Not quite lost in translation: Mistranslation alters adaptive landscape topography and the dynamics of evolution

Michael Schmutzer[1,2], Andreas Wagner[1,2,3*],

1 Department of Evolutionary Biology and Environmental Studies, University of Zürich, Zürich, Switzerland
2 Swiss Institute of Bioinformatics, Lausanne, Switzerland
3 Santa Fe Institute, Santa Fe, New Mexico, USA

* andreas.wagner@ieu.uzh.ch

Abstract

Mistranslation – the erroneous incorporation of amino acids into nascent proteins – is a source of protein variation that is orders of magnitude more frequent than DNA mutation. Like other sources of nongenetic variation, it can affect adaptive evolution. We study the evolutionary consequences of mistranslation with experimental data on mistranslation rates applied to three empirical adaptive landscapes. We find that mistranslation generally flattens adaptive landscapes by reducing the fitness of high fitness genotypes and increasing that of low fitness genotypes, but it does not affect all genotypes equally. Most importantly, it increases genetic variation available to selection by rendering many neutral DNA mutations non-neutral. Mistranslation also renders some beneficial mutations deleterious and vice versa. It increases the probability of fixation of 3 to 8 percent of beneficial mutations. Even though mistranslation increases the incidence of epistasis, it also allows populations evolving on a rugged landscape to evolve modestly higher fitness. Our observations show that mistranslation is an important source of non-genetic variation that can affect adaptive evolution on fitness landscapes in multiple ways.

Introduction

Living organisms are no clockworks. The mechanisms by which they convert the information stored in their genes into phenotypes are imprecise and error-prone. As a result, even genetically identical individuals raised in the same environment vary in traits that affect fitness (Ackermann, 2015; Govers et al., 2017; Kotte et al., 2014; Solopova

et al., 2014; van Boxtel et al., 2017). The resulting nongenetic variation in fitness cannot be inherited with high fidelity (if at all). Any selective advantages or disadvantages it may cause will thus generally disappear within one generation (Bonduriansky and Day, 2018).

Nongenetic fitness variation is ubiquitous and present at all levels of biological organisation, from the structure of macromolecules to the behaviour of whole organisms. For example, both RNA and protein molecules undergo rapid changes in folding due to Brownian motion with biological consequences that are readily acted upon by natural selection (Chen et al., 2010; Wagner, 2014). Another example is the accuracy of biological catalysts: Enzymes are often promiscuous, both in the sense that they can catalyse reactions with more than one kind of substrate, but also different reactions with the same substrate (Tawfik and S, 2010). Metabolic enzymes are sufficiently error-prone that cells have evolved specialised enzymes to remove undesirable metabolites (Peracchi, 2018). Further examples include gene expression noise (Bódi et al., 2017; Elowitz et al., 2002; McAdams and Arkin, 1999), errors in transcription (Giacomelli et al., 2007; Traverse and Ochman, 2016), and multiple alternative states in regulatory systems (Bruggeman and Teusink, 2018; Espinosa-Soto et al., 2011; van Heerden et al., 2014) that underlie behaviours such as bet-hedging (Carey et al., 2018; Solopova et al., 2014; van Boxtel et al., 2017), as well as mistranslation, i.e. the incorporation of amino acids different from those encoded in mRNA into a protein (Mordret et al., 2019). Here we

focus on mistranslation and its role for adaptive evolution.

Even though nongenetic fitness variation is not highly heritable, it can affect the rate of adaptation (Draghi, 2018; Frank, 2011; Rajon and Masel, 2011; Rocabert et al., 2020; Whitehead et al., 2008). On the one hand, it can decrease the power of selection to act on genetic variation, for example by increasing the influence of genetic drift on a population (Mineta et al., 2015; Wang and Zhang, 2011). On the other hand, nongenetic fitness variation can increase the fitness effects of some mutations. Specifically, it may increase the probability that beneficial mutations reach fixation, and it may increase the rate at which deleterious mutations are driven to extinction. In consequence, it may accelerate a populations's rate of adaptation (Frank, 2011; Rajon and Masel, 2011; Rocabert et al., 2020; Tănase-Nicola and ten Wolde, 2008; Zhang et al., 2009). For example, a yeast strain with higher variation than another strain in the expression of a protein conferring resistance to the antifungal drug fluconazole derives a larger fitness benefit from mutations in this protein, and thus evolves high resistance faster (Bódi et al., 2017). More generally, because of conflicting pertinent evidence, it is unclear whether nongenetic fitness variation is more likely to hinder or facilitate adaptation (Draghi, 2018), and quantification of their relative occurrence is overdue.

Whether nongenetic fitness variation benefits adaptation may depend on the dimensionality of a trait (Frank, 2011; Rocabert et al., 2020). For example, gene expression noise may be able to accelerate adaptation in the expression level of a single gene (a

single dimension; (Bódi et al., 2017)), but it might also slow down the evolution of a regulatory network involving expressiong of multiple genes (many dimensions). The reasion is that any benefit derived from the expression level of one particular gene may be eliminated by detrimental expression noise in other genes. This 'cost of complexity' (Rocabert et al., 2020) may prevent nongenetic fitness variation from accelerating adaptation (Frank, 2011; Rocabert et al., 2020).

Adaptive landscapes map genotypes within a collection or "space" of genotypes onto fitness or some proxy thereof. They are powerful conceptual tools for studying long-term evolution (Aguirre et al., 2018; Bataillon et al., 2022; Ferretti et al., 2018; Wright, 1932; Wu et al., 2016). Because genotype spaces are high-dimensional discrete spaces, they have geometrical properties different from those of low-dimensional continuous spaces (Aguirre et al., 2018). For example, epistasis leads to rugged landscapes with many local peaks (Wu et al., 2016; Zagorski et al., 2016). In low dimensional spaces, epistasis can be an important barrier to adaptation, as populations can get stuck at local fitness optima. As dimensionality increases, more opportunities for escaping or bypassing local optima become available, and thus a population evolving on a high dimensional, rugged landscape is more likely to reach higher fitness genotypes than in a low-dimensional landscape (Wu et al., 2016; Zagorski et al., 2016).

Mistranslation is an unavoidable property of the translation machinery (Ribas de Pouplana et al., 2014), which erroneously incorporates amino acids into nascent proteins

at rates between 10^{-5} to 10^{-2} per amino acid (Mordret et al., 2019; Ribas de Pouplana et al., 2014). As a result, about 10% of all proteins in *E. coli* will have an amino acid sequence that is different from their genetically encoded one (Ellis and Gallant, 1982; Ruan et al., 2008). Not all amino acid substitutions are equally likely, and some codons have a much higher rate of mistranslation than others (Mordret et al., 2019).

For three reasons, mistranslation is highly useful to investigate the evolutionary impact of nongenetic fitness variation. First, the process linking a DNA genotype to phenotype – a set of polypeptides produced through error-prone translation – is well understood (2019; Ribas de Pouplana et al., 2014). The second reason stems from recent breakthroughs in high-throughput fitness assays (Blanco et al., 2019), which allow the mapping of large experimentally determined adaptive landscapes containing up to more than 10^5 polypeptide sequences (Lite et al., 2020; Wu et al., 2016). Such landscapes permit modeling the fitness consequences of producing mistranslated polypeptides. Third, mistranslation can affect the course of adaptation. Even though it detrimentally impacts cell physiology and fitness (Bratulic et al., 2017; Javid et al., 2014; Ribas de Pouplana et al., 2014; Samhita et al., 2021), mistranslation can be beneficial for adaptation. For example, higher rates of mistranslation increase resistance to the antibiotic rifampicin in mycobacteria (Javid et al., 2014). Increased mistranslation of the *E. coli* antibiotic resistance gene TEM-1 β-lactamase facilitates the evolution of protein stability (Bratulic et al., 2015), and helps purge deleterious mutations (2017; Zheng et al., 2021).

Here we build a quantitative model of the fitness effects of mistranslation that draws on experimental measured mistranslation rates (Mordret et al., 2019) and experimentally characterised adaptive landscapes (Lite et al., 2020; Wu et al., 2016). We used this model to investigate how mistranslation affects adaptive evolution.

Results

Mistranslation causes the fitness of a cell expressing a given genotype to depend not only on the protein encoded in the nucleotide sequence of the genotype (figure 1a), but also on mistranslated variants of the same protein (figure 1b). Because translation errors are random, mistranslation also causes variation between cells in the kind of protein variants they contain. These protein variants may have different activities, and thus cause variation in fitness among genetically identical cells.

We combine data from experimentally mapped adaptive landscapes (figure 1c) with experimentally measured mistranslation rates (figure 1d) to quantify the average change in fitness and the variation in fitness between individuals that is caused by mistranslation. The same data also allows us to quantify how mistranslation changes the outcome of adaptive evolution. To this end we simulate adaptive evolution on the adaptive landscapes defined by the data. Adaptive evolution occurs at the nucleotide sequence level, and can be influenced by the mistranslation of individual codons on the mRNA into different amino acids. Our model allows us to estimate the protein variant compositions of a

