Open access • Posted Content • DOI:10.1101/2021.09.27.461914

Fitness landscape analysis reveals that the wild type allele is sub-optimal and mutationally robust — Source link

Tzahi Gabzi, Yitzhak Pilpel, Tamar Friedlander

Institutions: Weizmann Institute of Science, Hebrew University of Jerusalem

Published on: 27 Sep 2021 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Mutation (genetic algorithm), Mutation rate, Fitness landscape, Evolutionary dynamics and Population

Fitness landscape analysis reveals that the wild type allele is sub-optimal and mutationally robust

Tzahi Gabzi (1), Yitzhak Pilpel (1) and Tamar Friedlander (2)
(1) Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 7610001, Israel
(2) The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture
Faculty of Agriculture, Hebrew University of Jerusalem,
P.O. Box 12 Rehovot 7610001, Israel

Correspondence: tamar.friedlander@mail.huji.ac.il, pilpel@weizmann.ac.il.

September 27, 2021

Abstract

Fitness landscape mapping and the prediction of evolutionary trajectories on these landscapes are major tasks in evolutionary biology research. Evolutionary dynamics is tightly linked to the landscape topography, but this relation is not straightforward. Models predict different evolutionary outcomes depending on mutation rates: high-fitness genotypes should dominate the population under low mutation rates and lower-fitness, mutationally robust (also called 'flat') genotypes - at higher mutation rates. Yet, so far, flat genotypes have been demonstrated in very few cases, particularly in viruses. The quantitative conditions for their emergence were studied only in simplified single-locus, two-peak landscapes. In particular, it is unclear whether within the same genome some genes can be flat while the remaining ones are fit. Here, we analyze a previously measured fitness landscape of a yeast tRNA gene. We found that the wild type allele is sub-optimal, but is mutationally robust ('flat'). Using computer simulations, we estimated the critical mutation rate in which transition from fit to flat allele should occur for a gene with such characteristics. We then used a scaling argument to extrapolate this critical mutation rate for a full genome, assuming the same mutation rate for all genes. Finally, we propose that while the majority of genes are still selected to be fittest, there are a few mutation hot-spots like the tRNA, for which the mutationally robust flat allele is favored by selection.

Introduction

Fitness landscape mapping and prediction of evolutionary trajectories of these landscapes are major tasks in evolutionary biology [1]. While evolutionary theory predicts that population mean fitness should increase over time, it offers only few quantitative predictions for the dynamics of evolution and the possible evolutionary trajectories. The main hurdle for generally computing evolutionary

trajectories is their dependence on the underlying fitness landscape. Currently available fitness landscapes include between 16 and 100,000 different genotypes (for review see [2, 3]). Yet, even the largest datasets [4, 5, 6, 7] encompass only small fractions of the entire fitness landscape of even a single gene. As detailed fitness measurements have been unavailable until recently, most of the associated theory was developed in isolation from data [8, 9, 10, 11, 12, 13, 14, 15, 16]. Additionally, the development of a general theory is difficult, because fitness landscapes are diverse and differ in details.

Evolutionary dynamics on empirical fitness landscapes was studied in cases in which genotype-phenotype mapping was available, such as folded RNA molecules [17, 18, 19] and transcription-factor binding sites [20, 21, 22] or in computational fitness landscape models closely inspired by particular experimental systems, such as maturation of the immune response [13, 23] and molecular interactions [24, 25]. Evolutionary dynamics on phenotypic fitness landscapes was studied for bacterial metabolic networks [26] and antibiotic resistance [27]. Exploration of empirical fitness landscapes and extraction of their statistical features such as local correlation, epistasis, ruggedness and density of local maxima [8, 28, 2, 6, 29, 30, 31], were pursued in the belief that these statistical hallmarks will aid in translating evolutionary trajectories to more general landscapes [32, 33].

The focus of the studies surveyed above was genotype *fitness*. Genes however are thought to evolve not only to maximize fitness, but also to reduce crosstalk [34, 35], increase network modularity [36] and allow for desired signaling properties [37, 24]. Mutational robustness - the extent to which fitness changes due to mutations - has been demonstrated to be an additional driver of evolution [38, 39, 19, 40, 41, 42, 43, 44]. The quasi-species framework developed by Manfred Eigen and Peter Schuster [45, 46, 47] is a theoretical framework that describes mutation-selection evolutionary dynamics of a large number of distinct genotypes. This framework is suitable for studying evolution of genetic sequences with a large variety of alleles, as those captured by fitness landscapes. Quasi-species theory is an extension of the simple single-locus systems studied in population genetics [48]. While the above-mentioned models mostly assumed the strong-selection-weak-mutation (SSWM) regime, in which the population is nearly monomorphic, the quasi-species framework allows for high mutation rates such that the population is polymorphic. This theory predicts a failure to adapt (so-called "error catastrophe") if the mutation rate exceeds a threshold value. In intermediate mutation rates, it predicts that populations could (depending on the landscape) favor sub-optimal but mutationally robust genotypes over the fittest ones. This "survival of the flattest" result has been shown theoretically for the simple two-peak landscape case [49, 50]. It was demonstrated in simulations of digital organisms [51] and experimentally in plant viral pathogens [52] and RNA viruses [53].

The advent in sequencing technologies now enables measurement of increasingly larger fitness landscape datasets [6, 7]. It is then desirable to predict evolutionary trajectories on these empirical fitness landscapes, using the previously developed theory in this field.

A recent set of experiments characterized the fitness landscape of the $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene of *S. cerevisiae*. As this gene is relatively short (72 nucleotides), its landscape is significantly smaller than that of a typical protein (average of 1.4 kb in *S. cerevisiae*). It is a single-copy, non-essential gene, such that many of its mutants are viable. Li *et al.* measured the growth rates of $23,284$ different mutants of this gene in four different growth conditions (23°C, 30°C, 37°C and oxidative

stress) [54, 55]. The richness of this dataset renders it a highly valuable case study for analyzing topographic properties and evolutionary trajectories of an empirical landscape and for comparing them with theoretical predictions. Here, we comprehensively analyze this tRNA fitness landscape, in efforts to identify the properties that dictate whether a particular genotype can be the "wild type", namely the extant outcome of the evolutionary dynamics. We found that the wild type was not the fittest genotype, in any of the four conditions measured, nor was it the fittest on average over all, nor a local fitness maximum. We then defined a measure of genotype local flatness with respect to its single-point mutants and found that the wild type was one of the flattest genotypes in the dataset. Stochastic evolutionary simulations over this empirical fitness landscape showed a phase transition at a threshold mutation rate, from a population dominated by a high-fitness (non wild type) genotype at low mutation rates to a collection of many intermediate-fitness genotypes composed of the wild type and other genotypes of similar fitness. To estimate the full-genome mutation rate in which this transition is expected, we used the threshold mutation rate for the tRNA alone, as obtained in the simulations, and applied a scaling argument, assuming equal properties for all loci. Variation in either local mutation rate or gene susceptibility to mutation could however cause hybrid constructs with a mixture of fit and flat genes in the same genome.

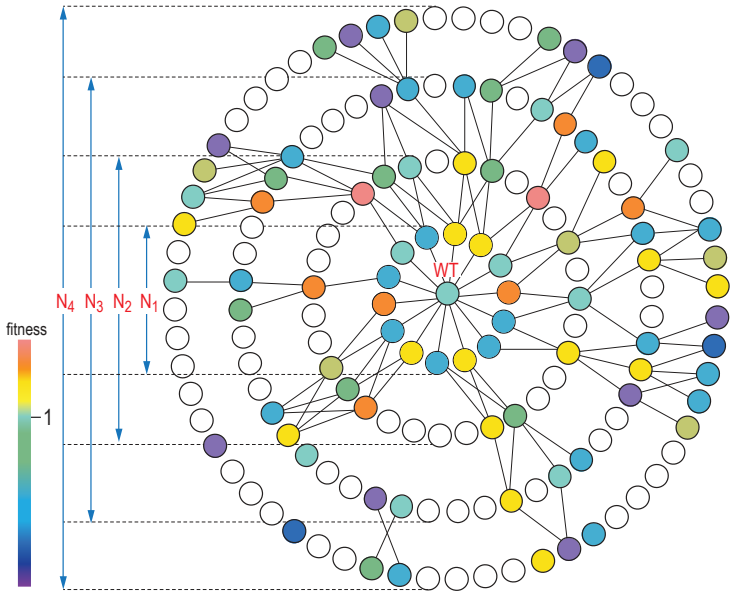
Results

The wild type is not the fittest genotype.

