore let mutations alter α and β by an increment drawn from a

normal distribution with mean $-\alpha/a$ or $-\beta/a$, with a set to 750, and with standard deviation of σ_m/K in both cases, with σ_m set to 0.5. In the case of neutrality, this mutational process eventually reaches a stationary distribution with mean 0 and standard deviation as calculated in Eq. S3 of Rajon and Masel (2011):

$$V(a, K, \sigma_m) = \frac{(\sigma_m/K)^2}{1-((a-1)/a)^2} \quad (7)$$

A co-option at gene k changes the gene's quantitative effect to

$$(1-\rho)(\alpha_k + \beta_k) + \rho B_k'(\alpha_k + \beta_k + \beta_k') \quad (8)$$

where B_k' and β_k' are the state and the quantitative effect of a new cryptic sequence created by co-option. Following a co-option mutation at locus k, we set the new B_k equal to 1 or 0 with probabilities proportional to P_{ben} and P_{del}, and resample the value of β_k from Normal(0, $V(a, K, \sigma_m)$).

Evolutionary simulations by origin-fixation

We model evolution using an approach known as "strong-selection-weak-mutation" (Gillespie 1983), or "origin-fixation" (McCandlish and Stoltzfus 2014). This approximation of population genetics is accurate in the limit where the waiting time until the appearance of the next mutation destined to fix is substantially longer than its subsequent fixation time. The population can then be approximated as genetically homogeneous in any moment in time.

While unrealistic for higher mutation rates and larger population sizes, origin-fixation models are computationally convenient. Still more importantly, origin-fixation models, unlike more realistic models with segregating variation, allow the location of the drift barrier to be set externally in the form of the value of the parameter *N*, rather than having the location of the drift barrier emerge from complicated linkage phenomena within the model. Fortunately, for quantitative traits affected by multiple cryptic loci, most evolvability arises from diversity of the effects of co-option of different loci, rather than among the diversity of the effects of co-option from different starting genotypes (Rajon and Masel 2013). This allows us to study evolvability (in the population sense of Wagner (2008)) even in the absence of genetic diversity that is imposed by the origin-fixation formulation.

Origin-fixation models are often implemented via a crude rejection algorithm; large numbers of mutations are simulated, and each is accepted as a successful fixation event if and only if a random number sample from the uniform [0, 1] distribution falls below its (fairly low) fixation probability. For large *N*, this method is computationally slow when significant numbers of nearly neutral mutations must be sampled before one fixes with probability ~1/*N*. Given that our model posits only a relatively small range of possible mutations, we instead sampled only mutations that go on to become fixed, by sampling according to the relative values of "fixation flux", proportional to mutation rate $\times$ fixation probability for each of our six categories of mutation. In other words, we used a form of the Gillespie (1977) algorithm.

335 In a haploid population of size N, the probability of fixation of a new mutant into a resident
336 population is given by
337
338 $$p_{fix} = \frac{1-e^{-s}}{1-e^{-Ns}} \quad (9)$$
339
340 where $s = w_{mutant} / w_{resident} - 1$. It is then straightforward to calculate fixation flux values for all
341 possible switches between benign and deleterious states. Matters are slightly more
342 complicated for quantitative mutations to ρ, α, and β, because we must integrate the fixation
343 flux over all possible sizes for a mutation, prior to summing the fixation flux per mutation to
344 arrive at the fixation flux for an entire mutational category.
345
346 We use the quadrature method to calculate the integral over these possibilities, using a grid of
347 2000, limited for $\Delta\alpha_k$ to the interval $[-\alpha_k/a-5\sigma_m/K, -\alpha_k/a+5\sigma_m/K]$, for $\Delta\beta_k$ to the interval $[-\beta_k/a-$
348 $5\sigma_m/K, -\beta_k/a+5\sigma_m/K]$, and for $\Delta\log_{10}\rho$, to the interval $[-2, \min(2, -\log_{10}\rho)]$. In the latter case, the
349 number of grid intervals is reduced proportional to any truncation of the interval at $-\log_{10}\rho$.
350
351 For mutational co-options of benign cryptic sequences, the effect of replacing the value of α_k
352 with that of $\alpha_k+\beta_k$ is fixed, but there is also a stochastic range of effects of initializing a new β_k
353 and a new B_k. We integrate Prob(new β_k) $\times$ Prob(fixation$|\beta_k$) over new values of β_k in the range
354 $[-5\sigma_m/K, 5\sigma_m/K]$, and multiply by the probability that the new B=1. To this we add Prob(B=0) $\times$
355 Prob(fixation$|B$=0), a case which does not require an integration step.
356

The expected waiting time before the current genotype is replaced by another is

$$\text{waiting time} = \frac{1}{total\ fixation\ flux\ over\ all\ six\ categories} \quad (10)$$

A standard Gillespie (1977) algorithm would calculate the realized waiting time as the inverse of a random number drawn from an exponential distribution with this mean. Since we are only interested in the outcome of evolution, and not the variation in its timecourse, we used the expected waiting time instead, decreasing our computation time. The waiting time can be interpreted as the time it takes for a mutation destined for fixation to appear, neglecting the time taken during the process of fixation itself. Using this interpretation, we specify waiting times in terms of numbers of generations, based on our assumptions about absolute mutation rates.