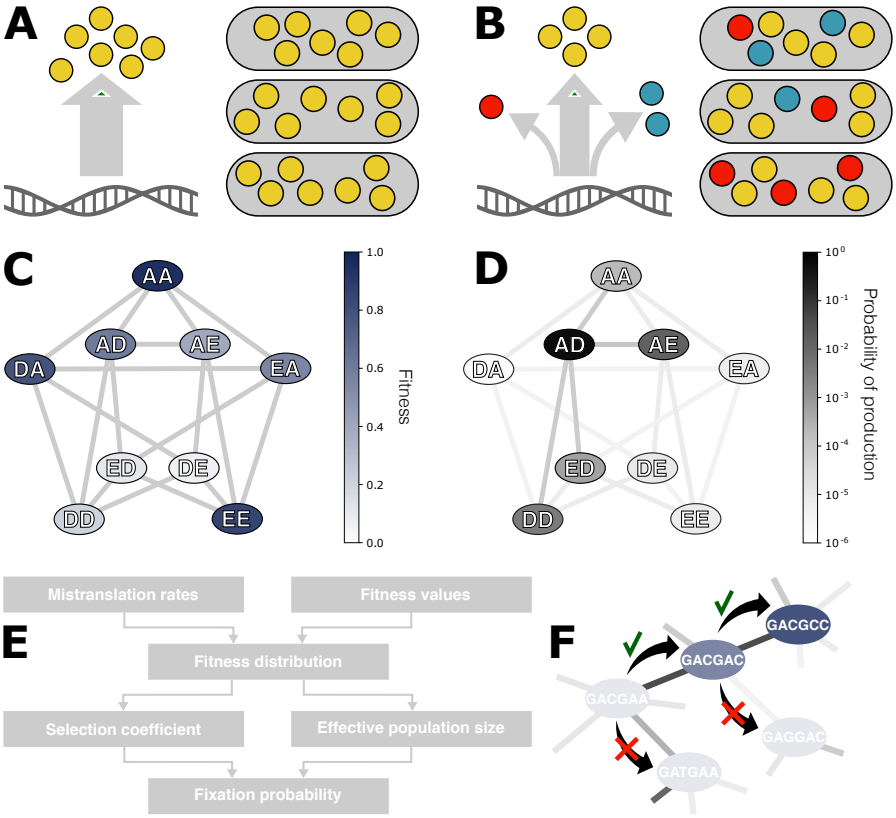

Figure 1: Modelling evolution under the influence of mistranslation. A) In the absence of mistranslation, cells produce proteins with exactly the amino acid sequence encoded in their genes. (B) Mistranslation causes the production of several protein variants (coloured circles) in different proportions (arrow thickness) from the same genotype. These protein variants do not necessarily have the same activity and can affect the fitness of a cell. Because mistranslation is a random process, cells with identical genotypes may still vary in protein variant composition and potentially in fitness. (C) We determined the effects of mistranslation from experimental data that quantified the fitness (or its proxy) of genotypes in an adaptive landscape at the polypeptide sequence level, as illustrated here with a toy example. Each polypeptide sequence (ellipses, linked to polypeptides that differ by one amino acid with grey edges) is encoded by a set of synonymous nucleotide sequences. (D) We simulate mistranslation using experimentally measured mistranslation rates (Mordret et al., 2019). Each codon differs in its mistranslation rates, and thus each nucleotide sequence has a different propensity to produce alternative protein variants . For example, a nucleotide sequence encoding the polypeptide "AD" (black ellipse) has a higher probability of producing certain polypeptide sequences with a single amino acid substitution than others (e.g. "AE" is more likely than "ED" in this example), and is even less likely to produce a protein variant with a double substitution ("DA"). We calculate the composition of protein variants expected under mistranslation for a given nucleotide sequence, and then use the fitness measurements of these protein variants from (C) to estimate a cell's expected fitness. (E) Model work flow. Using the experimental measurements of fitness and mistranslation rates, we estimate the fitness distributions of a cell expressing a given genotype and of a cell expressing a mutant genotype. From these two fitness distributions we calculate the mutant's selection coefficient and the effective population size, which is influenced by mistranslation (Methods). We then use these estimates and Kimura's (Kimura, 1962) equation to estimate the fixation probability of the mutation. (F) We simulate evolution at the nucleotide sequence level in a weak mutation-strong selection regime, where new mutant genotypes fix rapidly compared to the time they need to originate. In this regime populations harbour only one allele at the loci we simulate for most of the time. Consequently, evolution resembles a 'walk' from one genotype (ellipses, coloured by fitness, edges connect genotypes differing in one nucleotide position) to the next. Each step in the walk is equivalent to the fixation of a novel genotype. In our model, mutations change a single nucleotide position (black arrows). Every mutation is equally likely to occur, but its fixation probability (edge shading) depends on its fitness relative to the current genotype. Mutations that decrease fitness are likely (but not necessarily) lost (red crosses), while mutations that increase fitness are likely to reach fixation (green check-marks).

population of cells expressing a given nucleotide sequence (figure 1d) from these codon mistranslation rates (Mordret et al., 2019). We then estimate the fitness distribution of cells in the population based on the fitness of these protein variants (figure 1d). Using the fitness distributions (figure 1e) of a population carrying a given ("wild-type") genotype and another carrying a mutant genotype, we can then estimate the selection coefficient of the mutant, which we define as the relative difference in mean fitness between these genotypes. We also estimate the effective population size, which can be affected by the mistranslation rate (Wang and Zhang, 2011). These two quantities define the fixation probability of the mutant genotype when introduced at a low frequency into a population consisting only of wild-type genotypes.

Because the experimentally determined landscapes we consider (Lite et al., 2020; Wu et al., 2016) are based on a small number of nucleotide sites, we assumed that adaptive evoluiton occurs in the weak mutation, strong selection regime (Gillespie, 1983; 1984). In this regime, only one mutation segregates at any point in time, and adaptive evolution effectively becomes an adaptive walk in which a population steps from one genotype to the next (1984; Maynard Smith, 1970; Wright, 1932) (figure 1f). The probability of each step is defined by the fixation probabilities we calculated from the data. While simulating such adaptive walks, we quantified evolutionary time through the number of mutations that occurred since the beginning of the walk. We kept track of the small fraction of these mutations that go to fixation.

Using such adaptive walks, we studied how mistranslation affects adaptive evolution on three adaptive landscapes that are combinatorially complete for all 20 amino acids (Methods, (Wu et al., 2016)). The first of these landscapes comprises measurements of the binding affinity of almost 160,000 genotypic variants of the streptococcal immunoglobulin-binding protein (GB1), which can bind antibodies (2016). In using data from this landscape, we follow other authors in considering strong molecular binding as a proxy for high fitness (Aguilar-Rodríguez et al., 2017; Natarajan et al., 2013; Olson et al., 2014; Sarkisyan et al., 2016). We will refer to this landscape as the antibody-binding landscape.

The second and third landscapes comprise data on the growth rate of *E. coli* cells transformed with almost 8,000 variants of a toxin-antitoxin system (Methods, (Lite et al., 2020)). Specifically, they regard fitness data of the ParD3 antitoxin, part of a toxin-antitoxin system that targets topoisomerases (Jiang et al., 2002; Yuan et al., 2010). These variants of the ParD3 antitoxin were either expressed together with its native ParE3 toxin, or with its close homologue ParE2, resulting in two fitness measurements per antitoxin variant. We refer to these landscapes as the toxin-antitoxin landscapes, and differentiate them where necessary by the target toxins E3 or E2.

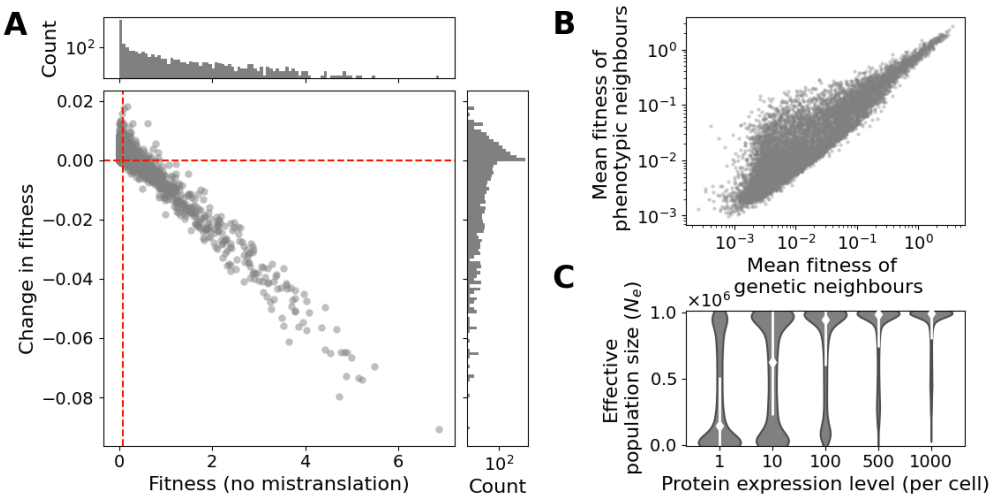

Figure 2: Mistranslation flattens the fitness distribution of the antibody-binding landscape. A) Main panel: The horizontal axis shows the fitness of 10^4 randomly chosen genotypes without mistranslation (distribution in small upper panel). The vertical axis shows the change in fitness when the same set of genotypes is expressed with and without mistranslation (the distribution of these changes is shown in the small panel to the right). The vertical red dotted line marks the average fitness in the antibody-binding landscape, and the horizontal red dotted line marks zero change in fitness. B) The mean fitness of a genotype's phenotypic neighbours (polypeptide sequences that differ in one amino acid position from the genotype, vertical axis) and the mean fitness of its genotypic neighbours (genotypes with a single nucleotide change, horizontal axis) are positively correlated. C) Mistranslation reduces the effective population size. White diamonds and lines show the median effective population size and the standard deviation, respectively. The violin plots are a Gaussian kernel density estimate of the distribution of effective population sizes. In all plots results are shown for the same set of 10^4 randomly chosen genotypes from the antibody-binding landscape, at a population size of $N = 10^6$, and an expression level of one protein per cell (unless where stated otherwise).

Mistranslation flattens adaptive landscapes

We first investigated the impact of mistranslation on properties of genotypes that are important for adaptive evolution. Specifically, we assessed how mistranslation affects (i) the expected (mean) fitness of genotypes, (ii) the effective population size (Methods), (iii) the fitness effects of mutations, and (iv) the fixation probabilities of mutations. We characterised the effect of mistranslation for 10^4 randomly chosen genotypes from each landscape and at three population sizes $N = 10^4$, 10^6 and 10^8. We found that many, but not all, of the effects of mistranslation are subtle, because the mistranslation rates we used are realistically small (Mordret et al., 2019). Our results in this section are very similar for all three landscapes, therefore we will focus on the antibody-binding landscape (see supplementary materials figures 1 and 2 for the toxin-antitoxin landscapes).

The production of mistranslated polypeptides can change the mean fitness of a population of cells, because these mistranslated polypeptides vary in their effect on fitness. The extent of this change depends on a genotype's fitness in the absence of mistranslation (figure 2a). In the antibody-binding landscape, the change in fitness of a genotype due to mistranslation is strongly negatively correlated with the fitness of that genotype in the absence of mistranslation (Pearson's $r = -0.94$, at a population size of 10^6), even though mistranslation causes little fitness change in the vast majority of genotypes. Mistranslation tends to increase the fitness of genotypes with a fitness below the mean fitness of 0.08 in the antibody-binding landscape (vertical red line in figure

2a). In contrast, mistranslation tends to decrease the fitness of genotypes with higher fitness. Overall, mistranslation causes the landscape to become "flattened", with high fitness genotypes decreasing in fitness due to mistranslation, and low fitness genotypes increasing in fitness.

We also found that the fitness values of the phenotypic neighbours of a given genotype (the polypeptide sequences that differ by one amino acid from the polypeptide encoded by the genotype) are correlated with the fitness values of its genotypic neighbours (figure 2b, Kendall's $\tau = 0.72$, $p = 0.0$, $n = 10^4$). In order words, the nongenetic fitness variation created through mistranslation is related to the genetic fitness variation present in the immediate neighbourhood of a genotype in sequence space. This is a necessary condition for mistranslation to facilitate adaptative evolution (Ancel and Fontana, 2000; Frank, 2011; Rocabert et al., 2020). As a result of this correlation, the fixation of a genotype with higher fitness will not only increase the mean fitness of a population, but is also likely to increase the probability that subsequent mutations will increase fitness even more.