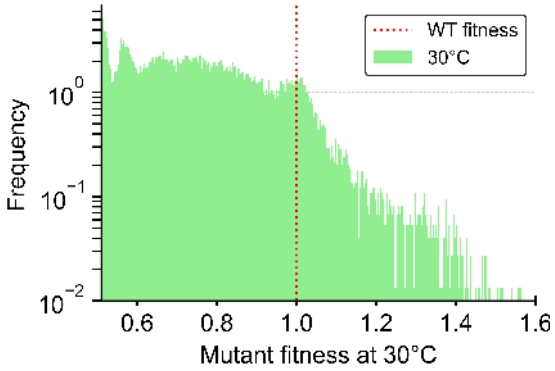
Our dataset consists of experimental fitness measurements of $\sim 65,000$ mutants of the *S. cerevisiae* $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ gene collected by Li *et al.* [54, 55]. Growth rates of 23,284 of these mutants were measured under four different environmental conditions: 23°C, 30°C, 37°C and oxidative stress. In the following, we refer only to the genotypes that were measured under all four conditions. The fitness of each genotype was defined as the base-2 exponent of its relative growth rate with respect to the wild type (see Methods). Hence, by definition the wild type fitness was set to 1, for each condition.

We began by closely examining the fitness values dataset. Our first remarkable finding was that the wild type was not the genotype with highest fitness under any of the four conditions, as one might expect from population-genetic models for single-locus selection, if the population is at steady state. Under each of the conditions, between 2000 and 2400 mutants (out of the 23,284) exhibited higher fitness than the wild type (Fig. 1b-e) [1]. We then analyzed possible sources for measurement errors, including read-count variability, as a source of inaccuracy in growth rate assessment and the possibility that the fitness effect was due to independent mutations that fortuitously occurred elsewhere in the genome (SI - Figs. S1-S2). While such error sources did exist, they could not fully account for the wild type's fitness sub-optimality.

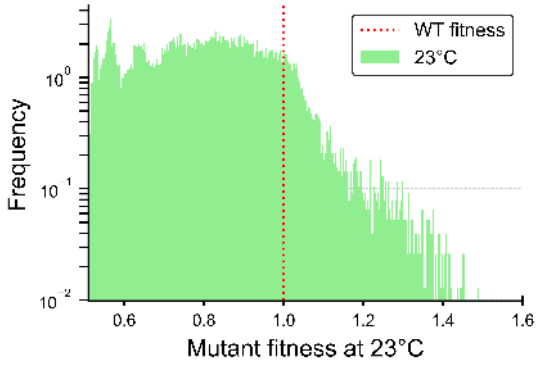
[1] 2441 genotypes in 23°C, 2075 genotypes in 30°C, 2008 genotypes in 37°C and 2236 genotypes in oxidative stress



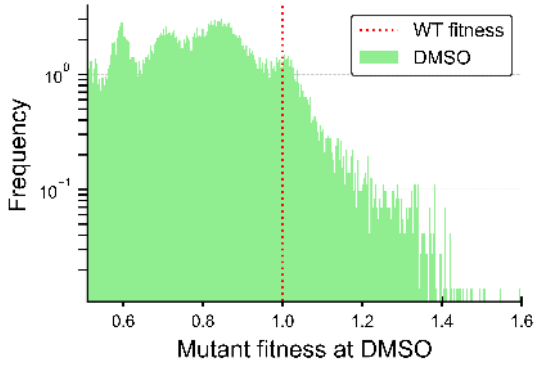
(a)



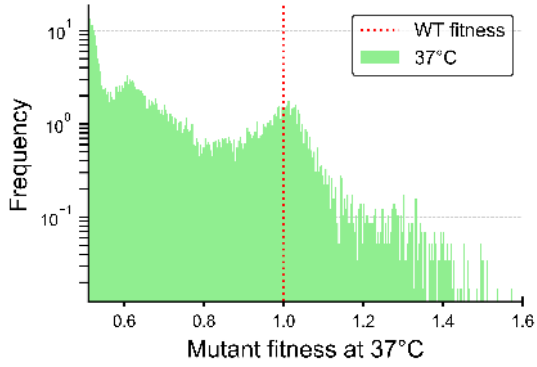
(b)



(c)



(d)



(e)

Figure 1 *(previous page)*: **Empirical fitness landscape of a tRNA gene (a)** A schematic visualization of the experimentally measured tRNA fitness landscape. Each circle represents a genotype. Filled circles represent genotypes whose fitness values (here encoded by different colors) were measured. Empty circles represent genotypes whose fitness values were not measured. We use here a concentric representation of the fitness landscape, centered around the wild type, where the minimal number of steps on the graph between any two genotypes is the number of point mutations separating them. The wild type is then surrounded by expanding circles of its single mutants (denoted by N_1), double mutants (N_2), etc. The experiment probed all the wild type's single-point mutants, but only decreasingly smaller proportions of the following mutational neighborhoods, N_i. **(b-e):** The distribution of all fitness values measured under four different conditions (30°C, 23°C, DMSO and 37°C), at semi-log scale. The wild type fitness value is shown in each by the red dotted line. Fitness was defined relative to the wild type's fitness, such that the wild type fitness was set to 1 for each condition. Under each of the conditions tested, 8%-10% of the genotypes in this dataset were fitter than the wild type. The relative weights of different fitness values were biased by the non-uniform sampling of the landscape, with dense sampling close to the wild type, and sparser sampling farther away.

The wild type is not the fittest on average across conditions

A possible explanation for the apparent sub-optimality of the wild type could be that while some mutants are fitter than the wild type under a specific condition, they are much less fit under other conditions, such that, *on average* the wild type is fittest. To test the applicability of this explanation for our case, we checked for each genotype the correlation between its fitness values under the various growth conditions. For high-fitness genotypes ($>$1.05 30°C), a high correlation was found between the fitness values measured under various conditions, $r = 0.75 - 0.91$ between fitness values at 30°C and fitness values under the other conditions. Namely, most genotypes which are fit under one condition are also fit under others (see Fig. 2a). In contrast, genotypes with low fitness in the range 0.6-0.8 at 30°C, showed a much lower correlation between their fitness values across conditions, $r = 0.28 - 0.49$ (see Fig. 2b). These results argue against the possibility that the wild type is the fittest on average, which would imply that genotypes having high-fitness under one condition should have low fitness under another.

To formally compare between fitness values averaged over multiple conditions, we considered the geometric mean fitness [56], $\langle f_i \rangle = (\prod_m f_i^m)^{1/M}$, where f_i^m is the fitness value of the i-th genotype in the m-th condition (out of M). The fitness values we have are relative to the wild type's, whose fitness was defined to be 1 under each of the conditions. Since growth rates differ between conditions, we must first transform the fitness values under the different conditions to a common baseline before we can calculate the geometric mean. To do so, we used the wild type growth rates reported by Li *et al.* for each of the conditions (See Methods section for details). Fig. 2c shows a histogram of the geometric mean fitness values $\langle f_i \rangle$ of all the genotypes in our dataset after transforming the original values. A possible caveat to this calculation is the underlying assumption that all four conditions are equally probable in the organism's natural habitat. Empirical fitness values might be inaccurate due to various reasons as discussed in the SI (Section 1). To reduce dependence on fitness value inaccuracies, we may also look at the fitness ranks: under each condition separately, all the genotypes are ranked according to their fitness values in ascending order (lowest fitness has

rank 1). The mean (and minimum) rank over the various conditions is then calculated for each genotype. If the fitter-than wild type mutants are fitter under one condition, but less fit under other conditions, it should be reflected in their mean rank being lower-than wild type mean rank. Fig. 2d shows the histogram of mean and minimum fitness ranks (over the four conditions) for all genotypes. The vertical dashed lines represent the wild type measures. We observe that the wild type is not the fittest across conditions in any of these calculations, but rather a sizable proportion of genotypes were fitter.