We assign the identity of the next fixation event among the six categories according to probabilities proportional to their relative fixation fluxes, then we assign the identity within the category. For switches between benign and deleterious states, allocating a fixation event within a category according to the relative values of fixation fluxes is straightforward. For mutations to ρ, α, and β, and mutational co-option, we relax the granularity and cutoff assumptions of the grid-integration method when choosing a mutation within the category. Instead, we sample a mutational value of $\Delta\log_{10}\rho$ from Normal(ρ_{bias}, 0.2^2). We reject and resample $\Delta\log_{10}\rho$ if $\Delta\log_{10}\rho \geq -\log_{10}\rho$. Otherwise, we accept vs. reject-resample according to the fixation probability of that exact mutation, by comparing this probability to a random number uniformly

379 distributed at [0, 1.1×the maximum fixation probability across the grid points previously
380 calculated for $\Delta\log_{10}\rho$ during our grid calculation]. For $\Delta\alpha$ (or $\Delta\beta$), the procedure is conceptually
381 similar but has a more complicated implementation. We first sample from Normal(0, $(\sigma_m/K)^2$).
382 We then add the random number to each of the values of $-\alpha_k/a$, and calculate the sum of
383 corresponding fixation probabilities across all loci k. We accept vs. reject-resample the
384 mutation by comparing this sum to a random sample from a uniform distribution at [0, 1.1×the
385 maximum corresponding fixation probability sum calculated during our grid calculation]. If the
386 mutation is accepted, we allocate it to a locus k with probability proportional to their relative
387 fixation probabilities. For mutational co-option of a benign cryptic sequence, the main effect is
388 to replace α_k with $\alpha_k+\beta_k$, but there are also more subtle effects arising from the reinitialization
389 of the new cryptic sequence. Any of the k loci for which $B = 1$ are eligible for co-option, the new
390 value of B may be either 0 to 1, and the new β_k may take a range of values. Each combination of
391 k and new B has its own fitness flux, and the first choice is among these {k, B} pairs. Next we
392 sample β_k from Normal(0, $(\sigma_m/K)^2$); for a new B equal to 0 we always accept the result, and for
393 new B equal to 1, we accept vs. reject-resample β_k by comparing its probability of fixation to a
394 random sample from a uniform distribution at [0, 1.1×the maximum corresponding fixation
395 probability sum calculated during our grid calculation].
396
397 **Initialization and convergence**
398 We initialized the trait optimum at $x_{opt} = 0$. We could have initialized all values of α_k and β_k at
399 zero. However, at steady state, variation in $\sum^K \alpha_k$ and $\sum^K \beta_k$ is far lower than would be
400 expected from variation in α_k and β_k – this emerges through a process of compensatory

evolution (Rajon and Masel 2013). Allowing a realistic steady state to emerge in this way is computationally slow under origin-fixation dynamics, especially when N is large. We instead sampled the initial values of α_k and β_k from Normal(0, $V(a, K, \sigma_m)$), where $V(a, K, \sigma_m)$ is defined by Eq. 7, and then subtracted $\bar{\alpha}$ from α_k and $\bar{\beta}$ from β_k, where $\bar{\alpha}$ and $\bar{\beta}$ are the means of a genotype across each of its quantitative loci k. This process initializes α_k and β_k to have variances equal to those of the stationary distributions, while the overall trait value is initialized at the optimal value, zero. This procedure greatly reduces the computation time needed to burn-in to achieve a somewhat subtle state of negative within-genotype among-loci correlations. We confirmed that subsequent convergence of the variance of $\sum^K \alpha_k$ was fast, occurring in less than 1000 steps, where a "step" is defined to be the fixation of one mutation. We expect $\log_{10}\rho$, D, and variation in β_k, to converge even faster than variation in α_k.

For the low-ρ initial conditions, ρ was initialized at 10^{-5}, and we initialized the benign vs. deleterious status of cryptic sequences at the neutral mutational equilibrium, choosing exactly $L \times P_{del}/(P_{del}+P_{ben})$ (rounded to the nearest integer) to be deleterious, independently of their different values of E. For the high-ρ initial conditions, we set ρ to 10^{-1}, and made all cryptic sequences benign.

We ran simulations for 10^5 steps, recording information at fixed times (measured in terms of waiting times), corresponding to approximately every 1000 steps on average, and hence yielding about 100 timepoints. To summarize the evolutionary outcome, we calculated the

arithmetic means of $\log_{10}p$, of L_{del}, and of D among the last 20 timepoints, i.e. approximating steps 0.8×10^5-1×10^5.

Evolvability

After adaptation to a trait optimum of $x_{opt} = 0$ had run to convergence (i.e. after 10^5 steps), we changed x_{opt} to 2, forcing the quantitative trait to evolve rapidly. This allows the co-option of benign cryptic sequences an opportunity to increase evolvability. We measured evolvability in two ways: as the inverse of the waiting time before trait x exceeded 1, and the inverse of the waiting time before the population recovered half of the fitness it lost after x_{opt} changed.

We want our measures of evolvability to reflect a genotype's potential to generate beneficial mutations, but this goal was complicated by population size. A large population finds a given mutation faster than a small population does, inflating the total fixation flux in direct proportion to population size. We therefore divided our evolvability measures by the population size to correct for this obvious effect. This normalization converts the population-level evolvability measure into a measure of the population-size-independent evolvability of an individual that has the genotype of interest.

Transcriptional errors rates as a function of gene expression

We used *E. coli* transcriptional error data from Traverse and Ochman (2016), using the authors' mapping to the *E. coli* genome and filtering to exclude low quality reads. We also excluded reads of structural RNA genes, *ompF*, *ompC*, *tufA*, and *tufB*. We pooled data acquired under

different media, growth phase, and replicates, and fit the pooled data to models with a different error rate, ρ_i, for each of the 12 different mistranscription types (i.e. U->T, A->G etc.). Using the glm function in R, we fit the null model

$$N_{site\_specific\_error} = \sum_i^{12} \rho_i N_{site\_specific\_reads} + \varepsilon \quad (11)$$

where $N_{site\_specific\_error}$ is the number of reads of non-consensus nucleic acids at a specific site of a transcript and $N_{site\_specific\_reads}$ is the total number of reads of that site. There is no intercept in the model and ε is Poisson-distributed. To detect whether transcript levels affect error rates, we compared this to the model

$$N_{site\_specific\_error} = \sum_i^{12} \rho_i N_{site\_specific\_reads} + C\sqrt{N_{site\_specific\_reads}} + \varepsilon \quad (12)$$

where C is a factor acting uniformly on each mistranscription type. The square root of $N_{site\_specific\_reads}$ gave a better fit than other functional forms that we tried. The best fit to Eq. 12 had $C = 1.2\times10^{-5}$, with the arithmetic mean of the 12 fitted ρ_i values being 2.4×10^{-5}. The two models are nested, and so were compared using a likelihood ratio test.