Mistranslation strengthens genetic drift

Mistranslation causes variation in the protein composition between genetically identical cells, which can induce small variations in fitness between cells. The degree of fitness variation between individuals depends on a protein's expression level. When few copies

of a protein exist in a cell, a single mistranslated copy can affect fitness more substantially than when many copies exist. Consequently, fitness will vary more strongly between cells that produce fewer copies of a protein. This kind of nongenetic fitness variation can weaken selection, an effect that can be quantified as a decrease in the effective population size (Methods, (Wang and Zhang, 2011)). To determine the strength of this effect, we thus quantified the effect of protein expression level on the effective population size, and did so for protein expression levels between 10^0 to 10^3 proteins per cell.

As predicted, we found that mistranslation causes a greater reduction in the mean effective population size at lower than at higher expression levels (figure 2c). For example, mistranslation causes effective population size in a population of $N = 10^6$ individuals to decrease on average by 66 percent to $N_e = 3.4 \times 10^5$ ($\pm 3.7 \times 10^5$ standard deviation) for a protein expressed at one copy per cell, but by only 11 percent to $N_e = 8.9 \times 10^5$ ($\pm 2.0 \times 10^5$) for a protein expressed at 10^3 copies. Effective population size also varies greatly even within one expression level depending on the protein genotype that is considered. The reason is that different genotypes vary in their rates of mistranslation and the fitness values of their phenotypic neighbours. Similar observations hold for the toxin-antitoxin landscapes (see supplementary materials figures 1 and 2). Throughout the rest of this paper we will distinguish between the population size N and the effective population size $N_e \leq N$ that is a consequence of mistranslation.

Mistranslation reduces neutrality and can cause deleterious mutations to turn beneficial and vice versa

Changes in the fitness of genotypes may also affect the fitness of mutations. To determine how strongly mistranslation affects the fitness of mutations, we sampled 10^4 genotypes at random from the antibody-binding landscape, and chose one of their mutational neighbours, i.e. a genotype that differs from the focal genotype by one nucleotide substitution. We classified the fitness effects of these mutants as beneficial, deleterious or (nearly) neutral (if $|s| < \frac{1}{4N_e}$ (Charlesworth, 2009)), compared to the wild-type genotype, both in the presence and absence of mistranslation (figure 3a).

This analysis showed that mistranslation can cause substantial changes in fitness effects. In particular, mistranslation dramatically reduces the number of nearly neutral mutations. For example, at a population size of $N = 10^6$ and a low expression level of one copy per cell, 37% of mutations in the antibody-binding landscape are nearly neutral in the absence of mistranslation, but only 4% are nearly neutral in the presence of mistranslation and low expression. 56% of the mutations that are nearly neutral in the absence of mistranslation are synonymous mutations, and 79% of these synonymous nearly neutral mutations come under selection in the presence of mistranslation. 16% of nonsynonymous mutations are neutral in the absence of mistranslation, and all of these mutations become non-neutral under mistranslation. Moreover, in the presence of mistranslation the fitness differences between nonsynonymous mutations are orders of

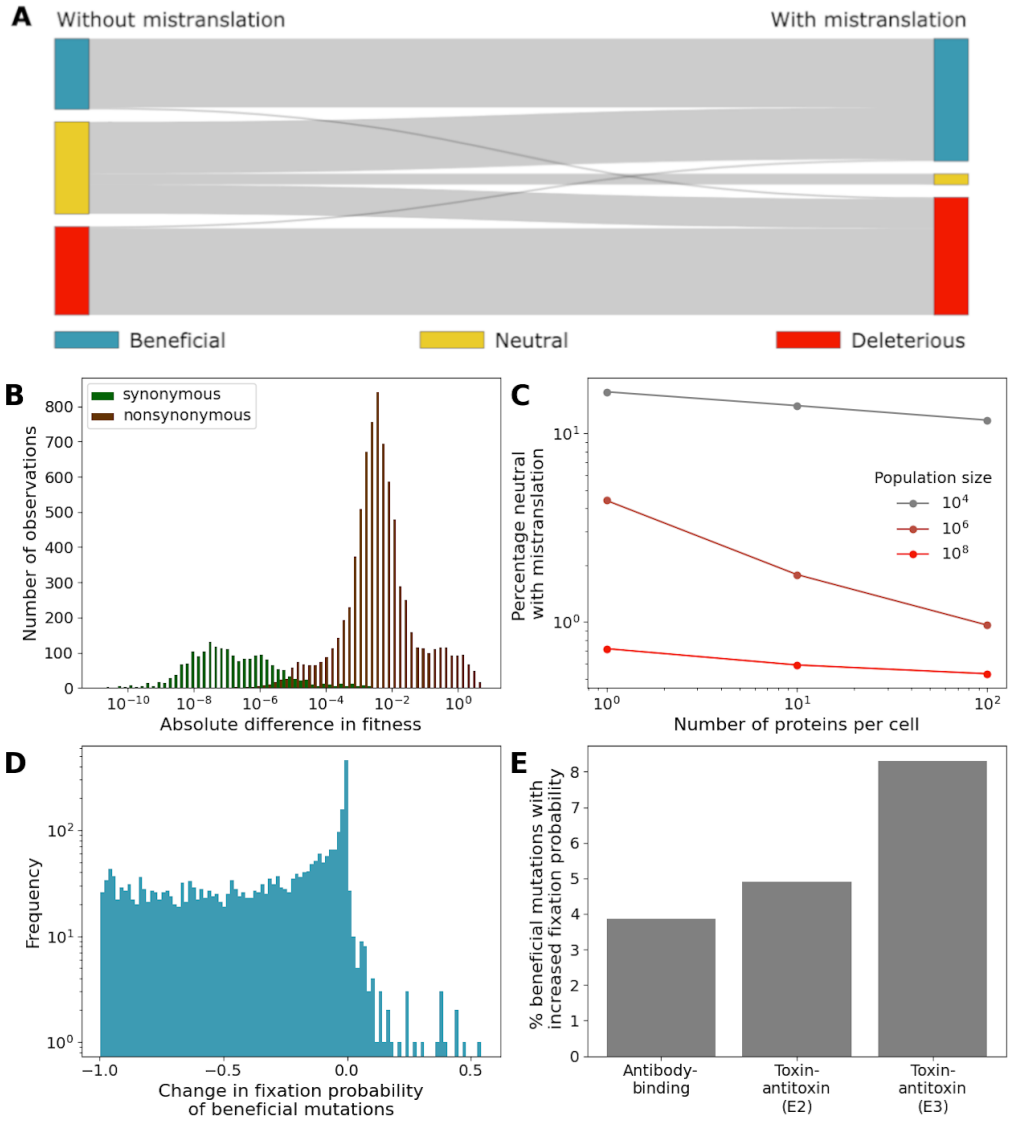

Figure 3: Mistranslation changes the fitness effects of some mutations. A) Changes due to mistranslation in the proportions of mutations classified according to their fitness effects as beneficial, (nearly) neutral, or deleterious. B) Distributions of the absolute differences in mean fitness between both synonymous (green) and nonsynonymous (brown) pairs of neighbouring genotypes sampled from the antibody-binding landscape in the presence of mistranslation at an expression level of one protein per cell. Only nonzero fitness differences are shown. C) Percentage of mutations classified as neutral in the presence of mistranslation for multiple combinations of the population size N and the protein expression level, both of which impact the effective population size N_e (and thus neutrality). D) Distribution of the change in fixation probabilities of beneficial mutations due to mistranslation. Only those mutations are shown that are beneficial both with and without mistranslation. The horizontal axis shows the difference in fixation probability due to mistranslation, with zero denoting no change. For example, changes in fixation probability close to negative one signify mutations that are almost certain to fix without mistranslation ($u_{fix} \approx 1$), but have almost no chance of reaching fixation with mistranslation ($u_{fix} \approx 0$). E) Percentage of beneficial mutations that increase in fixation probability due to mistranslation for each landscape. In all plots (except where stated otherwise), results are shown for the same set of 10^4 randomly chosen genotypes from the antibody-binding landscape as in figure 2, and randomly chosen one-step mutational neighbours, at a population size of $N = 10^6$, and an expression level of one protein per cell. The same parameters apply for the two toxin-antitoxin landscapes in (E).

magnitude larger than the differences between synonymous mutations (figure 3b). This loss of nearly neutral mutations is influenced both by the population size and the protein expression level (figure 3c). The loss of nearly neutral mutations also severely affects the size of nearly neutral networks (see supplementary materials section 3).

In addition, we observed that 2.5% of mutations that are beneficial in the absence of mistranslation become deleterious in its presence, and 2.1% of mutations change from deleterious to beneficial (figure 3a). The reason is that the flattening of the fitness landscape caused by mistranslation does not happen to the same extent for all genotypes, because genotypes differ both in their rate of mistranslation and their genetic neighbourhoods. For example, a mutation that is beneficial in the absence of mistranslation may become deleterious because it has a higher mistranslation rate than the wild-type, or because the mutation increases the probability of producing a highly deleterious protein variant through mistranslation.

Even when mistranslation does not change the fitness effect of a mutation from beneficial to deleterious, it can affect the mutation's fixation probability (figure 3d). For example, in the antibody-binding landscape 95.7 percent of beneficial mutations experience a reduction in this probability, and only 3.9 percent of mutants experience an increase. In this respect, the toxin-antitoxin landscapes are different from the antibody-binding landscape. For example, the fraction of beneficial mutations whose fixation probability increases is much greater (8.3 percent) in the toxin-antitoxin (E3) landscapes

(figure 3e). For the deleterious mutations we examined, the fixation probability is largely unaffected with a mean increase of $2.81 \times 10^{-79} \pm 1.65 \times 10^{-77}$ in the antibody-binding landscape, and a mean increase of $6.69 \times 10^{-15} \pm 3.64 \times 10^{-13}$ in the E3 toxin-antitoxin landscape, but a mean decrease of $6.89 \times 10^{-101} \pm 3.62 \times 10^{-99}$ in the E2 toxin-antitoxin landscape ($N = 10^6$).

Mistranslation helps evolving populations attain higher fitness on two landscapes

We next investigated how mistranslation influences adaptive evolution on all three adaptive landscapes. To this end, we chose 10^4 starting genotypes from each landscapes at random among the viable genotypes that ranked in the bottom 10% of the fitness distribution. We performed adaptive walks (Methods) for three population sizes (10^4, 10^6, 10^8), and two levels of protein expression. Specifically, we considered expression levels of one and 500 proteins per cell, where 500 proteins is the median expression level in *E. coli* (Ishihama et al., 2008). We will refer to these expression levels as low and high expression respectively. We will refer to the number of mutations (both fixed and extinct) since the beginning of a walk as a proxy for evolutionary time.

We found that mistranslation facilitates adaptation in the antibody-binding landscape, but much less so or not at all in the two toxin-antitoxin landscapes. Specifically, at the end of adaptive walks on the antibody-binding landscape, populations with mistranslation

attain slightly higher mean fitness, and this advantage increases with larger population sizes. For example, for populations with $N = 10^4$ individuals and low expression, the average adaptive walk ended at genotypes that have 52.3 ($\pm$ 26.3, one standard deviation) and 50.2 ($\pm$ 27.7) percent of the maximum fitness, respectively, in the presence and absence of mistranslation. These differences are small but statistically significant (Dunn's test with Bonferroni correction, $p \leq 4.09 \times 10^{-9}$).