Evolutionary trajectories exist to fitter-than-wild type genotypes

The wild type sub-optimality could hypothetically be rooted in the fitness landscape topography. If, for example, the wild type were an isolated local maximum, separated from the global fitness maximum by fitness valleys, the population could be hindered from reaching the global maximum (at least temporarily) [8]. To test this hypothesis, we characterized the genotypes reached by single-point mutations in the wild type (or other focal genotypes) and then searched for evolutionarily accessible trajectories, namely trajectories along which fitness did not decrease. We began by exploring the location of the high-fitness genotypes with respect to the wild type's. "Mutational neighborhoods" $N_i(\mathrm{WT})$ surrounding the wild type were defined as the set of genotypes reached by i point mutations (shortest path) from the wild type (see Fig. 1a). Fig. 3b shows the fitness distributions of the four mutational neighborhoods $N_1 - N_4$ (single to quadruple mutants). We found that all four mutational neighborhoods contained fitter-than wild type genotypes, but the largest proportion of such fitter genotypes was in N_2, only two point mutations away from the wild type. Low-fitness genotypes were also found in all mutational neighborhoods. The fitness distribution of N_1 genotypes was much narrower than those of the other mutational neighborhoods, suggesting some level of correlation in fitness values between nearest neighbors.

Dissection of each mutational neighborhood into one of four fitness categories (Fig. 3a) found that 73% of the wild type's single mutants were nearly neutral ($0.9 < f < 1.05$) with respect to the wild type, 25% of them were much less fit ($0.6 < f < 0.9$), and 2% (4 genotypes) were significantly fitter than the wild type ($f > 1.05$). Amongst the $N_2(\mathrm{WT})$ genotypes (wild type's double mutants) the proportion of fitter-than wild type genotypes was even larger (944 out of 8101; 11%). We then checked for the existence of evolutionary trajectories of non-decreasing fitness, leading from the wild type to the fitter genotypes in N_2. To find whether direct access from the wild type to the fitter N_2 genotypes is possible via single-point mutations trajectory, we mapped all 2-step trajectories of strictly increasing fitness, originating from the wild type. We found ~1000 such trajectories, made possible due to the small number of fitter-than wild type single mutants. These comprise 2% of all possible 2-step trajectories originating from the wild type (total of $(69 \cdot 3)^2 \approx 4 \cdot 10^4$). Consider a yeast population of a typical size of 10^8. Over the course of time, it is highly likely that such trajectories will be visited and produce novel fitter genotypes [57]. This estimate is only a lower bound to the actual number of trajectories leading to fitter genotypes. If we also include trajectories passing through neutral single mutants, the number of trajectories reaching fitter double-mutants will be much higher.

We conclude that the wild type is not a local maximum. It is worth mentioning in this context

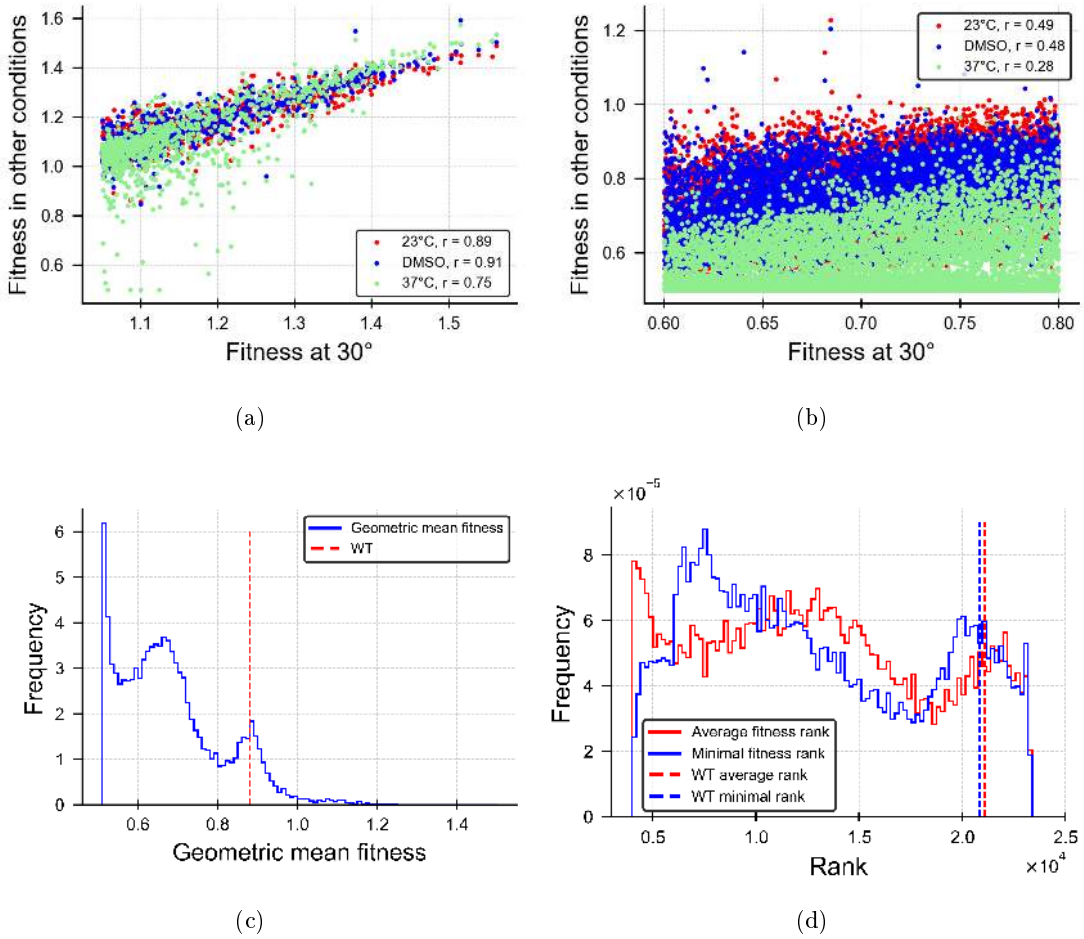


Figure 2: **The wild type is not the fittest across conditions.** Fitness values of mutants with fitness in the range [1.05, 1.6] **(a)** and in the range [0.6, 0.8] **(b)** at 30°C plotted against the fitness values of these genotypes under the other three conditions (23°C, oxidative stress and 37°C). The correlation coefficient between fitness values under different conditions were $r = 0.89, 0.91, 0.75$ respectively for the high-fitness range, but only $r = 0.49, 0.48, 0.28$ for the low-fitness range. We conclude that mutants that have high fitness in one environment usually have high fitness in all four of them. In contrast, for low-fitness genotypes, there is a much lower correlation between fitness values under different conditions. This rules out the possibility that the high-fitness genotypes "specialize" in one environment, while remaining inferior in others, whereas the wild type is the fittest on average. **(c)** Distribution of $\langle f_i \rangle$ - the geometric mean fitness (over the four environments) values for all genotypes. Note that after alignment of fitness values under different conditions to a common baseline, the wild type fitness is no longer 1. **(d)** Distribution of genotype average (red) and minimal (blue) fitness ranks over the four conditions (ascending fitness order, lowest fitness is assigned rank 1). In both **(c)** and **(d)**, the wild type fitness is denoted by vertical dashed lines. In either measure, a sizable proportion of the genotypes was fitter than the wild type.

that the higher the landscape dimension, the larger the number of possible single mutants for each genotype. A genotype is only a local maximum if *all* its single mutants are less fit. Hence, with the increase of landscape dimensionality, it is less likely to find local maxima [3].