RESULTS

Recall that in the absence of variation in expression among genes, there are two solutions to handle erroneous expression due to stop codon readthrough: at high population size N, the local solution purges all deleterious cryptic sequences, making high rates of readthrough

466 harmless, while at low *N*, the global solution reduces the rate of readthrough, allowing
467 deleterious cryptic sequences to accumulate near-neutrally. At intermediate *N*, we see
468 bistability, with either solution possible, depending on starting conditions (Fig. 1, $\sigma_E = 0$). It is
469 important to note that we use the word "bistability" loosely. Strictly speaking, bistability means
470 that the system has two stable steady states (here a state is defined by readthrough rate and
471 the exact property of each cryptic sequence), i.e. two attractors. But in a stochastic model,
472 there are no attractors in the strict sense of the word, only a stationary distribution of states.
473 We use the term bistability to refer to the case where the stationary distributions of states has
474 two modes. Transitions between the two modes are rare, therefore the two modes can be
475 loosely interpreted as the two attractors of the system.
476
477 Our results qualitatively reproduce the bistability reported by Rajon and Masel (2011) for the
478 case where there is no expression variation among genes, though the range of values of *N*
479 leading to bistability is smaller than that found in Rajon and Masel (2011) in which a full Wright-
480 Fisher simulation is used. The smaller range of bistability in our model could be caused by the
481 ease with which long-term evolution is captured using an origin-fixation framework, or by other
482 subtle differences between the approaches, e.g. the greater ease of compensatory evolution
483 under Wright-Fisher dynamics than under origin-fixation. We chose origin-fixation mainly to
484 reduce the computational burden, which for our study was increased by the need to track
485 individual loci, in contrast to previous work that needed only to track the number of loci with
486 deleterious cryptic sequence, without distinguishing their identities (Rajon and Masel 2011;
487 Rajon and Masel 2013).

However, bistability vanishes with variation in expression among genes (Fig. 1, σ_E = 2.25 and σ_E = 3.5). To understand why, consider a population initialized at low readthrough rate (ρ) and many deleterious cryptic sequences. Because the strength of selection against a deleterious cryptic sequence at locus i is proportional to ρE_i (the effect of a locus i on D in Eq. 3 is proportional to E_i), purging works at the most highly expressed loci, even when ρ is low. This lowers the proportion D of readthrough events that are deleterious, which relaxes selection for high fidelity, leading to an increase in ρ. As ρ increases, loci with lower E_i become subject to effective purging, which further reduces D, which feeds back to increase ρ further. Because E_i is log-normally distributed, but contributes linearly to selection via D, each round of the feedback loop involves smaller changes than the last. Eventually, the changes are too small for selection on them to overcome mutation bias in favor of deleterious sequences. Similarly, when a population is initialized at high ρ, mutational degradation begins at low E_i sites and arrests when selection is strong enough to kick in. The point of balance between mutation bias and selection defines a single intermediate attractor for $\sigma_E \geq 2.25$, instead of the bistable pair of attractors found for uniform E_i (σ_E = 0). For $\sigma_E < 2.25$, bistability is still found, but for a narrower range of population sizes than in the absence of variation (Fig. S2).

Even though bistability is not found for σ_E = 2.25, there is still a fairly sharp dichotomy, with solutions being either local (high ρ and low L_{del}) or global (low ρ and high L_{del}), and intermediate solutions found only for a very restrictive range of N, following a sigmoidal curve (Fig. **1a** and **1c**). Increasing variation in expression among genes blurs the boundary between the local

510 solution and the global solution. Intermediate solutions are found for broader ranges of N as
511 expression variance σ_E increases to 3.5. The trend, as expression variance σ_E increases from 0, is
512 to first replace bistability with a limited range of intermediate solutions (σ_E = 2.25), and then for
513 the intermediate solutions to become more prevalent, with extreme local and global solutions
514 becoming less attainable as σ_E > 2.25.
515
516 The breakdown of the local solution begins with intermediate values of L_{del}, while the
517 breakdown of the global solution begins with intermediate values of ρ and D (Fig. 1 **a**-**c**). The
518 breakdown of global solutions involves high-expression loci (Fig. 2), which affect D more than
519 L_{del}. In contrast, the breakdown of local solutions involves low-expression loci (Fig. 2), which
520 affect L_{del} more than D. Because ρ is better described as co-evolving with D than with L_{del}, as
521 explained earlier, intermediate values of ρ are seen more in the breakdown of global than local
522 solutions.
523
524 The local solution promotes evolvability by making benign cryptic sequences available for co-
525 option. Differences in evolvability between genotypes should therefore be largely determined
526 by the fraction of quantitative trait loci that carry benign rather than deleterious cryptic
527 sequences. In agreement with this, evolvability inversely mirrors L_{del}, as a function of population
528 size, i.e., evolvability (Fig. 1**d**) resembles L_{del} (Fig. 1**c**) far more than it resembles ρ (Fig. 1**a**) or D
529 (Fig. 1**b**).
530

Reducing the total number of loci L tips the balance toward local solutions, because a higher average expression level across loci helps stop deleterious mutations from passing through the drift barrier even in smaller populations, but this does not qualitatively change our results (Fig. S4). The distinction between global and location solutions becomes more extreme when the mutation bias toward deleterious rather than benign cryptic sequences is increased from 4:1 ratio to a 99:1 ratio, but persists even when the mutation bias is eliminated in favor of a 1:1 ratio (Fig. 3). In the absence of mutation bias, there is less evolvability to be gained by the local relative to the global solution, since half the quantitative loci are available for co-option regardless (Fig. 3**c**). Nevertheless, a small evolvability advantage to the local solution can still be observed (Fig. 3**d**). In any case, assuming mutation bias toward deleterious options is biologically reasonable, and Fig. 3 shows that results are not sensitive to the strength of our assumptions on this count.