At a population size of 10^8 adaptive walks with mistranslation reach a fitness that is 3.5 percent higher than in its absence (50.8 $\pm$ 25.8 and 47.3 $\pm$ 27.2, percent of the maximum, respectively, with significant differences only between walks with and without mistranslation (Dunn's test with Bonferroni correction, $p = 1.03 \times 10^{-19}$ or smaller). Compared to the antibody-binding landscape the effect of mistranslation on the outcome of adaptive walks on the toxin-antitoxin (E3) landscape is an order of magnitude smaller, and we did not observe a consistent pattern in the toxin-antitoxin (E2) landscape (supplementary materials table 1 and section 4).

Protein expression level has little effect on the outcome of adaptive evolution with mistranslation. For example, for adaptive walks on the antibody-binding landscape, populations with $N = 10^8$ individuals reached 50.8$\pm$25.8 and 50.2$\pm$26.2 percent of the maximum fitness, a difference in outcome that is not statistically significant (Dunn's test with Bonferroni correction, $p = 0.944$). Similar observations hold for other population sizes on the antibody-binding landscape and on the toxin-antitoxin landscapes (see

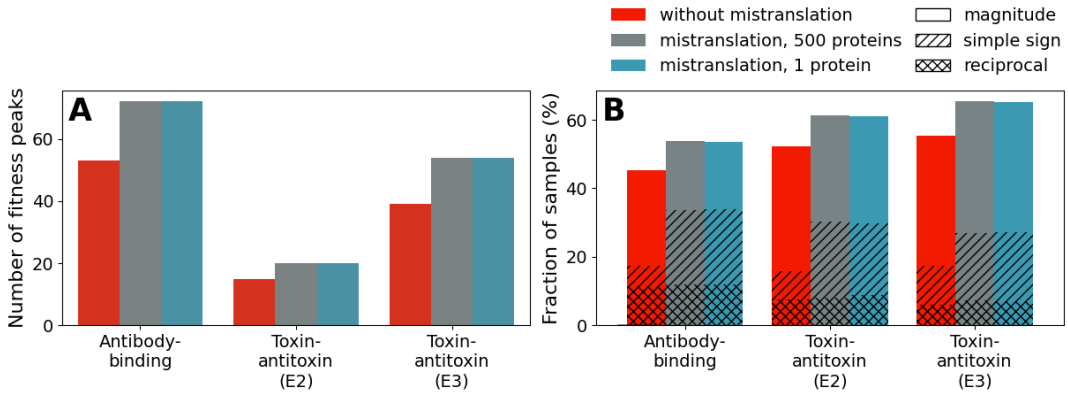

Figure 4: Mistranslation increases the ruggedness of all three landscapes. A) Number of fitness peaks among 10^5 genotypes with the highest fitness in all three landscapes, without mistranslation (red), and with mistranslation at both high and low protein expression levels (grey and blue, respectively). Fitness peaks were evaluated at large population sizes ($N = 10^8$). B) Percentage of 10^4 genotype samples classified as one of three kinds of epistasis. Each sample consists of a "square" of four genotypes, two of which are double nucleotide mutants of each other, and the other two are single mutant intermediates. If the fitness effects of the double mutant is not additive with respect to the fitness effects of the single mutants, the two mutations are epistatic, with the kind of epistasis depending on the nature of the deviation from additivity (Methods, (2007), see also supplementary figure 5). For all three landscapes, we determined epistasis for the same set of 10^4 sequences without and with mistranslation, and in the latter case for high and low protein expression.

supplementary materials tables 1 and 2). However, protein expression level does affect how fast adaptive walks attain high fitness (see supplementary materials section 5).

Mistranslation increases the ruggedness of landscapes

Mistranslation may change the ruggedness of a fitness landscape with multiple peaks, and thus also how accessible high fitness genotypes are. The reason is that in a rugged landscape, evolving populations are more likely to become stuck at peaks of low fitness (2007). Because mistranslation may 'smooth over' some of the roughness of a landscape, it may render high fitness genotypes genotypes more accessible through fitness increasing paths (Frank, 2011). Intriguingly, the following analysis does not support this prediction. We estimated the ruggedness of the three landscapes with two complementary measures,

the number of fitness peaks in a landscape and the frequency of different kinds of epistasis. For the purpose of our analysis, a fitness peak can consist of a single genotype or a network of nearly neutral genotypes that are single-nucleotide mutational neighbours of one another (Methods). We consider any such nearly neutral network a fitness peak if all single-step mutational neighbours of the genotypes in the network have a lower fitness.

Because the size of nearly neutral networks is affected by population size, we evaluated the number of fitness peaks at different population sizes. Mistranslation increases the number of fitness peaks (figure 4e), except at low population sizes (supplementary materials figure 3). Most of these additional fitness peaks arise because synonymous mutations come under selection in large populations.

Another measure of ruggedness is the frequency of multiple kinds of non-additive (epistatic) interactions between pairs of single nucleotide mutants (Poelwijk et al., 2007). One can distinguish three kinds of such non-additivity (epistasis): Magnitude epistasis, simple sign epistasis and reciprocal sign epistasis. The latter two are associated with the existance of fitness valleys that may decrease the number of accessible (monotonically fitness-increasing) paths towards high fitness peaks (2007). Our analysis shows that mistranslation makes all kinds of epistasis more common, particularly magnitude and simple sign epistasis, and to a lesser extent also reciprocal sign epistasis (figure 4f). Taken together, these observations show that mistranslation increases the ruggedness of a landscape. Thus, if mistranslation increases the accessibility of high fitness peaks, it

does not do so by landscape smoothing.

Synonymous mutations facilitate continued adaptation at high fitness

Most of the increase in fitness we discussed above happens within the first 10^3 mutations of an adaptive walk. This increase is largely driven by nonsynonymous mutations. For example, nonsynonymous mutations contributed on average 99.98 ± 0.03 percent of the change in fitness during walks with mistranslation on the antibody-binding landscape ($N = 10^8$, high protein expression). However, evolution involving both selection and drift may continue for a long time among high fitness genotypes. Specifically, we hypothesised that mistranslation can help to lengthen adaptive walks, because high fitness genotypes subject to mistranslation largely produce deleterious protein variants. For this reason, selection may act to prefer codons associated with low mistranslation. In other words, populations with mistranslation may not only experience purifying selection of deleterious genotypes (like populations without mistranslation do), but in addition also positive selection of mutations that decrease mistranslation rates.

To test this hypothesis, we studied our adaptive walks in greater detail. In support of it, we found that in all three landscapes, adaptive walks with mistranslation fixed significantly more mutations than walks without mistranslation, except at small population sizes and in walks on the antibody-binding landscape with mistranslation and low expression (supplementary materials table 3). The difference is small, partly because

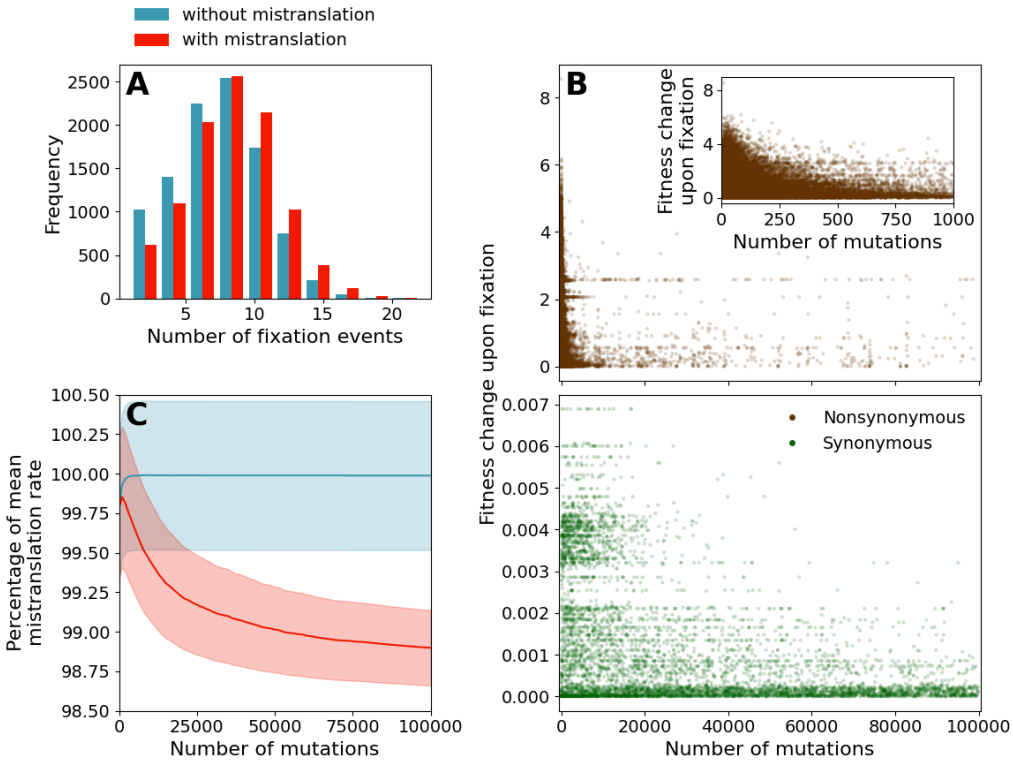

Figure 5: Mistranslation causes longer adaptive walks due to selection against high mistranslation rates acting on synonymous mutations. A) Distributions of the number of mutations that reach fixation during adaptive walks with (red) and without mistranslation (blue). B) Fitness effects (vertical axis) of all nonsynonymous (brown, upper panel) and synonymous (green, lower panel) mutations fixed during adaptive walks with mistranslation. The horizontal axis shows the number of mutations (both extinct and fixed) since the beginning of the adaptive walk. Inset: Fitness effects of nonsynonymous mutations that reached fixation within the first 1000 mutations of the adaptive walk. C) Mean and standard error (lines and shaded areas, respectively) of the mistranslation rates (red) and post hoc mistranslation rates (blue) of the genotypes encountered during adaptive walks with and without mistranslation rates, respectively. (Post hoc) mistranslation rates are shown as a percentage of the mean mistranslation rate of all genotypes in the antibody-binding landscape (100 percent is equivalent to the mean mistranslation rate). Results are shown for 10^4 adaptive walks on the antibody-binding landscape at high expression (500 proteins per cell) and a population size of $N = 10^8$.

only few of a walks 10^5 mutations go to fixation (mean number of fixations: 8.5±3.1 vs 7.8±3.1, figure 5a, antibody-binding landscape, $N = 10^8$, high expression, Wilcoxon test $z = 15703133$, $p = 1 \times 10^{-73}$, $n = 10^4$). These differences come from an increase in the number of synonymous mutational steps in adaptive walks with mistranslation. For example, on the antibody-binding landscape, without mistranslation and a population size of $N = 10^8$, we observed on average only 3.0×10^{-4} ($\pm 1.7 \times 10^{-2}$) synonymous and 7.7 ($\pm$ 3.1) nonsynonymous mutations per adaptive walk. In contrast, with mistranslation at high expression we observed on average 0.6 ($\pm$ 0.7) synonymous and 7.9 ($\pm$ 3.0) nonsynonymous mutations per walk. In other words, mistranslation helps increase the number of synonymous substitutions by three orders of magnitude.