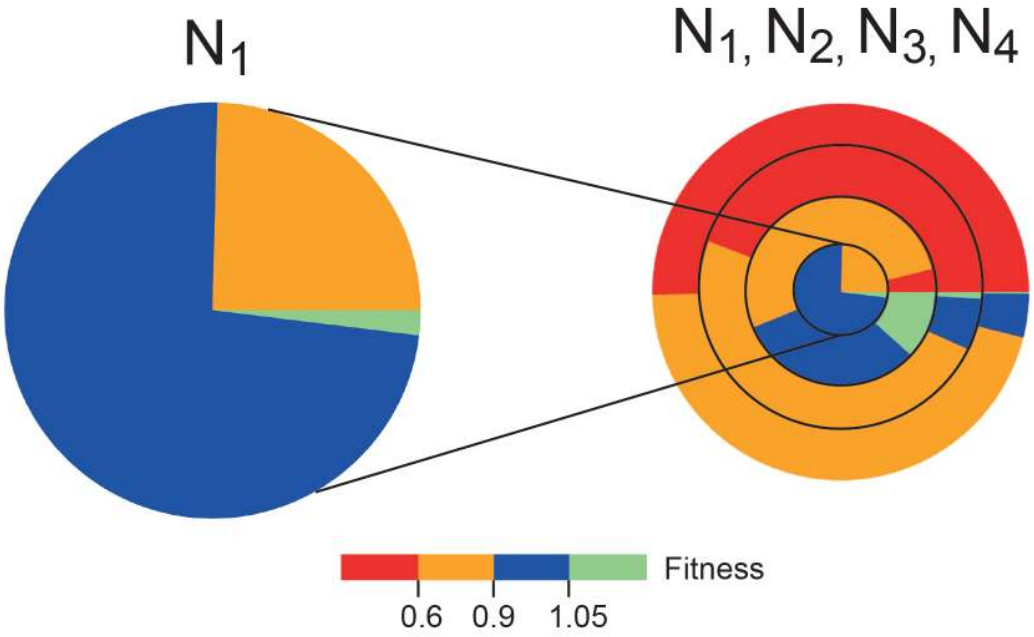
The wild type resides in a flat mutational neighborhood

Until this point, we saw that fitness considerations alone cannot explain how a sub-optimal genotype evolved to be the wild type. Recent literature suggests that two forms of selection may be at play. First-order selection drives populations towards higher fitness, where second-order selection promotes mutational robustness and adaptability [43] or alternatively, evolution maximizes the reproductive value rather than the fitness [44]. Hence, we next sought to characterize the relative robustness to mutations of different genotypes in our dataset. To quantify a genotype's evolutionary stability, we defined genotype steepness as the average fitness difference (absolute-value) between the genotype and its single mutants, $N_1(i)$:

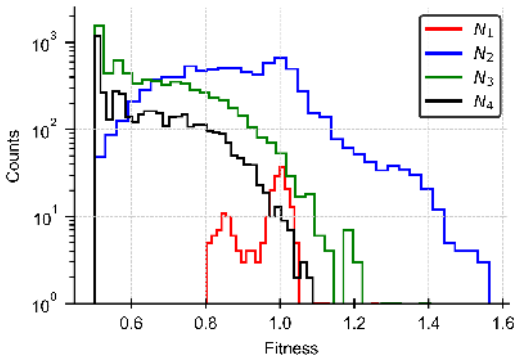
$$s_i = \frac{1}{|N_1(i)|} \sum_{j \in N_1(i)} |f_i - f_j|, \tag{1}$$

where $|N_1(i)|$ is the number of single mutants of genotype i. This definition is based on fitness information of all the single mutants of genotype i. In practice, with the exception of the wild type, we only have measurements of a subset of the single mutants and estimate s_i using partial data. To handle the non-uniformity (in number and in relative location) of single mutant sets amongst different genotypes, we only compared steepness between pairs of genotypes. This was achieved by sampling their corresponding single-mutant sets such that the set sizes are equal (see Methods section). To minimize biases because of small numbers of single mutants, steepness was only calculated for genotypes having at least 5 single mutants. This limitation enabled calculation of steepness only for 5615 genotypes (out of 23,284), yet the wild type and its single and double mutants were included in this calculation. As single-mutant sets can be randomly sampled in multiple ways (potentially yielding different s values), we sampled them multiple times, resulting in a distribution of steepness difference values between every two genotypes. Fig. 4A presents different statistical measures (mean, median, maximum etc.) of this steepness difference distribution, and shows the histogram of each measure over all the genotypes, relative to the wild type. Apart from the maximum, all measures had positive values only. Thus, all genotypes are *steeper* than the wild type, or in other words, the wild type is the flattest genotype in the dataset.

What is the relation between steepness and fitness? Assuming many additive contributions of different positions in the gene to its fitness, the fitness value distribution can be well approximated by a normal distribution, following the central limit theorem. In such a distribution, the most probable fitness value is very close (in the normal distribution it is exactly equal) to the median fitness value. We demonstrate here, that such fitness distributions yield a characteristic crescent-shaped fitness-steepness relation. For simplicity, assume an uncorrelated fitness landscape, such that the fitness values of a genotype 1-neighbors are randomly drawn from the fitness distribution. Then, genotypes with extreme fitness values (either very high or very low) are more likely to have single mutants with very different fitness values than their owns, and thus have high steepness. In contrast,



(a)



(b)

Figure 3: **The wild type is not a local maximum. (a)** We illustrate the fitness values of the wild type four mutational neighborhoods N_1, N_2, N_3, N_4 (right) using a color code. The majority of fitter than wild type genotypes (green section) are double mutants (in the second ring). Zooming into the composition of the inner circle N_1 (left), we find that a small subset of the wild type single mutants are fitter than the wild type (green). The vast majority of the single mutants are nearly neutral (blue). **(b)** Fitness value histograms of genotypes in the wild type's four mutational neighborhoods $N_1 - N_4$. Notably, all four mutational neighborhoods contain fitter than wild type genotypes ($f>1$), but the largest proportion of fitter genotypes is in N_2.

intermediate-fitness genotypes are more likely to have single mutants with fitness similar to their owns and consequently have low steepness. If the fitness landscape is correlated, the fitness values of the 1-neighbors are correlated to the focal genotype fitness, such that the above effect is weaker but still exists to some extent (depending on the level of correlation). We demonstrate this idea using the NK model [13], which allows for different degrees of landscape ruggedness and correlation by tuning of the model parameter K. Here we show simulation results with parameter values $N = 14$, $K = 6, 14$ (Fig. 4c-f). In this example, the fitness values of all genotypes are known and steepness can be calculated with no sampling bias. Fig. 4e-f presents the steepness vs. fitness values for all the genotypes in these examples. $K = 14$ represents a maximally rugged and uncorrelated landscape and $K = 6$ represents a partially correlated one. In both cases, the lowest steepness genotype was one that had an intermediate fitness value, and high-steepness genotypes were the ones with extreme fitness values (either high or low). The differences in steepness between the intermediate and the extreme fitness genotypes were larger in the uncorrelated landscape (Fig. 4f) than in the partially-correlated one (Fig. 4e), while their fitness values distributions were similar (Fig. 4c-d). In addition, a variety of steepness values were observed for many of the fitness values, in particular for intermediate fitness, where some had steepness as high as the steepness of the extreme fitness genotypes. It was thus intriguing to examine the fitness-steepness relation in the tRNA data. Fig. 4b presents a scatter plot of the steepness vs. fitness values of the different tRNA mutants and shows a picture similar to that observed in the simulated NK model landscape. Here too the lowest steepness was obtained for intermediate-fitness genotypes. Notably, among all genotypes with fitness values around 1, the wild type (blue circle) was one of the least steep genotypes.

For additional discussion of the effect of data incompleteness on steepness values and examination of differences in steepness calculated using different subsets of a genotype's single mutants, the reader is referred to the SI (Section 2, Fig. S3).