When we also account for mutation bias that tends to increase rather than decrease the error rate ρ, our model can explain the previously puzzling observation that the rate of transcriptional errors in small-N_e endosymbiont bacteria *Buchnera* is as high as that in large-N_e *E. coli* (McCandlish and Plotkin 2016; Traverse and Ochman 2016). In extremely small populations, even the global solution is subject to a drift barrier making ρ higher than its optimal value. For N so small such that most ρ-increasing mutations pass through the drift barrier, ρ can be as large as that in large populations (Fig. 4**a**). Despite their high error rates, these extremely small populations also carry heavy loads of deleterious cryptic products (Fig. 4**b** and **c**), consistent with the fact that in *Buchnera*, unlike *E. coli*, selection is unable to reduce the fraction of non-

553 synonymous transcriptional errors that are non-synonymous (Traverse and Ochman 2016). High
554 ρ shows the absence of a global solution, while high D and L_{del} show the absence of a local
555 solution; neither solution is found for a sufficiently small population. Equal error rates in large
556 and small populations can also be found, given bias in mutations to ρ, when there is no
557 variation in expression levels (Fig. S5).
558
559 DISCUSSION
560 When genes vary in their expression levels, the dichotomy between the local and global
561 solutions is replaced by a continuous transition. Very large populations still resemble the local
562 solution, although mutations making cryptic sequences deleterious may still pass through the
563 drift barrier in the occasional low-expression gene. Very small populations still resemble the
564 global solution, although mutations making cryptic sequences deleterious may still be
565 effectively purged in a few high-expression genes; because their high expression
566 disproportionately affects the burden from misexpression, this relaxes expression for high
567 fidelity, leading to less extreme proofreading.
568
569 In agreement with drift barrier theory, large-N_e *E. coli* exhibits a local solution of a tendency for
570 transcription errors to have synonymous effects, while small-N_e *Buchnera* does not (Traverse
571 and Ochman 2016). Interestingly, while as predicted, the global solution of low transcriptional
572 error rates does not obey the naïve drift barrier expectation of being weaker in *Buchnera* than
573 in *E. coli* (Traverse and Ochman 2016), nor is it stronger in *Buchnera* as predicted by previous
574 theory on the interplay between global and local solutions (Rajon and Masel 2011; McCandlish

575 and Plotkin 2016). Here we explain this by also taking into account a drift barrier on the global
576 solution of low error rates, a drift barrier seen in the presence of mutation bias towards higher
577 error rates. Small *Buchnera* populations have high error rates because they can't manage
578 better; large *E. coli* populations have equally high error rates because with the consequences of
579 error already purged, they don't need to incur the cost that proofreading entails.
580
581 With small amounts of variation in expression among genes, the range of intermediate values
582 of N_e for which bistability is found shrinks. With more variation, bistability vanishes in favor of a
583 sigmoidal transition between global and local solutions. With still more, the sigmoid is
584 smoothed out, and intermediate solutions are found for most values of N_e. To interpret our
585 results correctly, we must therefore estimate the degree to which genes vary. The results
586 presented here focus on two estimates of the variation in log-expression in yeast, namely
587 standard deviations of 2.25 and 3.5. However, other variation in the consequences of genes'
588 erroneous expression, in addition to variation in expression level, might make larger standard
589 deviations a better model of reality, further supporting a continuum of intermediate solutions.
590
591 Our model makes three critical assumptions, which must be understood for the results to be
592 interpreted appropriately. First, a "locus" in our model consists of one regular and one cryptic
593 sequence. The primary example that we used to parameterize the results posits an entire
594 protein-coding gene as the regular sequence, and the extended polypeptide resulting from stop
595 codon readthrough as the cryptic alternative. In the example of transcriptional errors, a locus is
596 a single codon, with its corresponding amino acid being the regular sequence and the most

common consequence of a transcriptional error, given their distribution, as the cryptic. The case of one regular sequence and many alternative cryptic ones has not been modeled. Similarly, proteins may each have a regular fold or binding partner, and our model considers the contrast between this and a single cryptic alternative.

Second, we assume that the rate of gene expression errors is set globally, across all loci. In reality, individual context may also affect the error rate, giving error rates a local solution aspect as well. A model of three rather than two interacting solutions – global error rates, local error rates, and local robustness to the consequences of error – remains for future work. Perhaps highly expressed genes will have both more benign cryptic sequences and lower rates of error, or perhaps the evolution of one kind of local solution will alleviate the need for another. Testing this empirically requires data on site-specific error rates and on a credible marker for the benign status of members of an identifiable class of cryptic sequences. Such tools are now becoming available, and indeed we recently found a positive correlation between a large number of readthrough errors at a particular stop codon and the benign status of the readthrough translation product (Kosinski et al., manuscript in preparation). We also reanalyzed the data of Traverse and Ochman (2016) to find that highly expressed transcripts have lower transcriptional error rates ($p = 3.95 \times 10^{-21}$, see Methods). This corresponds to an effect size in which the error rate is reduced by 33% at the most highly transcribed site, compared to a site transcribed at a 5000-fold lower level and hence detectable on average in only a single read per site.

619 Finally, we assume that the consequences of errors have a bimodal distribution: either highly
620 deleterious or largely benign, but rarely in between. In other words, we assume that a basic
621 phenomenon in biology is that changes tend to either break something, or to tinker with it.
622 There are a variety of lines of evidence supporting this intuitively reasonable assumption
623 (Fudala and Korona 2009; Wylie and Shakhnovich 2011).
624

625 ## ACKNOWLEDGEMENTS

626 This work was supported by the John Templeton Foundation [grant number 39667]. DJP was
627 also funded by the Undergraduate Biology Research Program at the University of Arizona. We
628 thank Lilach Hadany, Yoav Ram, and other members of the Hadany group for helpful
629 discussions that prompted us to explore variation in expression. We thank Paul Nelson for
630 developing the idea of local error rates as an expansion of our model, Ben Wilson for help with
631 R, and Etienne Rajon for helpful comments on the manuscript. We also thank Charles Traverse
632 and Howard Ochman for sharing their data on transcriptional errors. An allocation of computer
633 time from the UA Research Computing High Performance Computing (HPC) and High
634 Throughput Computing (HTC) at the University of Arizona is gratefully acknowledged.
635

LITERATURE CITED

Adachi, M., and A. R. O. Cavalcanti, 2009 Tandem Stop Codons in Ciliates That Reassign Stop Codons. J. Mol. Evol. 68: 424-431.