A second test of our hypothesis is based on its prediction that some synonymous mutations come under positive selection only under mistranslation. To detect this selection pressure we considered what mistranslation rates would have been in genotypes that are traversed by an adaptive walk without mistranslation. We call these inferred mistranslation rates post hoc mistranslation rates. By comparing the true and the post hoc mistranslation rates, we wanted to find out whether genotypes encountered by adaptive walks with mistranslation have different mistranslation rates than genotypes encountered by adaptive walks without mistranslation.

Indeed, while many of the synonymous mutations that reached fixation during adaptive walks with mistranslation had very weak fitness effects (both beneficial and deleteri-

ous), some had large fitness effects (figure 5b). Such mutations tend to appear after at least 10^3 mutations have occurred. They increase fitness by switching to synonymous codons with lower mistranslation rates and therefore higher fitness. For a population of $N = 10^8$ individuals and with high expression, every synonymous mutation with a fitness effect greater than 0.002 reduced mistranslation rates on average by 4%. Synonymous mutations with larger fitness effects also decreased mistranslation rates more strongly (Kendall's $\tau = -0.554$, $p = 0.0$ $n = 6000$), which is not the case for nonsynonymous mutations (Kendall's tau 2.15×10^{-3}, $p = 0.366$, $n = 78686$). Consequently, in adaptive walks with mistranslation, we observed a long-term decrease in mistranslation rates that is largely (but not entirely, see supplementary materials section 6) absent from post hoc mistranslation rates in walks without mistranslation (figure 5c). A similar pattern also holds for the two toxin-antitoxin landscapes, except that there nonsynonymous mutations also decreased mistranslation rates. For example, in the toxin-antitoxin (E3) landscape, nonsynonymous mutations with larger fitness effects also decrease mistranslation rates (Kendall's $\tau = -0.0521$, $p = 1.08 \times 10^{-95}$, $n = 70660$), although the association is an order of magnitude stronger for synonymous mutations (Kendall's $\tau = -0.523$, $p = 2.91 \times 10^{-248}$, $n = 1874$ both correlations $N = 10^8$ and high expression). In the toxin-antitoxin (E2) landscape, nonsynonymous mutations decrease mistranslation rates weakly (Kendall's $\tau = -0.0343$, $p = 1.56 \times 10^{-44}$, $n = 73982$), and the association is stronger for synonymous mutations (Kendall's $\tau = -0.558$, $p = 1.20 \times 10^{-250}$, $n = 1646$

both correlations $N = 10^8$ and high expression).

We also found evidence that the fixation of synonymous mutations is not only in itself adaptive, but that it can facilitate adaptation by increasing the supply of beneficial nonsynonymous mutations. For example, in the antibody-binding landscape for populations of $N = 10^4$ individuals and with mistranslation at low expression, synonymous mutations increased the average number of nonsynonymous genetic neighbours with higher fitness than the current genotype from 5.30×10^{-2} before fixation of the synonymous mutation to 1.83×10^{-1} afterwards (Wilcoxon test $z = 7928.5$, $p = 2.78 \times 10^{-223}$, $n = 9584$). This effect is particularly strong once populations have reached high fitness and the supply of beneficial nonsynonymous mutations is limited. In adaptive walks after the first 10^4 mutations, the average number of beneficial nonsynonymous mutations increases by 96.9 percent from 3.75×10^{-3} before the fixation of a synonymous mutation to 1.20×10^{-1} afterwards (Wilcoxon test $z = 2135$, $p = 1.20 \times 10^{-186}$, $n = 8787$). A similar pattern holds for the other landscapes and in adaptive walks without mistranslation (see supplementary materials table 4). Consequently, adaptive walks that fix few synonymous mutations, such as walks at high population sizes and without mistranslation, experience less adaptive change at high fitnesses.

Discussion

Because the phenotypic mutations caused by mistranslation are much more frequent than DNA mutations, they represent an important source of variation that may affect adaptive evolution despite its low heritability. To find out how mistranslation may affect evolution, we used experimentally measured mistranslation rates to study its effects on the topography of experimentally characterised adaptive landscapes, and how populations evolve on such landscapes. We found that mistranslation broadly decreases the fitness of high fitness genotypes and increases the fitness of low fitness genotypes, leading to a flattening of the adaptive landscape. However, mistranslation does not affect all genotypes equally. It therefore changes the fitness of some genotypes relative to others. The most dramatic effect of mistranslation is a large decrease in the number of mutations that are nearly neutral, i.e. mutations whose impact on fitness is so small that they are effectively invisible to selection (Ohta, 1973). More than half of these mutations are synonymous. We predict that mistranslation causes small fitness differences between synonymous mutations. Most of those fitness differences are orders of magnitude smaller than the differences caused by nonsynonymous mutations, but they are still large enough to become subject to selection in species with large effective population sizes (Sung et al., 2012).

Previous theoretical work (Rocabert et al., 2020; Wang and Zhang, 2011; Whitehead et al., 2008) has established that nonheritable variation in fitness can sometimes favour

the fixation of beneficial mutations. However, it is not clear how frequently this occurs. On the landscapes we study, mistranslation increases the fixation probability of roughly three to eight percent of beneficial mutations. The reason is that mistranslation results in the production of protein variants that are more beneficial than the variant genetically encoded by the mutant. This advantage comes at the cost of decreasing the fixation probability of many other beneficial mutations. However, considering experimental studies that demonstrate the fitness benefits of mistranslation in populations under stress (Javid et al., 2014; Mohler and Ibba, 2017; Ribas de Pouplana et al., 2014; Samhita et al., 2021) we suggest that mistranslation may often help the fixation of beneficial mutations in nature. We also found that mistranslation renders a small fraction of beneficial mutations deleterious, or vice versa (roughly 2 percent both ways). Thus, mistranslation can open new paths to higher fitness, and close others.

Our results also suggest that mistranslation helps evolving populations reach high fitness regions in rugged landscapes that are difficult to navigate, but conveys little to no benefit on landscapes where high fitness genotypes are easily found. Specifically, in adaptive walks on the antibody-binding landscape, where reaching high fitness genotypes is more difficult than on the toxin-antitoxin landscapes, populations with mistranslation reached on average a slightly higher fitness than populations without mistranslation. These advantages are small in part because the mistranslation rates we used in our study come from a population of *E. coli* growing in a favourable environment (Mordret et al.,

2019). However, mistranslation rates increase in stressful conditons (Netzer et al., 2009), and consequently the benefits we report here may be larger in nature (Samhita et al., 2021).

Our analysis also supports a prediction by Rocabert *et al.* (Rocabert et al., 2020) that nonheritable variation can accelerate adaptation not only at low fitness, but also at high fitness. We found two causes for this phenomenon. First, selection favours the reduction of mistranslation rates at high fitness, because mistranslation is typically detrimental for high fitness genotypes. This reduction occurs predominantly through synonymous mutations. This cause is consistent with the experimental observation that selection favours synonymous substitutions that decrease mistranslation rates (Akashi, 1994; Drummond and Wilke, 2008; Porceddu et al., 2013; Zaborske et al., 2014). Second, the fixation of synonymous mutations can open new paths to even higher fitness. Once a population reaches high fitness, some synonymous mutations can act as potentiating mutations (Blount et al., 2008), i.e. mutations that themselves have a small or no impact on fitness but which increase the supply of beneficial nonsynonymous mutations. Mistranslation increases the probability that synonymous mutations become fixed, and therefore make the discovery of these beneficial nonsynonymous mutations easier.

As a previous theoretical study on mistranslation has also observed (Drummond and Wilke, 2008), our results suggest that mistranslation can render the ratio of nonsynonymous to synonymous mutations (dN/dS) unreliable as an indicator of the absence of

positive selection. More precisely, a large dN/dS ratio indicates directional selection only early during adaptive evolution, when a population's fitness increases dramatically through nonsynonymous mutations. This is no longer the case later during adaptive evolution, when most genotypes in a population have attained high fitness. The reason is that mistranslation is deleterious for high fitness genotypes, and selection favours synonymous mutations that decrease mistranslation rates. Since the dN/dS ratio is designed to report selection on nonsynonymous mutations, it thus under-reports the extent of directional selection.

Like any other computational analysis, ours is limited by inevitable simplifying assumptions. First, we assumed that without mistranslation, synonymous mutations are neutral. That is not necessarily the case, because synonymous DNA mutations can alter mRNA stability (Chamary and Hurst, 2005; Duan et al., 2003; Kristofich et al., 2018) or ribosome binding (Eyre-Walker and Bulmer, 1993; Knöppel et al., 2016). One consequence of this assumption is that we may overestimate by how much mistranslation reduces the number of nearly neutral mutations, because many synonymous mutations may not be neutral to begin with. Second, our analysis requires a large amount of data. The few available sufficiently large empirical landscapes may not be representative of protein adaptive landscapes in general. This limitation will be alleviated as more and larger landscapes will be characterised in the future.

Third, the effects of mistransltion are more complicated in nature than in our model.

For example, they may be buffered by metabolism (Kacser and Burns, 1981), changes in gene expression (Bratulic et al., 2015), preferential degradation of misfolded proteins (Mohler and Ibba, 2017), and DNA mutations that increase protein stability (Bratulic et al., 2017; Zheng et al., 2021). Conversely, they may also be enhanced, for example through protein aggregation (Bucciantini et al., 2002; Drummond and Wilke, 2008). These complicating factors are an exciting research opportunity to develop more realistic models. Because the phenomena we study are difficult to observe directly by experiment, such data-driven models will remain essential to study how mistranslation changes the evolutionary dynamics on adaptive landscapes.

Methods

Mistranslation rate estimates

Following the literature in the field (Bratulic et al., 2015; Bürger et al., 2006; Yanagida et al., 2015), we define a phenotypic mutation as the mistranslation of a specific codon into a specific (wrong) amino acid. We used experimentally estimated mistranslation rates from a previous study relying on proteome-wide mass-spectrometry in *E. coli* (Mordret et al., 2019). The study computed a ratio of mistranslated to correctly translated proteins. This "intensity ratio" provides a direct estimate of the mistranslation rate per codon, which is reliable for common (but not for rare) mistranslation errors (2019).