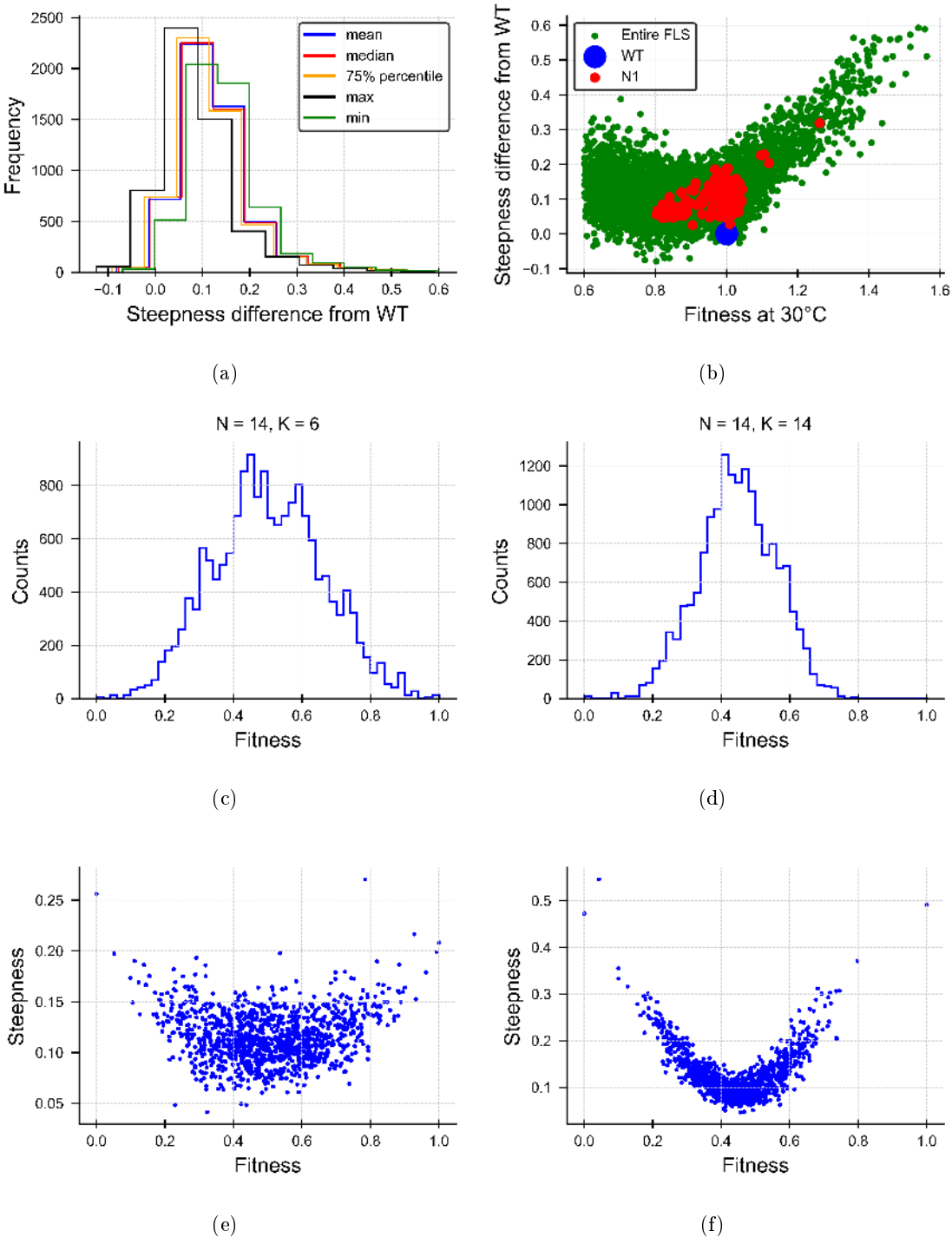


Figure 4

Figure 4 *(previous page)*: **The wild type is mutationally robust (flat). (a)**: The distributions of excess genotype steepness with respect to the wild type (see Methods for details). The different statistical measures (mean, median, etc.) were taken over different random samples of the genotype mutant sets, used for steepness calculation. These excess steepness distributions have almost exclusively positive values - namely the wild type is one of the least-steep genotypes in the tRNA dataset. **(b)** Steepness vs. fitness scatter plot of 5615 genotypes in the tRNA dataset. The wild type (blue circle) has one of the lowest steepness values in the whole dataset. In particular, it has very low steepness compared to other genotypes with similar fitness values. The wild type's single mutants (N_1) are shown for reference (red circles). Steepness was only calculated for genotypes with at least 5 single-mutants. **(c-f)** Fitness distributions **(c-d)** and steepness vs. fitness scatter plots **(e-f)** in a simulated NK landscape ($N = 14$, $K = 6, 14$). Both high and low fitness genotypes exhibit higher steepness. The genotypes with lowest steepness have intermediate fitness values. This effect is more pronounced in the uncorrelated landscape ($K = 14$, (d), (f)) but is still observed in the partially correlated landscape with $K = 6$ (c), (e).

Evolutionary dynamics on the empirical fitness landscape

We found that while the wild type was not the fittest in the landscape, it was among the flattest genotypes and specifically it was flatter than most other genotypes with similar fitness values. This suggests that this particular genotype could have become the wild type owing to its relative mutational robustness. To test this hypothesis and estimate the mutation rate at which the wild type should be stably maintained, we used stochastic simulations mimicking the evolutionary dynamics over the empirical tRNA fitness landscape. We used the measured fitness values and tested the effect of different mutation rates. Since we know the fitness values for only a subset of the actual landscape, mutations in our simulations could only reach genotypes included in this partial fitness landscape. Isolated genotypes with no single mutants cannot be reached or left via a single-point mutation and were therefor excluded from the simulations. See Methods for details on the simulation.

A unique technical challenge arises because different genotypes in the dataset can potentially have different numbers of single mutants. To resolve that, we defined the probability that a genotype is mutated independently of the number of its 1-neighbors. Yet, inequality in the number of single mutants could still bias the evolutionary dynamics if mutations affecting one genotype funnel into fewer genotypes than the mutations of another genotype. To describe that, a genotype's connectivity was defined as the number of its single mutants that were contained in our dataset $C(i) = |N_1(i)|$. The average population connectivity is then $\sum_i C(i)x_i$, where x_i is the population proportion of genotype i.

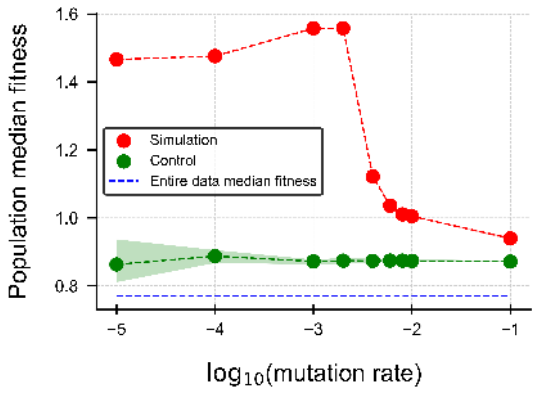
To disentangle fitness from connectivity effects in the simulation results, we also ran as a control a neutral evolutionary process on the very same genotype dataset, as if all genotypes were equally fit. The dynamics in these control runs is then purely neutral and only reflects the landscape connectivity structure. We then assigned quantitative measures (fitness, steepness, etc.) only in retrospect, such that they had no effect on the population dynamics during the simulation. Should the effects we observe in the evolutionary simulations appear also in the control, we can conclude that it is mostly due to the landscape non-uniform connectivity rather than due to selection.

Fig. 5a-d presents the final population median fitness, median steepness, average population

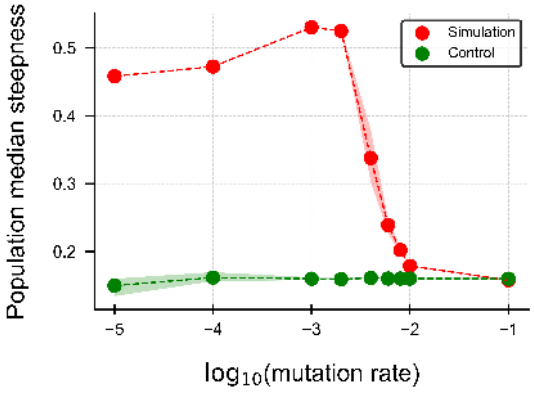
connectivity and the fraction of the population located in $N_1(\mathrm{WT})$ (the wild type's single-mutants set), for different mutation rates (red points) and compares them to the neutral simulation results (green points). Shaded regions around the curves represent the 25 and 75 percentiles over simulation repeats. All measures exhibited a sharp change at a critical mutation rate around $\mu_c \approx 10^{-3}$, where the population median fitness and steepness decrease and the population average connectivity and $N_1(\mathrm{WT})$ fraction increase. In contrast, the control simulations showed very little sensitivity, if any, to the mutation rate. Fig. 5c shows that at all mutation rates, the control population evolved to the highly connected regions of the landscape, in agreement with a previous study that addressed evolution on neutral networks [38]. As the majority of the dependence on mutation rate disappears in the neutral simulation, we conclude that most of the effect observed here is due to selection, rather than to the structure of the dataset.