Andreatta, M. E., J. A. Levine, S. G. Foy, L. D. Guzman, L. J. Kosinski *et al.*, 2015 The Recent De Novo Origin of Protein C-Termini. Genome Biology and Evolution 7: 1686-1701.

Brettner, L. M., and J. Masel, 2012 Protein stickiness, rather than number of functional protein-protein interactions, predicts expression noise and plasticity in yeast. BMC Systems Biology 6: 128.

Drummond, D. A., and C. O. Wilke, 2008 Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution. Cell 134: 341-352.

Frank, S. A., 2007 Maladaptation and the Paradox of Robustness in Evolution. PLoS ONE 2: e1021.

Fudala, A., and R. Korona, 2009 Low frequency of mutations with strongly deleterious but nonlethal fitness effects. Evolution 63: 2164-2171.

Geiler-Samerotte, K. A., M. F. Dion, B. A. Budnik, S. M. Wang, D. L. Hartl *et al.*, 2011 Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. Proc. Natl. Acad. Sci. U.S.A. 108: 680-685.

Giacomelli, M. G., A. S. Hancock and J. Masel, 2007 The conversion of 3' UTRs into coding regions. Mol. Biol. Evol. 24: 457-464.

Gillespie, D. T., 1977 Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81: 2340-2361.

Gillespie, J. H., 1983 Some properties of finite populations experiencing strong selection and weak mutation. Am. Nat. 121: 691-708.

Goldberg, A. L., 2003 Protein degradation and protection against misfolded or damaged proteins. Nature 426: 895-899.

Good, B. H., and M. M. Desai, 2014 Deleterious Passengers in Adapting Populations. Genetics 198: 1183-1208.

Haldane, J. B. S., 1957 The cost of natural selection. J. Genetic. 55: 511-524.

Kourie, J. I., and C. L. Henry, 2002 Ion channel formation and membrane-linked pathologies of misfolded hydrophobic proteins: The role of dangerous unchaperoned molecules. Clin. Exp. Pharmacol. Physiol. 29: 741-753.

Krakauer, D. C., and J. B. Plotkin, 2002 Redundancy, antiredundancy, and the robustness of genomes. Proc. Natl. Acad. Sci. U.S.A. 99: 1405-1409.

Lande, R., 1975 The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. 26: 221-235.

Liang, H., A. R. O. Cavalcanti and L. F. Landweber, 2005 Conservation of tandem stop codons in yeasts. Genome Biol. 6: R31.

Lynch, M., 2007 *The origins of genome architecture*. Sinauer Associates, Sunderland.

Lynch, M., 2010 Evolution of the mutation rate. Trends Genet. 26: 345-352.

Lynch, M., 2012 Evolutionary layering and the limits to cellular perfection. Proc. Natl. Acad. Sci. U.S.A. 109: 18851-18856.

Lynch, M., and W. Gabriel, 1983 Phenotypic evolution and parthenogenesis. Am. Nat. 122: 745-764.

McCandlish, D. M., and J. B. Plotkin, 2016 Transcriptional errors and the drift barrier. Proc. Natl. Acad. Sci. U.S.A. 113 136-3138

McCandlish, D. M., and A. Stoltzfus, 2014 Modeling Evolution Using the Probability of Fixation: History and Implications. Q. Rev. Biol. 89: 225-252.

Nichols, J. L., 1970 Nucleotide Sequence from the Polypeptide Chain Termination Region of the Coat Protein Cistron in Bacteriophage R17 RNA. Nature 225: 147-151.

Rajon, E., and J. Masel, 2011 The evolution of molecular error rates and the consequences for evolvability. Proc. Natl. Acad. Sci. U.S.A. 108: 1082-1087.

Rajon, E., and J. Masel, 2013 Compensatory Evolution and the Origins of Innovations. Genetics 193: 1209-1220.

Thomas, P. J., B.-H. Qu and P. L. Pedersen, 1995 Defective protein folding as a basis of human disease. Trends Biochem. Sci. 20: 456-459.

Traverse, C. C., and H. Ochman, 2016 Conserved rates and patterns of transcription errors across bacterial growth states and lifestyles. Proc. Natl. Acad. Sci. U.S.A. 113: 3311-3316

Vakhrusheva, A., M. Kazanov, A. Mironov and G. Bazykin, 2011 Evolution of Prokaryotic Genes by Shift of Stop Codons. J. Mol. Evol. 72: 138-146.

Wagner, A., 2008 Robustness and evolvability: a paradox resolved. Proc. R. Soc. Lond. Ser. B-Biol. Sci. 275: 91-100.

Wang, M., C. J. Herrmann, M. Simonovic, D. Szklarczyk and C. von Mering, 2015 Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. Proteomics 15: 3163-3168.

Wang, M., M. Weiss, M. Simonovic, G. Haertinger, S. P. Schrimpf *et al.*, 2012 PaxDb, a Database of Protein Abundance Averages Across All Three Domains of Life. Mol. Cell. Proteomics 11: 492-500.

Warnecke, T., and L. D. Hurst, 2011 Error prevention and mitigation as forces in the evolution of genes and genomes. Nat Rev Genet 12: 875-881.

Williams, I., J. Richardson, A. Starkey and I. Stansfield, 2004 Genome-wide prediction of stop codon readthrough during translation in the yeast *Saccharomyces cerevisiae*. Nucleic Acids Res. 32: 6605-6616.