As an estimate for the rate at which any one phenotypic mutation occurs, we used the median intensity ratios of each phenotypic mutation for the wild-type *E. coli* strain growing in MOPS (3-(N-morpholino)propanesulfonic acid) complete medium during mid-exponential growth. As one might expect, these rate estimates decrease with the number of times any one phenotypic mutation is observed, except for rare phenotypic mutations, where mistranslation rate estimates *increase* as events become more rare, probably due to sampling error (see supplementary materials figure 4a). Because direct rate estimates are therefore unreliable for rare phenotypic mutations, we did not use such estimates in our analysis. Instead, we estimated the rates of rare events through linear regression on the relationship between the number of times a phenotypic mutation was observed and the $\log_{10}$-transformed mistranslation rate estimates of common phenotypic mutations (supplementary materials figure 4a). After investigating multiple thresholds for the exclusion of direct rate estimates, we chose to exclude all mistranslation rate estimates with fewer than 18 observations, because this threshold minimised the intercept of the linear regression (supplementary materials figure 4b), and therefore minimised our estimate of the mistranslation rate for phenotypic mutations too rare to be observed in the empirical data (2019). This threshold excluded 86% of phenotypic mutations, for which we used indirect regression estimates. We assigned unobserved phenotypic mutations the rate of 2.7×10^{-4} per translated codon, the intercept of the regression line (supplementary materials figure 4a).

Landscape data

We drew on three experimentally determined landscapes for all our analyses, as described in more detail below. The first of these is an antibody-binding landscape (Wu et al., 2016). The second and third are toxin-antitoxin fitness landscapes that quantify the fitness of *E. coli* cells in which variants of an antitoxin bind to a toxin protein (Lite et al., 2020). All landscapes are represented at the level of amino acids, because the nucleotide sequences encoding these amino acids have not been reported (2020; Wu et al., 2016). We simulate evolution at the level of nucleotide sequences, so we map each nucleotide sequence to the amino acid it encodes. We chose these landscapes for two reasons. First, they contain phenotypic data for all 20 amino acids at each of the protein sites for which amino acid variation was examined. Second, they also contain data for all possible combinations of the 20 amino acids. In other words, these landscapes are combinatorially complete – they comprise data on all 20^L genotypes, where L is the number of protein sites (2016). Because mistranslation can cause phenotypic mutations from any amino acid to any other amino acid (Mordret et al., 2019), both characteristics are important for our purpose. We note that mistranslation inevitably affects these experimentally determined landscapes (albeit weakly, see supplementary materials section 7). Although it is not possible to correct for this influence, we work with this data, because it allows us to (i) study adaptation on an empirical landscape, and (ii) study the effect of mistranslation on adaptation on such a landscape.

The antibody-binding landscape (Wu et al., 2016) comprises *in vitro* data on the affinity of sequence variants of the protein GB1 to bind the antibody IgG-Fc. Specifically, it contains such data for all possible 20 amino acid variants at four sites in GB1, which is a streptococcal protein capable of binding immunoglobulins. Briefly, the authors of the pertinent study (2016) created a mutant library encoding nearly all 160,000 GB1 sequence variants at the four amino acid positions, translated the library, and tagged the resulting proteins with the mRNA encoding them in order to identify the proteins later. They then incubated these tagged proteins with immunoglobulin-bearing beads, and washed away all non-bound proteins. They released the bound proteins and sequenced their mRNA tags. By comparing the frequencies of mRNA tags before and after exposure to immunoglobulin, they determined the binding activity of the GB1 mutants relative to the wild-type sequence. While this landscape is strictly speaking no fitness landscape, we follow other authors that have used a phenotype such as molecular binding data as proxies for fitness (Aguilar-Rodríguez et al., 2017; Natarajan et al., 2013; Olson et al., 2014; Sarkisyan et al., 2016). We thus tacitly assume that natural selection can favour strong molecular binding. The 160,000 amino acid sequences are encoded by more than 1.3×10^7 nucleotide sequences.

The two toxin-antitoxin landscapes (Lite et al., 2020) comprise fitness measures for almost 8,000 variants of the ParD3 antitoxin, a protein that is co-expressed with the toxin ParE3, which it binds and inhibits. To generate this data, amino acids at each of

three sites in ParD3 were systematically replaced with all 19 other amino acids, and the resulting sequences were co-expressed in *E. coli* either with ParE3 (landscape 1) or its close homolog ParE2 (landscape 2 (2020)). The fitness of *E. coli* cells expressing these variants was measured in mass selection experiments combined with deep sequencing, by measuring the change in frequency of each variant before and after selection for toxin-antitoxin binding (2020). Fitness values were transformed such that the most fit ParD3 variant was assigned a fitness of one, and variants of ParD3 with a nonsense mutation were assigned a fitness of zero. This transformation caused 8% and 17% of variants in the of the toxin-antitoxin (E3) and the toxin-antitoxin (E2) landscape, respectively, to have a fitness below zero. For the purpose of our analysis, we set all negative fitness values to zero, reasoning that at least some of them may have resulted from measurement error. The 8,000 amino acid sequences are encoded by about 2.2×10^5 nucleotide sequences.

Simulating the fitness effects of mistranslation

Our simulations start from a given genotype, that is, a nucleotide sequence encoding the part of the focal protein for which fitness data on protein variants is available (2020; Wu et al., 2016). Each codon in this nucleotide sequence has a given probability of being translated into either the cognate amino acid or into one of the 19 noncognate amino acids at a rate of mistranslation that is experimentally determined (Mordret et al., 2019). We assume that the probability of correct translation is one minus the sum of all

mistranslation rates. For every polypeptide sequence of length L encoded by a nucleotide sequence of length $3L$, (mis)translation produces from this nucleotide sequence one of $k = 20^L$ protein variants with a unique polypeptide sequence.

Any one cell that expresses the focal protein will harbour multiple variants of this protein, one resulting from correct translation, and all others from mistranslation. We estimate the number of each mistranslated variant in the total pool of n copies of the focal protein in the cell through a multinomial distribution with n "trials". Each trial corresponds to the biosynthesis of a protein, and can have one of k possible outcomes – a protein variant. Each protein variant i (with $i = 1, \ldots, k$) has its own probability of being produced p_i, which is determined by the mistranslation rates of the encoding nucleotide sequence. We assume that the mistranslation rate at each codon is independent from that at other codons, because experimental measurements suggest that most translation errors are due to codon-anticodon mispairing in the ribosome (Kramer and Farabaugh, 2007; Mordret et al., 2019). Consequently, the probability p_i of producing a given protein variant i is the product of the probabilities that translation incorporates the amino acids that make up the protein's sequence, given the encoding nucleotide sequence. More specifically, the multinomial distribution for the number of copies x_i of each protein

variant $i,\ldots,k$ is given by

$$f(x_1,\ldots,x_k) = \begin{cases} \frac{n!}{x_1!\ldots x_k!} p_1^{x_1} \times \cdots \times p_k^{x_k}, & \text{if } \sum_{i=1}^{k} x_i = n \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

A single cell producing n copies of the focal protein will contain a tiny fraction of all possible protein variants k. However, large populations of such cells will produce a larger fraction of variants. Over multiple generations or evolutionary time-scales, even rare protein variants will appear. Consequently, for the purpose of deriving our model we assume that all protein variants k may be produced from a given nucleotide sequence. In practice, some of these variants will have such low probability of being produced through mistranslation that they have no influence on the fitness or evolution of a population, and in our simulations we ignore these variants (see below).

With this multinomial distribution we then calculate the expected number of copies of each protein variant, together with the variation in the number of copies of a given variant and the covariation in the number of copies between different protein variants. We need all three quantities to estimate the fitness distribution of a population of cells under the influence of mistranslation. The expected number of copies $E(x_i)$ of protein variant i produced in a cell is

$$E(x_i) = np_i \tag{2}$$

and its variance is

$$Var(x_i) = np_i(1 - p_i). \tag{3}$$

We calculate these two quantities for the number of copies of each possible protein variant $x_1, \ldots, x_k$. Because the number of proteins produced per cell is finite, the production of a copy of one protein variant decreases the probability that a cell will contain a copy of another protein variant. Consequently, the covariance between the number of copies of different variants is negative. Specifically, the covariance of the number of copies x_i of variant i with the number of copies x_j of another variant j is

$$Cov(x_i, x_j) = -np_i p_j,\ i \neq j. \tag{4}$$

Using the multinomial distribution of a cell's protein composition together with experimental fitness measurements we then determine the fitness distribution f of a population of cells expressing the protein. Specifically, we aim to estimate the mean and variance of the fitness distribution, because these quantities are needed later to estimate the fixation probability of mutations. The fitness of an individual cell depends upon the kind of protein variants it produces. For simplicity, we assume that every protein copy contributes equally to fitness, and that the fitness contribution w_i of copy i is independent of that of all other copies produced. The fitness of a cell then becomes the mean of the fitness contributions of all copies it produces. In consequence, the expected fitness of a

cell $E(f)$ is given by

$$E(f) = \frac{1}{n}\sum_{i=1}^{k} n \cdot p_i \cdot w_i \tag{5}$$

where n is the total number of proteins per cell, k the number of all possible protein variants, p_i the probability of translation producing protein variant i, and w_i the fitness contribution of that protein variant. As indicated in this expression for $E(f)$, we calculate the expected (mean) fitness of a cell by summing over all the fitness contributions w_i of each possible protein variant (k in total), with variants that are more frequent having a greater impact, and then take the average by dividing this sum by the number of copies per cell n.

To arrive at an expression for the variance in fitness among cells, we start by observing in equation 5 that the expression for the expected fitness of a cell is a sum over the expected number of protein copies $n \cdot p_i$ for each protein variant i, multiplied by the term w_i/n. With this observation in mind, we compute the fitness variance by drawing on three well-known statistical principles. First, the variance of a random variable multiplied by a constant a is $Var(a \cdot X) = a^2 \cdot Var(X)$. Consequently, every protein variant i contributes to the variance in fitness according to the expression $Var(w_i/n \cdot x_i) = (w_i/n)^2 \cdot np_i(1-p_i)$. Second, the covariance of two random variables multiplied by the constants a and b is given by $Cov(a \cdot X, b \cdot Y) = a \cdot b \cdot Cov(X,Y)$. Consequently, the covariance in the fitness contributions of two protein variants i and j becomes $Cov(w_i/n \cdot x_i, w_j/n \cdot x_j) = -n \cdot p_i \cdot p_j \cdot w_i/n \cdot w_j/n$. Third, the variance of the sum of two random variables X and

Y is $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$. Consequently, we can sum over the variances of the fitness contributions of all k protein variants and adjust the resulting quantity by their pairwise (negative) covariances. Putting all these considerations together, the expected variance in the fitness of cells becomes

$$Var(f) = \frac{1}{n}\sum_{i=1}^{k}\left[p_i \cdot (1-p_i) \cdot w_i^2 - \sum_{j=i+1}^{k} 2 \cdot p_i \cdot p_j \cdot w_i \cdot w_j\right]. \tag{6}$$

During our simulations we saved computational resources by ignoring the contributions of all protein variants whose probability of being produced by translation was less than 10^{-9} when computing the mean fitness of cells and its variance.