To further study the effect of mutation rates on the population, we examined the population composition at the end of the simulation for different mutation rates. Fig. 5e-g shows the population fitness distribution for the adaptive simulation accounting for fitness (red) and for the control (green) at three different mutation rates. At the low mutation rate $\mu = 10^{-3}$ (Fig. 5e), the population is dominated by a few high-fitness genotypes ($f > 1.5$). From Figs. 5b-d, we learn that they have high steepness, few neighbors and are not single mutants of the wild type. As the mutation rate increases these genotypes are gradually replaced by others having lower fitness $f \approx 1$, lower steepness and higher connectivity (Fig. 5f-g). A significant proportion (≈ 0.25) of them are wild type single mutants (Fig. 5d). At all three mutation rates, the control simulation spans the whole range of fitness values with a peak at $f \approx 1$. Importantly, at the higher mutation rates $\mu = 10^{-2} - 10^{-1}$ when the adaptive simulation population is no longer dominated by a fitter than wild type genotype, it does show an enrichment at $f \approx 1$ *beyond* the control population. The control simulation represents the effect of data connectivity alone, and the peak it shows at $f \approx 1$ is due to the higher connectivity of the wild type's neighborhood. Thus, we conclude, that at the low mutation rates (Fig. 5e) the population dynamics is dominated by fitness, and connectivity has a minor effect. At the higher mutation rates (Fig. 5f-g) the population dynamics in our simulations is affected by a combination of genotype fitness and connectivity. The over-representation of $f \approx 1$ genotypes on top of the control population, suggests an additional role for the flatness of the wild type's neighborhood, beyond what is expected due to its connectivity alone.

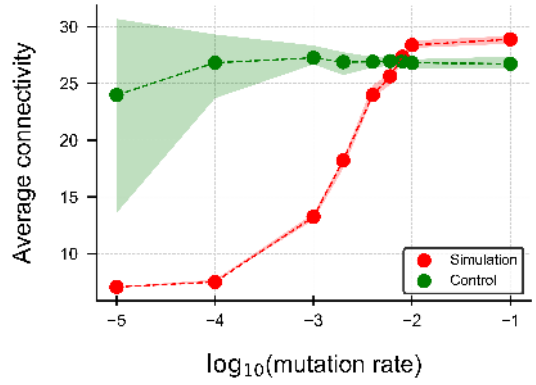
Fig. 5a also compares the population median fitness to the median fitness value of the entire dataset (horizontal blue dashed line). Had the neutral simulation uniformly sampled all the genotypes in the dataset, we would have expected it to overlap with the dataset median fitness value. The population median fitness obtained in the control simulation was very close to the median fitness value of the entire dataset, but was consistently slightly higher. For comparison, we ran a similar evolutionary simulation with a neutral control on an artificially fabricated NK model landscape for which we have full data of all genotypes and equal connectivity for all. There, the fitness of the control simulation exactly overlapped with the dataset median fitness value (SI, Fig. S4). Thus, we attribute the small gap between the control (green points) and the dataset median to some level of correlation between high connectivity and high fitness in the dataset.



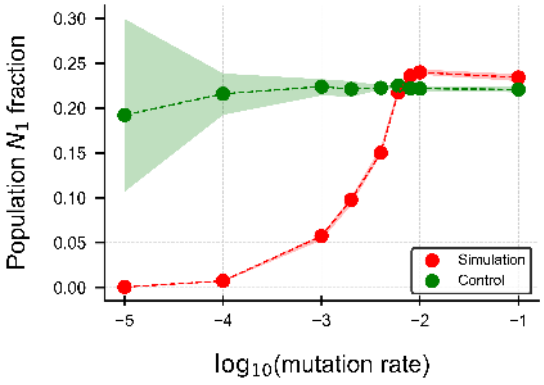
(a)



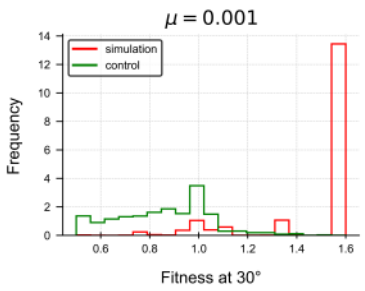
(b)



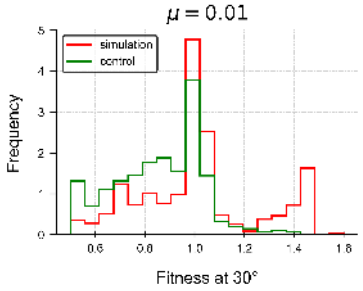
(c)



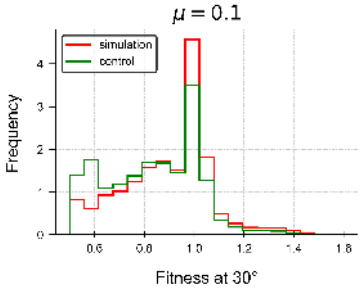
(d)



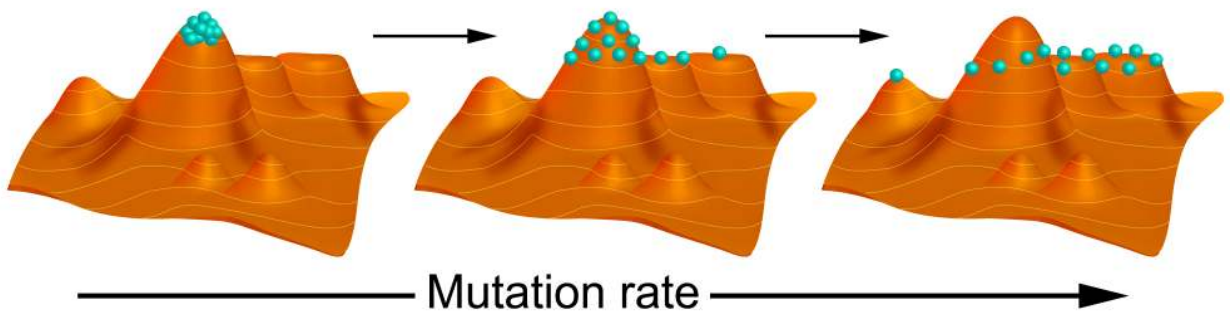
(e)



(f)



(g)



(h)

Figure 5 *(previous page)*: **Evolutionary simulations on the tRNA fitness landscape exhibit fitness decline at a threshold mutation rate.** We simulated the evolutionary dynamics of the experimentally measured tRNA fitness landscape at different mutation rates (red circles). To test whether the landscape incompleteness of our dataset affects the results, we also ran a control simulation (green circles). There, we simulated a neutral evolutionary process on the same dataset, assigning equal fitness values to all genotypes throughout the simulation. The quantitative measures (fitness, steepness, etc.) shown in the graphs were only assigned to the genotypes in retrospect and had no effect on the dynamics. We plot the final population median fitness **(a)**, median steepness **(b)**, average connectivity (number of single mutants per genotype) **(c)** and fraction of the population located in N_1 (the wild type's single mutants) **(d)**. All measures are plotted against the mutation rate per base pair per generation (log-scale). Shaded regions represent the 25 and 75 percentiles over repeated runs with the same parameters. All measures exhibit a sharp change at a critical mutation rate around $\mu_c \approx 10^{-3}$. In contrast, the control simulation results are almost independent of the mutation rate for all measures, but exhibit a larger variation at low mutation rates. Median population fitness (**a**) in the control is consistently slightly higher than the median fitness of the whole dataset, suggesting some degree of correlation between connectivity and fitness in this dataset. Simulation parameters: population size = 10,000. Simulations were run for 2500 generations each with 15 repeats for each parameter combination. See Methods for more details. **(e-g)** Examples of population fitness distributions at the end of the simulations for the adaptive simulation (red) and control (green) at three mutation rates $\mu = 10^{-3}, 10^{-2}$ and 10^{-1} respectively. We observed transition in the landscape occupancy from few high-fitness ($f > 1.5$) genotypes at $\mu = 10^{-3}$ to a multitude of genotypes with $f \approx 1$ at $\mu = 10^{-2}$. At the lowest mutation rate $\mu = 10^{-3}$ the adaptive simulation population exhibits a markedly different distribution compared to the control. At the higher mutation rates, the adaptive simulation population shifts to lower fitness values, similar to the control. Yet, the adaptive simulation has a higher population proportion around fitness values of ≈ 1 (the wild type's) compared to the control population. Thus the high proportion of wild type-like genotypes in the adaptive simulation cannot be fully accounted for by the wild type's higher connectivity. **(h)** A cartoon of the mutation-selection balance on a rugged fitness landscape at different mutation rates. At low mutation rates (left), the population is composed mostly of the fittest genotype. The higher the mutation rate, the more diverse the population becomes (middle, right). At the highest mutation rate, we exemplify the "survival of the flattest" where intermediate fitness genotypes survive because their mutational neighborhood is relatively flat.