Wu, X., and L. D. Hurst, 2015 Why Selection Might Be Stronger When Populations Are Small: Intron Size and Density Predict within and between-Species Usage of Exonic Splice Associated cis-Motifs. Mol. Biol. Evol. 32: 1847-1861.

Wylie, C. S., and E. I. Shakhnovich, 2011 A biophysical protein folding model accounts for most mutational fitness effects in viruses. Proc. Natl. Acad. Sci. U.S.A. 108: 9916-9921.

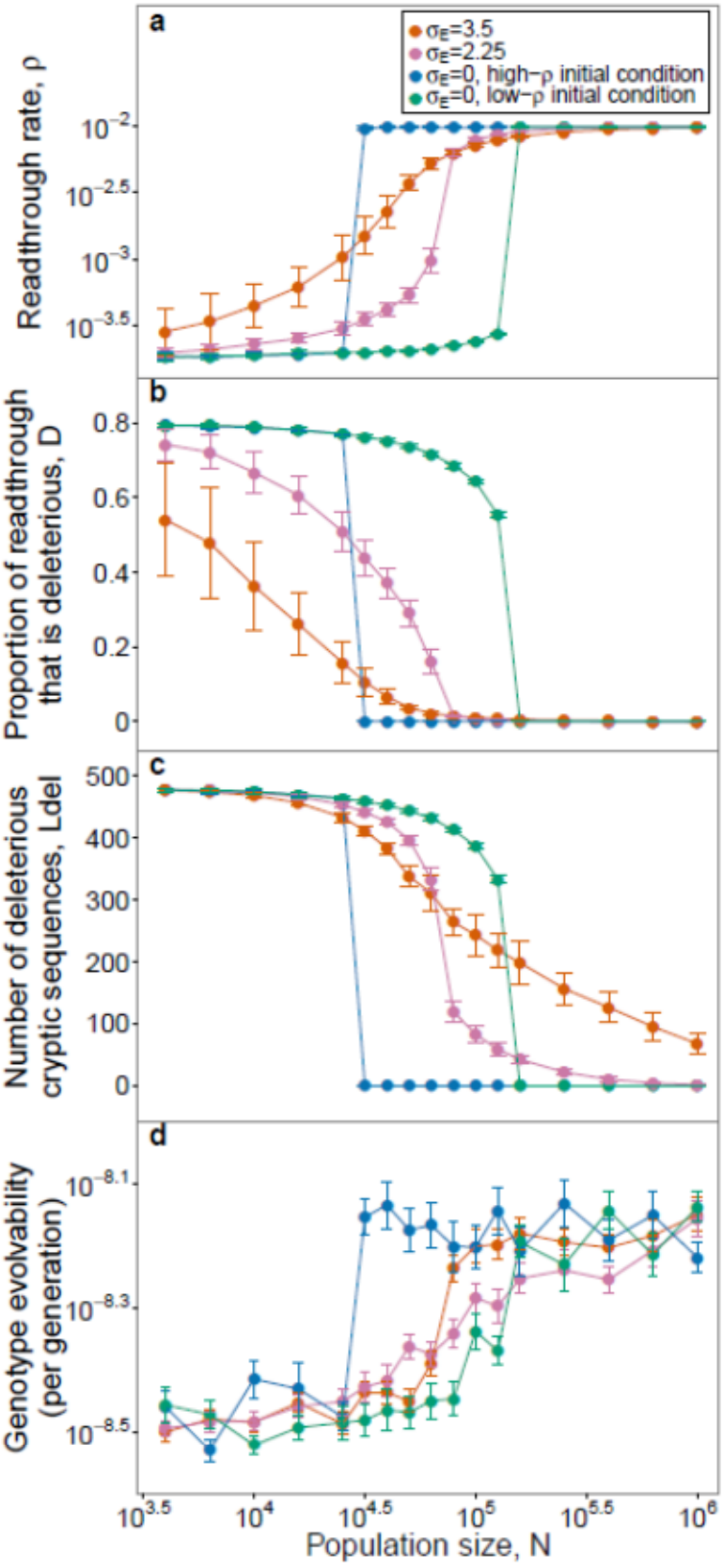

Figure 1: Evolutionary dynamics are bistable in the absence of variation in gene expression ($\sigma_E = 0$), but not with variation in gene expression ($\sigma_E = 2.25$ and $\sigma_E = 3.5$). We calculated the average values of ρ, D, and L_{del} towards the end of the simulations, and then measured the genotype evolvability after changing the optimal trait value (see Methods for details). For each value of N, 20 simulations were initialized at high-ρ conditions and 15 at low-ρ conditions. For $\sigma_E = 2.25$ and $\sigma_E = 3.5$, simulations from the two initial conditions reached indistinguishable endpoints (Fig. S1), so the results were pooled. The increment in N is $10^{0.1}$ between $10^{4.4}$ and $10^{5.2}$ to increase resolution, and is $10^{0.2}$ elsewhere. At $\sigma_E = 0$, D is indistinguishable from zero for $N \geq 10^{5.2}$ under high-ρ conditions and for $N \geq 10^{4.7}$ under low-ρ conditions, corresponding to L_{del} being effectively zero. In contrast, when $\sigma_E = 2.25$ or 3.5, because the weakness of selection on low-expression genes prevents L_{del} from falling all the way to zero, D never quite reaches zero either, despite appearing superimposable in **b**. For **a** to **c**, data is shown as mean$\pm$SD. For evolvability (**d**), data is shown as mean$\pm$SE. Evolvability is based on time to fitness recovery; see Fig. S3 for similar results based on time to trait recovery. $L = 600$.