Estimating the fixation probability of novel mutations

Given an estimate for the mean and variance in fitness for both a wild-type genotype *wt* and a single-nucleotide DNA mutant *mt* of the wild-type, we here estimate the probability of fixation of the mutant genotype if it is introduced into a population that consists of only wild-type genotypes. To do so, we need three quantities. The first one is the selection coefficient s, which we defined as the relative difference between the expected fitness values of the invading mutant $E(f_{mt})$ and of the wild type $E(f_{wt})$:

$$s = E(f_{mt})/E(f_{wt}) - 1. \tag{7}$$

The second quantity is the effective population size N_e. Mistranslation can decrease the power of selection by adding non-genetic variation in fitness. As a result, in a population under the influence of mistranslation the amount of variation in allele frequencies from one generation to the next is larger than in a population without mistranslation. Previous authors (Wang and Zhang, 2011) have modelled this increase in the variation of allele frequencies between generations as being equivalent to a decrease in the effective population size of a haploid population from N to N_e according to

$$N_e = N/(1+e^2) \tag{8}$$

where N is the population size in the absence of mistranslation and e^2 is the squared coefficient of variation in fitness ($Var(f_{wt})/E(f_{wt})^2$). We follow this approach. The new effective population size N_e and the selection coefficient s can then be used to estimate the probability of fixation of the invading mutation, where the initial allele frequency is $p = 1/N$. (We note that the expression $p = 1/N_e$ would be incorrect here, because the initial allele frequency is not determined by the effective population size N_e but by the actual number of individuals N). The third quantity is the fixation probability u_{fix}. We use Kimura's equation (Kimura, 1962) to estimate the fixation probability of a new mutant in a haploid population:

$$u_{fix} = \frac{1-e^{-2N_e sp}}{1-e^{-2N_e s}} \tag{9}$$

Simulating adaptive walks on an adaptive landscape

We simulated adaptive evolution on our three landscapes at the level of nucleotide sequences. However, the landscapes reported their measurements at the amino acid sequence level. We therefore made the simplifying assumption that synonymous mutations are neutral in the absence of mistranslation.

We simulated the occurrence and either fixation or loss of mutations, and not the gradual change in allele frequency of a population from one generation to the next. Behind this procedure lies the assumption that the evolutionary dynamics on our landscapes fall into the weak mutation-strong selection regime (Gillespie, 1983; 1984). The reason is that we simulate evolution at a small number of nucleotide sites, where it is unlikely that multiple alleles will be segregating in a population at these sites at any one time.

During each step of our simulations, we chose a one-step mutational neighbour of the current (wild-type) genotype at random, and evaluated its probability of fixation. Based on this probability, the mutation either reached fixation or was lost at random. We continued this process for 10^5 successive mutations per walk, and note that only a small fraction of these mutations reach fixation. We performed 10^4 replicate adaptive walks for each evolutionary scenario we modelled, each with its own starting genotype chosen at random from among 10% of genotypes with lowest fitness.

Counting the number of fitness peaks and the incidence of epistasis

In our analysis, a fitness peak can either consist of a single genotype or a network of multiple genotypes that are connected by single mutations, and in which genotypes that differ in a single mutation are nearly neutral (Ohta, 1973), i.e. mutations with such small fitness effects that they effectively evolve by genetic drift. We define a nearly neutral mutation as a mutation with a selection coefficient $|s| < \frac{1}{4N_e}$ relative to the wild-type (Charlesworth, 2009). We consider such a genotype or network of genotypes a peak if all its single-mutation, non-neutral neighbours have lower fitness.

To search for fitness peaks in our three landscapes, we chose a set of 10^5 genotypes with the highest fitness as seeds for the search. For each seed, we examined whether every one of the one-step mutational neighbours of the seed is nearly neutral or not. We tested for neutrality both going from one genotype to another and in reverse, as the effective population size and selection coefficient can differ depending on the starting genotype. The reason is that genotypes may differ in the extent of between-individual fitness variation caused by mistranslation. We only considered a neighbouring genotype nearly neutral if mutations in both directions were nearly neutral. Sequences that matched this criterion are members of the same nearly neutral network of genotypes. We then assessed the near neutrality of *their* neighbours, repeating this process until we could not find any further nearly neutral neighbours. If one of the genotypes we chose as seeds was found to be a member of a previously explored nearly neutral network we continued

on to another seed. During each step of this iterative search, we also kept track of the current nearly neutral network's non-neutral mutational neighbours. If all the non-neutral neighbours of each genotype in the network had lower fitness than the genotype itself, we considered the set of genotypes in the network to constitute a fitness peak.

We estimated the incidence of epistasis by sampling 10^4 genotypes from each of our landscapes. For each of these starting genotypes, we created a "square" of four genotypes (see also supplementary materials figure 5). To do so, we chose two sites at random in the starting genotype (which we denote as 00) and replaced either one or both sites with alternative nucleotides, generating an additional three nucleotide sequences for each starting genotype. Two of these sequences are single mutants (01 and 10), and one a double mutant (11). We quantify any deviation from additivity by using the linear combination of the fitness w of each genotype (Srivastava and Payne, 2022):

$$\varepsilon = w_{00} + w_{11} - w_{01} - w_{10}. \tag{10}$$

Any square that deviates from additivity ($\varepsilon \neq 0$) is epistatic. To differentiate between magnitude epistasis, simple sign epistasis, or reciprocal epistasis, we assessed whether a given mutation causes the same sign change (positive or negative) in fitness in different genetic backgrounds, specifically the starting genotype (00) and the single mutant for the other site. If the background did not affect the sign change of neither mutation, we classified the square as a case of magnitude epistasis. If the background caused a

sign reversal of one or both mutations, we classified the square as a case of simple sign epistasis or reciprocal sign epistasis, respectively.

References

Ackermann, M. (2015). A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Reviews Microbiology*, *13*(8), 497–508.

Aguilar-Rodríguez, J., Payne, J. L., & Wagner, A. (2017). A thousand empirical adaptive landscapes and their navigability. *Nature Ecology & Evolution*, *1*(2), 1–9.

Aguirre, J., Catalán, P., Cuesta, J. A., & Manrubia, S. (2018). On the networked architecture of genotype spaces and its critical effects on molecular evolution. *Open Biology*, *8*(7), 180069.

Akashi, H. (1994). Synonymous Codon Usage in Drosophila Melanogaster: Natural Selection and Translational Accuracy. *Genetics*, *136*(3), 927–935.

Ancel, L. W., & Fontana, W. (2000). Plasticity, evolvability, and modularity in RNA. *Journal of Experimental Zoology*, *288*(3), 242–283.

Bataillon, T., Ezard, T. H. G., Kopp, M., & Masel, J. (2022). Genetics of adaptation and fitness landscapes: From toy models to testable quantitative predictions. *Evolution*, *76*(6), 1104–1107.

Blanco, C., Janzen, E., Pressman, A., Saha, R., & Chen, I. A. (2019). Molecular Fitness Landscapes from High-Coverage Sequence Profiling. *Annual Review of Biophysics*, *48*(1), 1–18.

Blount, Z. D., Borland, C. Z., & Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. *Proceedings of the National Academy of Sciences*, *105*(23), 7899–7906.

Bódi, Z., Farkas, Z., Nevozhay, D., Kalapis, D., Lázár, V., Csörgő, B., Nyerges, Á., Szamecz, B., Fekete, G., Papp, B., Araújo, H., Oliveira, J. L., Moura, G., Santos, M. A. S., Jr, T. S., Balázsi, G., & Pál, C. (2017). Phenotypic heterogeneity promotes adaptive evolution. *PLOS Biology*, *15*(5), e2000644.

Bonduriansky, R., & Day, T. (2018). *Extended heredity: A new understanding of inheritance and evolution*. Princeton University Press.

Bratulic, S., Gerber, F., & Wagner, A. (2015). Mistranslation drives the evolution of robustness in TEM-1 β-lactamase. *Proceedings of the National Academy of Sciences*, *112*(41), 12758–12763.

Bratulic, S., Toll-Riera, M., & Wagner, A. (2017). Mistranslation can enhance fitness through purging of deleterious mutations. *Nature Communications*, *8*(1), 15410.

Bruggeman, F. J., & Teusink, B. (2018). Living with noise: On the propagation of noise from molecules to phenotype and fitness. *Current Opinion in Systems Biology*, *8*, 144–150.

Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C. M., & Stefani, M. (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, *416*(6880), 507–511.

Bürger, R., Willensdorfer, M., & Nowak, M. A. (2006). Why Are Phenotypic Mutation Rates Much Higher Than Genotypic Mutation Rates? *Genetics*, *172*(1), 197–206.

Carey, J. N., Mettert, E. L., Roggiani, M., Myers, K. S., Kiley, P. J., & Goulian, M. (2018). Regulated Stochasticity in a Bacterial Signaling Network Permits Tolerance to a Rapid Environmental Change. *Cell*, *173*(1), 196–207.e14.

Chamary, J., & Hurst, L. D. (2005). Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. *Genome Biology*, *6*(9), R75.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195–205.

Chen, T., Vernazobres, D., Yomo, T., Bornberg-Bauer, E., & Chan, H. S. (2010). Evolvability and Single-Genotype Fluctuation in Phenotypic Properties: A Simple Heteropolymer Model. *Biophysical Journal*, *98*(11), 2487–2496.

Draghi, J. (2018). Links between evolutionary processes and phenotypic robustness in microbes. *Seminars in Cell & Developmental Biology*, *88*, 46–53.

Drummond, D. A., & Wilke, C. O. (2008). Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution. *Cell*, *134*(2), 341–352.

Duan, J., Wainwright, M. S., Comeron, J. M., Saitou, N., Sanders, A. R., Gelernter, J., & Gejman, P. V. (2003). Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Human Molecular Genetics*, *12*(3), 205–216.

Ellis, N., & Gallant, J. (1982). An estimate of the global error frequency in translation. *Molecular and General Genetics MGG*, *188*(2), 169–172.

Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic Gene Expression in a Single Cell. *Science*, *297*(5584), 1183–1186.

Espinosa-Soto, C., Martin, O. C., & Wagner, A. (2011). Phenotypic plasticity can facilitate adaptive evolution in gene regulatory circuits. *BMC Evolutionary Biology*, *11*(1), 5.

Eyre-Walker, A., & Bulmer, M. (1993). Reduced synonymous substitution rate at the start of enterobacterial genes. *Nucleic Acids Research*, *21*(19), 4599–4603.

Ferretti, L., Weinreich, D., Tajima, F., & Achaz, G. (2018). Evolutionary constraints in fitness landscapes. *Heredity*, *121*(5), 466–481.

Frank, S. A. (2011). Natural selection. II. Developmental variability and evolutionary rate*. *Journal of Evolutionary Biology*, *24*(11), 2310–2320.

Giacomelli, M. G., Hancock, A. S., & Masel, J. (2007). The Conversion of 3' UTRs into Coding Regions. *Molecular Biology and Evolution*, *24*(2), 457–464.

Gillespie, J. H. (1983). A simple stochastic gene substitution model. *Theoretical Population Biology*, *23*(2), 202–215.