Scaling of the critical mutation rate

The mutation threshold phenomenon we found in the simulations occurred at a lower mutation rate than the "error catastrophe" predicted by quasi-species theory. "Error catastrophe" is a phase transition between a single (or few) high-fitness genotype to a cloud of highly connected lower-fitness genotypes. In the absence of crossing-over, it is expected at a mutation rate $\mu_e \approx 1/L$, where L is the genome length [45]. Intuitively, when the mutation rate exceeds $1/L$, each individual receives, on average, one mutation per generation, and hence, no error-free copies of the genome remain. In our simulations, $L = 72$ and error catastrophe is expected at $\mu_e = 0.014$.

Our estimate of the critical mutation rate $\mu_c \approx 10^{-3} < \mu_e$ - beyond which the fittest genotypes no longer dominated - was determined based on only a small segment of the genome (72 nt). In order to compare it to known mutation rates in yeast, we need first to check how it scales with the number of genes or with the genome size. We use here a highly simplified model of the genome. Assume all genes have either one of two fitness values: fit (high fitness, low robustness) or flat (sub-

optimal fitness, high robustness). We use a similar rationale as in the error catastrophe argument. Assume a genome consisting of K genes having equal properties: equally likely to be mutated and equal contribution to the total organismal fitness. If a single gene requires a mutation rate $\mu_c^{(1)}$ to become flat, then the mutation rate needed for one out of K genes, on average, to become flat is $\mu_c^{(K)} = \frac{1}{K}\mu_c^{(1)}$.

In the complete absence of crossing-over, although genomes carrying a flat gene are less fit, they cannot be purged by selection, because the fully fit genotype no longer exists at this mutation rate. Mutations then keep accumulating, one following the other, and then additional genes should transition from fit to flat. This process goes on, until all genes should become flat. Similarly to the error catastrophe argument, here too the critical mutation rate for *all* genes to become flat inversely depends on the number of genes. At the other extreme of highly frequent crossing-over, every gene evolves independently of the others, because the fully-fit genome can be restored by crossing-over distinct genomes having different genes that are non-fit. At a mutation rate $\mu_c^{(K)}$ we then expect to find one flat gene on average per individual (which can be a different gene for different individuals), where the majority of the genes remain fit.

Returning to the tRNA, the evolutionary simulation of the single gene fitness landscape showed transition from fit to flat genotypes at a mutation rate of $\mu_c^{(1)} \approx 10^{-3}$. The yeast genome is known to contain roughly 6600 genes [58]. Add to that an equal number of non-coding regions affecting fitness and consider each of them twice because the genome is diploid. If yeast were purely asexuals, we would obtain that $\mu_c(\text{full genome}) \approx \frac{10^{-3}}{4\times 6600} = 4 \times 10^{-8}$. Alternatively, when considering the relative weight of the 72 nt tRNA gene in the 1.2×10^7 nt long full genome, we obtain $\mu_c(\text{full genome}) \approx 6 \times 10^{-9}$. Yeast are not obligatory sexuals, but do occasionally pursue sexual reproduction. The linkage disequilibrium of yeast laboratory strains was reported to fall to half at a distance of 23 kb [59]. Thus, yeast should fall in between the two extremes with possibly a few flat genes per genome at $\mu_c(\text{full genome})$. The previous estimate assumed equal properties to all genes, and in particular a uniform mutation rate genome-wide. Yet, mutation rates are not uniform across the genome, and specifically tRNA genes were shown to be 7-10-fold more mutable compared to the background genome [60, 61]. Thus, we expect that highly mutable genes, such as the tRNA, are likely to be amongst these few flat genes.

Discussion

Recent advances in high-throughput experimental methods have allowed for large-scale characterization of empirical fitness landscapes [4, 5, 6, 7], which can be applied to test hypotheses about the driving forces of evolutionary dynamics. Here, we test the "survival of the flattest" hypothesis, which suggests that beyond a certain level of mutation rates, it is not the fittest genotype that becomes dominant but rather the flattest one - the genotype that is the most mutationally robust (though often sub-optimal).

In this study, we analyzed the fitness measurements of 23, 284 genetic variants of the $\text{tRNA}^{\text{Arg}}_{\text{CCU}}$ *S. cerevisiae* gene made by Li *et al.* [54, 55]. We found that the wild type allele is not the fittest, but instead, approximately 8%-10% of the measured mutants (2000-2400) were fitter. The wild type allele was not only sub-optimal in each of the four conditions measured, but also sub-optimal on

average on all four of them. Still, as fitness was only measured under four different conditions, it is possible that some of the high-fitness mutants are inferior in another condition not included in this experiment [62]. Characterization of the fitness landscape showed that the wild type is not a local fitness maximum either and that fitness-increasing trajectories leading to fitter genotypes only two mutations away, are feasible. Instead we find, that the wild type is mutationally robust. We used stochastic evolutionary simulations on the empirical fitness landscape to study the conditions at which the wild type allele should be favored by selection. The simulations showed transition from high-fitness but mutationally sensitive to sub-optimal but mutationally robust genotypes (among them is the wild type) at a critical mutation rate of $\mu_c \approx 10^{-3}$ [per position, per cell division]. This estimate refers to the tRNA alone. The *S. cerevisiae* genome contains ~6600 genes as well as non-coding regions contributing to fitness. Extrapolation of the critical mutation rate to the full genome is complicated, because different genes have different properties (such as propensity to mutate and contribution to organismal fitness) and yeast occasionally engage is sexual reproduction. Under several simplifying assumptions, we estimated that a mutation rate of $4 \times 10^{-8} - 6 \times 10^{-9}$ several flat genes are expected per genome. For reference, the genome-wide mutation rate in *S. cerevisiae* was estimated to be $10^{-9} - 10^{-10}$ [63, 64, 57]. However, genes vary in their mutation rate and, in particular, tRNA genes were shown to have a mutation rate 7-10-fold higher than the background genome [60, 61]. We propose that while the majority of yeast genes are still selected to be fittest, there are a few mutation hot-spots like the tRNA for which the mutationally robust flat allele is favored by selection.

Analysis of different error sources as potential explanations for the reported high-fitness mutants was included in the original publications [54, 55]. We repeated some of them, e.g. fitness inaccuracy due to read-count noise, an estimate of the possibility that the high fitness is due to other mutations in the genome and the violation of exponential growth assumption (see SI, Section 1). We found that they could explain 10% of the fitter-than-wild type mutants, at most.