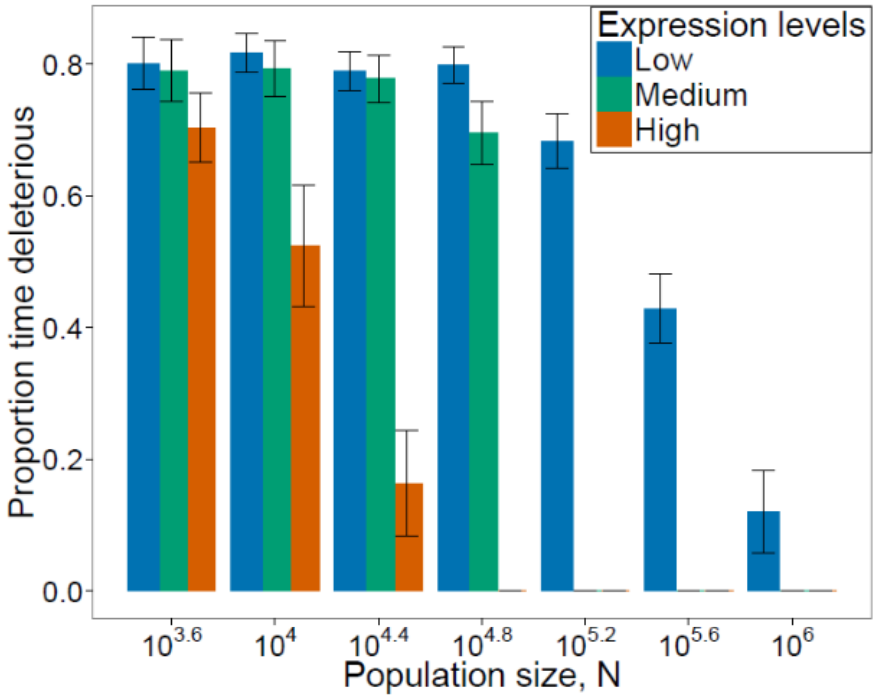

Figure 2: The effectiveness of purging a deleterious cryptic sequences of deleterious mutations depends on its expression level. We examined the states of the cryptic sequences of the loci with the 10 highest, the 10 lowest, and the 10 median expression levels among the 600 loci in each of the simulations showed in Fig. 1 (σ_E = 2.25). We counted how often each locus contained a deleterious cryptic sequence among the last 20 timepoints we had collected from that simulation. Bars represent the proportion of time that each of the 10 loci carried a deleterious cryptic sequences, averaged over 20 replicates, and shown as mean $\pm$SD. Simulations were initialized at low-ρ conditions.