Gillespie, J. H. (1984). Molecular Evolution Over the Mutational Landscape. *Evolution*, *38*(5), 1116–1129.

Govers, S. K., Adam, A., Blockeel, H., & Aertsen, A. (2017). Rapid phenotypic individualization of bacterial sister cells. *Scientific Reports*, *7*(1), 8473.

Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F. U., Kerner, M. J., & Frishman, D. (2008). Protein abundance profiling of the Escherichia coli cytosol. *BMC Genomics*, *9*(1), 102.

Javid, B., Sorrentino, F., Toosky, M., Zheng, W., Pinkham, J. T., Jain, N., Pan, M., Deighan, P., & Rubin, E. J. (2014). Mycobacterial mistranslation is necessary and sufficient for rifampicin phenotypic resistance. *Proceedings of the National Academy of Sciences*, *111*(3), 1132–1137.

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of Escherichia coli gyrase. *Molecular Microbiology*, *44*(4), 971–979.

Kacser, H., & Burns, J. A. (1981). The Molecular Basis of Dominance. *Genetics*, *97*(3-4), 639–666.

Kimura, M. (1962). On the Probability of Fixation of Mutant Genes in a Population. *Genetics*, *47*(6), 713–719.

Knöppel, A., Näsvall, J., & Andersson, D. I. (2016). Compensating the Fitness Costs of Synonymous Mutations. *Molecular Biology and Evolution*, *33*(6), 1461–1477.

Kotte, O., Volkmer, B., Radzikowski, J. L., & Heinemann, M. (2014). Phenotypic bistability in Escherichia coli's central carbon metabolism. *Molecular Systems Biology*, *10*(7), 736.

Kramer, E. B., & Farabaugh, P. J. (2007). The frequency of translational misreading errors in E. coli is largely determined by tRNA competition. *RNA*, *13*(1), 87–96.

Kristofich, J., Morgenthaler, A. B., Kinney, W. R., Ebmeier, C. C., Snyder, D. J., Old, W. M., Cooper, V. S., & Copley, S. D. (2018). Synonymous mutations make dramatic contributions to fitness when growth is limited by a weak-link enzyme. *PLOS Genetics*, *14*(8), e1007615.

Lite, T.-L. V., Grant, R. A., Nocedal, I., Littlehale, M. L., Guo, M. S., & Laub, M. T. (2020). Uncovering the basis of protein-protein interaction specificity with a combinatorially complete library. *eLife*, *9*, e60924.

Maynard Smith, J. (1970). Natural Selection and the Concept of a Protein Space. *Nature*, *225*(5232), 563–564.

McAdams, H. H., & Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolar scale. *Trends in Genetics*, *15*(2), 65–69.

Mineta, K., Matsumoto, T., Osada, N., & Araki, H. (2015). Population genetics of non-genetic traits: Evolutionary roles of stochasticity in gene expression. *Gene*, *562*(1), 16–21.

Mohler, K., & Ibba, M. (2017). Translational fidelity and mistranslation in the cellular response to stress. *Nature Microbiology*, *2*(9), 1–9.

Mordret, E., Dahan, O., Asraf, O., Rak, R., Yehonadav, A., Barnabas, G. D., Cox, J., Geiger, T., Lindner, A. B., & Pilpel, Y. (2019). Systematic Detection of Amino Acid Substitutions in Proteomes Reveals Mechanistic Basis of Ribosome Errors and Selection for Translation Fidelity. *Molecular Cell*, *75*(3), 427–441.e5.

Natarajan, C., Inoguchi, N., Weber, R. E., Fago, A., Moriyama, H., & Storz, J. F. (2013). Epistasis Among Adaptive Mutations in Deer Mouse Hemoglobin. *Science*, *340*(6138), 1324–1327.

Netzer, N., Goodenbour, J. M., David, A., Dittmar, K. A., Jones, R. B., Schneider, J. R., Boone, D., Eves, E. M., Rosner, M. R., Gibbs, J. S., Embry, A., Dolan, B., Das, S., Hickman, H. D., Berglund, P., Bennink, J. R., Yewdell, J. W., & Pan, T. (2009). Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature*, *462*(7272), 522–526.

Ohta, T. (1973). Slightly Deleterious Mutant Substitutions in Evolution. *Nature*, *246*(5428), 96–98.

Olson, C. A., Wu, N. C., & Sun, R. (2014). A Comprehensive Biophysical Description of Pairwise Epistasis throughout an Entire Protein Domain. *Current Biology*, *24*(22), 2643–2651.

Peracchi, A. (2018). The Limits of Enzyme Specificity and the Evolution of Metabolism. *Trends in Biochemical Sciences*, *43*(12), 984–996.

Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M., & Tans, S. J. (2007). Empirical fitness landscapes reveal accessible evolutionary paths. *Nature*, *445*(7126), 383–386.

Porceddu, A., Zenoni, S., & Camiolo, S. (2013). The Signatures of Selection for Translational Accuracy in Plant Genes. *Genome Biology and Evolution*, *5*(6), 1117–1126.

Rajon, E., & Masel, J. (2011). Evolution of molecular error rates and the consequences for evolvability. *Proceedings of the National Academy of Sciences*, *108*(3), 1082–1087.

Ribas de Pouplana, L., Santos, M. A. S., Zhu, J.-H., Farabaugh, P. J., & Javid, B. (2014). Protein mistranslation: Friend or foe? *Trends in Biochemical Sciences*, *39*(8), 355–362.

Rocabert, C., Beslon, G., Knibbe, C., & Bernard, S. (2020). Phenotypic noise and the cost of complexity. *Evolution*, *74*(10), 2221–2237.

Ruan, B., Palioura, S., Sabina, J., Marvin-Guy, L., Kochhar, S., LaRossa, R. A., & Söll, D. (2008). Quality control despite mistranslation caused by an ambiguous genetic code. *Proceedings of the National Academy of Sciences*, *105*(43), 16502–16507.

Samhita, L., K Raval, P., Stephenson, G., Thutupalli, S., & Agashe, D. (2021). The impact of mistranslation on phenotypic variability and fitness. *Evolution*, *75*(5), 1201–1217.

Sarkisyan, K. S., Bolotin, D. A., Meer, M. V., Usmanova, D. R., Mishin, A. S., Sharonov, G. V., Ivankov, D. N., Bozhanova, N. G., Baranov, M. S., Soylemez, O., Bogatyreva, N. S., Vlasov, P. K., Egorov, E. S., Logacheva, M. D., Kondrashov, A. S., Chudakov, D. M., Putintseva, E. V., Mamedov, I. Z., Tawfik, D. S., … Kondrashov, F. A. (2016). Local fitness landscape of the green fluorescent protein. *Nature*, *533*(7603), 397–401.

Solopova, A., van Gestel, J., Weissing, F. J., Bachmann, H., Teusink, B., Kok, J., & Kuipers, O. P. (2014). Bet-hedging during bacterial diauxic shift. *Proceedings of the National Academy of Sciences*, *111*(20), 7427–7432.

Srivastava, M., & Payne, J. L. (2022). On the incongruence of genotype-phenotype and fitness landscapes. *PLOS Computational Biology*, *18*(9), e1010524.

Sung, W., Ackerman, M. S., Miller, S. F., Doak, T. G., & Lynch, M. (2012). Drift-barrier hypothesis and mutation-rate evolution. *Proceedings of the National Academy of Sciences*, *109*(45), 18488–18492.

Tănase-Nicola, S., & ten Wolde, P. R. (2008). Regulatory Control and the Costs and Benefits of Biochemical Noise. *PLOS Computational Biology*, *4*(8), e1000125.

Tawfik, O. K., & S, D. (2010). Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annual Review of Biochemistry*, *79*(1), 471–505.

Traverse, C. C., & Ochman, H. (2016). Conserved rates and patterns of transcription errors across bacterial growth states and lifestyles. *Proceedings of the National Academy of Sciences*, *113*(12), 3311–3316.

van Boxtel, C., van Heerden, J. H., Nordholt, N., Schmidt, P., & Bruggeman, F. J. (2017). Taking chances and making mistakes: Non-genetic phenotypic heterogeneity and its consequences for surviving in dynamic environments. *Journal of The Royal Society Interface*, *14*(132), 20170141.

van Heerden, J. H., Wortel, M. T., Bruggeman, F. J., Heijnen, J. J., Bollen, Y. J. M., Planqué, R., Hulshof, J., O'Toole, T. G., Wahl, S. A., & Teusink, B. (2014). Lost in Transition: Start-Up of Glycolysis Yields Subpopulations of Nongrowing Cells. *Science*, *343*(6174), 1245114.

Wagner, A. (2014). Mutational Robustness Accelerates the Origin of Novel RNA Phenotypes through Phenotypic Plasticity. *Biophysical Journal*, *106*(4), 955–965.

Wang, Z., & Zhang, J. (2011). Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proceedings of the National Academy of Sciences*, *108*(16), E67–E76.

Whitehead, D. J., Wilke, C. O., Vernazobres, D., & Bornberg-Bauer, E. (2008). The look-ahead effect of phenotypic mutations. *Biology Direct*, *3*(1), 18.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. In D. F. Jones (Ed.), *Proceedings of the Sixth International Congress of Genetics* (pp. 356–366).

Wu, N. C., Dai, L., Olson, C. A., Lloyd-Smith, J. O., & Sun, R. (2016). Adaptation in protein fitness landscapes is facilitated by indirect paths. *eLife*, *5*, e16965.

Yanagida, H., Gispan, A., Kadouri, N., Rozen, S., Sharon, M., Barkai, N., & Tawfik, D. S. (2015). The Evolutionary Potential of Phenotypic Mutations. *PLOS Genetics*, *11*(8), e1005445.

Yuan, J., Sterckx, Y., Mitchenall, L. A., Maxwell, A., Loris, R., & Waldor, M. K. (2010). Vibrio cholerae ParE2 Poisons DNA Gyrase via a Mechanism Distinct from Other Gyrase Inhibitors*. *Journal of Biological Chemistry*, *285*(51), 40397–40408.

Zaborske, J. M., DuMont, V. L. B., Wallace, E. W. J., Pan, T., Aquadro, C. F., & Drummond, D. A. (2014). A nutrient-driven tRNA modification alters translational fidelity and genome-wide protein coding across an animal genus. *PLoS biology*, *12*(12), e1002015.

Zagorski, M., Burda, Z., & Waclaw, B. (2016). Beyond the Hypercube: Evolutionary Accessibility of Fitness Landscapes with Realistic Mutational Networks. *PLOS Computational Biology*, *12*(12), e1005218.

Zhang, Z., Qian, W., & Zhang, J. (2009). Positive selection for elevated gene expression noise in yeast. *Molecular Systems Biology*, *5*(1), 299.

Zheng, J., Guo, N., & Wagner, A. (2021). Mistranslation Reduces Mutation Load in Evolving Proteins through Negative Epistasis with DNA Mutations. *Molecular Biology and Evolution*, *38*(11), 4792–4804.