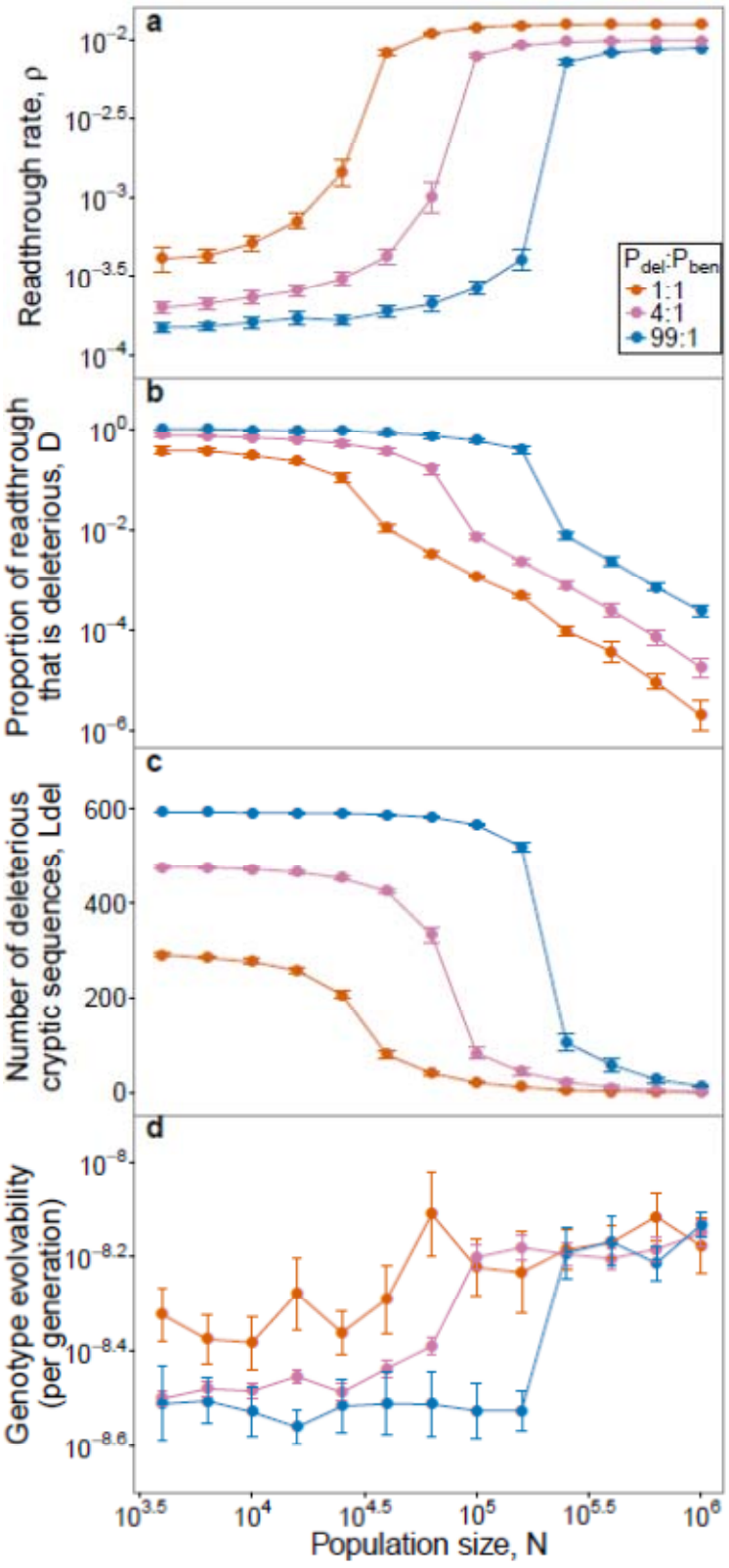

Figure 3: Results become more extreme when the mutation bias in the state of a cryptic sequence is increased from 4:1 ratio to a 99:1 ratio, but do not disappear completely when the mutation bias is eliminated in favor of a 1:1 ratio. The location of the drift barrier shifts as a function of mutation bias, but the dichotomy between local and global solutions (as seen in values of ρ and D) is not sensitive to relaxing the mutation bias. The advantage of the local solution with respect to evolvability (as seen in **d** and mirrored in L_{del} (**c**)) is more sensitive to lack of mutation bias, but is still visible even with a 1:1 ratio. To compare results across different mutation biases, we kept the sum of the two mutation rates constant. For the low-ρ initial conditions, the number of deleterious cryptic sequences was initialized at the neutral mutational equilibrium of $L \times P_{del}/(P_{del}+P_{ben})$ (rounded to the nearest integer). For P_{del}:P_{ben} = 4:1, we reused the results shown in Fig. 1. For the other ratios, five replicates were run for each initial condition, and pooled. For panels **a** to **c**, data is shown as mean±SD. For panel **d**, data is based on time to fitness recovery and is shown as mean±SE. L = 600 and σ_E = 2.25.

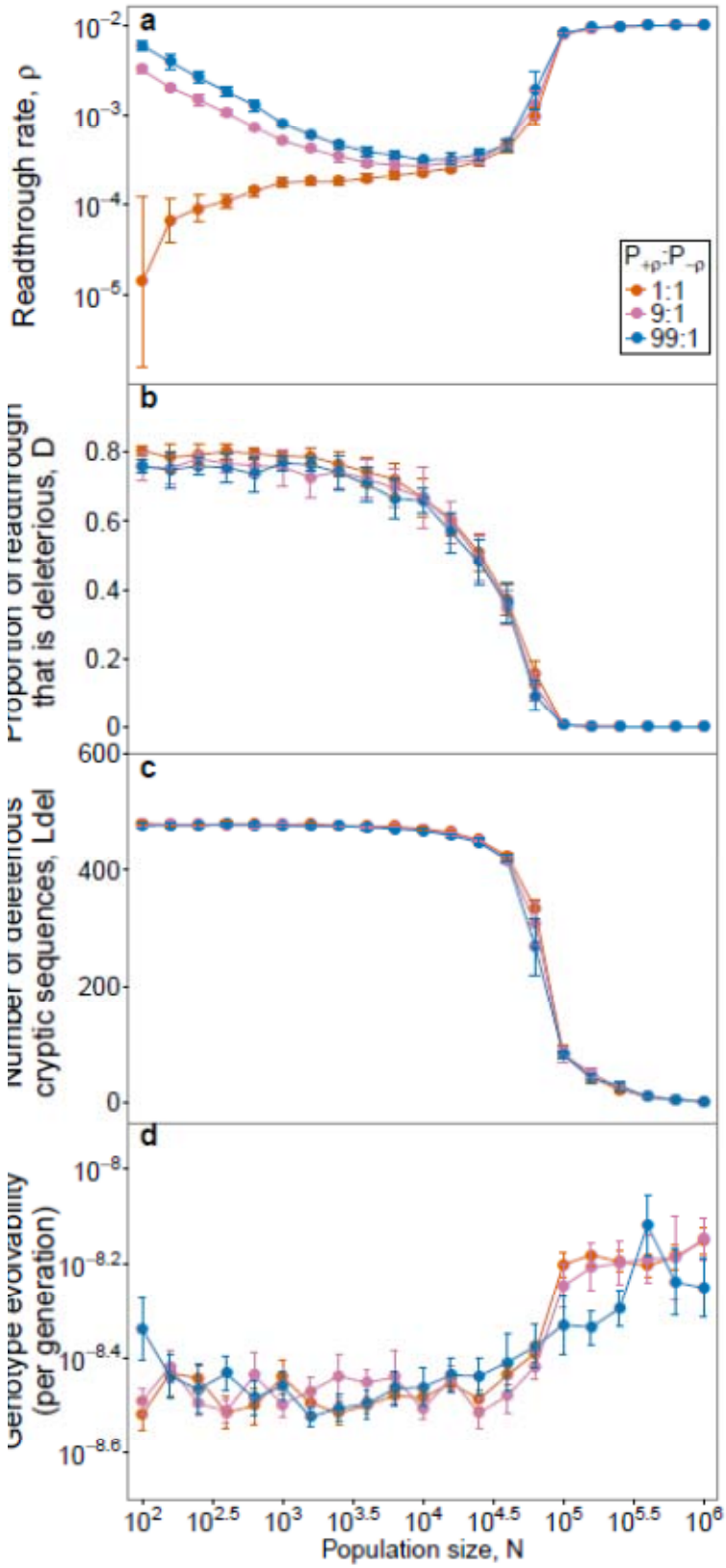

Figure 4: Mutation bias tends to increase ρ, such that even the global solution breaks down in sufficiently small populations. $P_{+\rho}$ is the probability that a mutation increases ρ, and $P_{-\rho}$ is the probability of a decrease. Each data point, (except those taken from Fig. 1 with $P_{+\rho}$:$P_{-\rho}$ = 1:1 and $N = 10^{3.6}$ to $N = 10^{6.0}$), is pooled from 5 replicates of high-ρ initial conditions and 5 replicates of low-ρ initial conditions. Because we assume multiplicative mutational effects to ρ, its value converges even for extremely small N. I.e., as ρ increases, the additive effect size $\Delta\rho$ of a typical mutation also increases, preventing it from passing through the drift barrier. For **a**, **b**, and **c**, data is shown as mean$\pm$SD. For **d**, data is shown as mean$\pm$SE. L = 600 and σ_E = 2.25. Note that large N evolvability goes down with extreme mutation bias: this is because reductions to ρ are sometimes beneficial when $\sum \beta_k$ is in a favorable direction for trait evolution; such opportunities are foregone in

these populations.