Inherent to the astronomical dimensionality of fitness landscapes is our inability to fully measure them, or even get close to a full measurement. Even the high-throughput measurements possible today only sample specific regions of the fitness landscapes, concentrated around the wild type gene or alternatively sample randomly scattered mutants [65]. In our case, the mutant library was produced by mutating the wild type allele. Its coverage was almost full near the wild type and became sparser further away. This non-uniform coverage introduces technical challenges, because different genotypes in the dataset potentially have different numbers of single mutants. This could have potentially biased our evolutionary simulations on the landscape and the steepness calculation. To alleviate this concern, we used a neutral simulation as a control and a special sampling procedure for the steepness calculation. Reassuringly, we found the effect of non-uniform landscape sampling to be minor at low mutation rates and only partial at higher mutation rates (see Fig. 5 and SI Fig. S3). Usage of fragments of a fitness landscapes to draw general conclusions is a common practice in the field. However, it does raise the fundamental question whether indeed it represents the entirety of the landscape and hence, should be used with caution. Full landscape mappings are possible only for computationally fabricated and relatively small landscapes [8, 24, 25]. Alternatively, massive random sampling of fitness landscapes, not concentrated around the wild type, could be used to

test the validity of this common approach. Recent works also handled the sparse sampling of fitness datasets by interpolating between the measured points to estimate fitness values of missing genotypes [66, 29] with some success. While these techniques are computationally very demanding, it would be interesting to test in the future whether they are applicable for computing evolutionary dynamics on incomplete landscapes.

"Survival of the flattest" is a theoretical prediction which is a direct outcome of quasi-species theory [49]. It was demonstrated in simulations on digital organisms [51] and experimentally in RNA viruses [53]. Previous theory focused on the simplest case of competition between two species, each located on a single peak fitness landscape: one which has high fitness but is steep and the other that has lower fitness but is less steep [49, 50], however actual fitness landscapes are significantly more complex. A general theory for the emergence of flatness and, in particular, calculation of the critical mutation rate at which this transition occurs, is still lacking. Here, we estimated this critical mutation rate using evolutionary simulations. A theoretical framework unifying selection for high fitness and selection for mutational robustness was recently introduced [67, 44].

As the number of large-scale fitness measurements of particular landscapes is still limited, additional examples for wild type genes being sub-optimal are scarce. Bank *et al.* characterized the fitness landscape of the yeast heat shock protein Hsp90 [30] and found some mutants with higher than wild type fitness. Experiments mimicking horizontal gene transfers found that replacement of a gene by orthologs from another species could, in a few cases, increase the organismal fitness [68].

Evidence of 'second-order' selection for adaptability and mutational robustness was reported [43, 41]. Multiple experimental works have demonstrated that high-fitness genotypes have less access to beneficial mutations or that the very same beneficial mutations have a smaller fitness effect, whereas deleterious mutations have a larger effect on the background of high-fitness genotypes compared to low-fitness ones [69, 43]. As most mutations of high-fitness genotypes are deleterious, these observations are in line with our findings that high-fitness genotypes are usually also steeper. This global pattern of high-fitness genotypes being more sensitive to mutations compared to lower fitness genotypes could be a simple statistical outcome of high-fitness ones being scarce [8] or the outcome of global epistasis patterns [70].

Mutation rates were previously thought to be uniform across the genome and hence, genome-wide elevated mutation rates were considered necessary to detect flatness [53]. Recent measurements reveal a more intricate picture of mutations, with variation in mutation rate at all scales: between strains [64], between individuals and across the genome [71, 72, 73]. A variety of factors, such as sequence context, transcription level, nucleosome occupancy, DNase hypersensitivity, and recombination rate, have been linked to increased fine-scale susceptibility to mutation [74, 72]. Additionally, different genes can be under different selection pressures. Building on these non-uniformities in both selection intensity and mutation rate, we propose that sporadic instances of sub-optimal and mutationally robust alleles exist in nature, while the majority of the genome is still fittest. Testing this hypothesis would require probing the fitness landscapes of both high- and low-mutation rate genes in the same organism.

Methods

Fitness data

We used the fitness measurements as published in [54]. For completeness, we briefly summarize how fitness was measured and defined there:

1. Cells were sampled and sequenced at time T_0, right before competition. The frequencies of the different genotypes $x_i(0)$ were then calculated $x_i(0) = \frac{R_i(0)}{\sum_j R_j(0)}$, where $R_j(t)$ is the number of reads of genotype j at time t. These baseline frequencies were then used for all conditions.

2. The original cell pool was then split to the four different conditions, with at least 3 replicates for each (30°C and 23°C were replicated 6 times, the rest 3 times).

3. All genetic variants were grown together for 24 h, where after 12 h, the culture was diluted by a factor of 1/100.

4. After 24 h (T24) cells from each of the growth conditions were sampled and sequenced.

The number of wild type generations in 24 h under condition m during competition was calculated as:

$$G^m_{\mathrm{WT}} = \log_2 \left(d \cdot \frac{g^m(24) x^m_{\mathrm{WT}}(24)}{g(0) x_{\mathrm{WT}}(0)} \right),$$

where $g^m(t)$ is the total number of cells at time t (calculated using the culture cell density (OD) measure) at condition m and d is the dilution factor. The measurements at time $t = 0$ were common to all conditions. The per-generation fitness of variant i at condition m was defined there as

$$f^m_i = \left(\frac{R^m_i(24)/R_i(0)}{R^m_{\mathrm{WT}}(24)/R_{\mathrm{WT}}(0)} \right)^{1/G^m_{\mathrm{WT}}}. \tag{2}$$

G^m_{WT} is the number of wild type generations under condition m, $G^m_{\mathrm{WT}} = r^m_{\mathrm{WT}} \cdot t / \log 2$, where r^m_{WT} is the wild type exponential growth rate under that condition. Turning to continuous time and assuming that all variants grow exponentially during the entire experiment (neglecting lag and yield phases), Eq. (2) can be written as

$$f^m_i = \left(\frac{\exp(r^m_i \cdot t)}{\exp(r^m_{\mathrm{WT}} \cdot t)} \right)^{1/G^m_{\mathrm{WT}}} = \exp\left[(r^m_i - r^m_{\mathrm{WT}}) t \cdot \frac{\log 2}{r^m_{\mathrm{WT}} t} \right] = 2^{\left(\frac{r^m_i}{r^m_{\mathrm{WT}}} - 1 \right)} = 2^{\left(\frac{G^m_i}{G^m_{\mathrm{WT}}} - 1 \right)}, \tag{3}$$

where r^m_i is the growth rate of the $i-th$ genotype under condition m. Hence, the fitness of a genotype i under condition m is its exponentiated relative growth rate difference with respect to the wild type's under that condition. In the common notation of population genetics, a mutant has fitness advantage s over the wild type, if it has on average $(1 + s)$-times more offspring per-generation. Thus, $2^{\left(\frac{r^m_i}{r^m_{\mathrm{WT}}} - 1 \right)} = 2^s$. The wild type's fitness equals 1 by definition, under each of the conditions.

Fitness value re-scaling between conditions

The fitness values of genotypes were defined relative to the wild type's under each of the conditions and thus are incomparable between conditions. In order to calculate an average fitness over all conditions, we first needed to define a common baseline to compare values referring to different conditions. Fitness of genotype i under condition m_2 relative to the wild type's fitness at m_2 was defined as:

$$f_i^{m_2} := 2^{\left(\frac{r_i^{m_2}}{r_{\mathrm{WT}}^{m_2}}-1\right)}. \tag{4}$$

Now we would like to define it when the reference is the wild type's fitness at condition m_1, namely:

$$\tilde{f}_i^{m_2} := 2^{\left(\frac{r_i^{m_2}}{r_{\mathrm{WT}}^{m_1}}-1\right)}. \tag{5}$$

Substituting Eq. (4) we obtain

$$r_i^{m_2} = \log_2(f_i^{m_2}) \cdot r_{\mathrm{WT}}^{m_2} + r_{\mathrm{WT}}^{m_2} = r_{\mathrm{WT}}^{m_2}(1 + \log_2(f_i^{m_2})). \tag{6}$$

Substituting $r_i^{m_2}$ into Eq. (5) we obtain:

$$\tilde{f}_i^{m_2} = 2^{(\log_2(f_i^{m_2}) \cdot r_{\mathrm{WT}}^{m_2} + r_{\mathrm{WT}}^{m_2} - r_{\mathrm{WT}}^{m_1})/r_{\mathrm{WT}}^{m_1}}. \tag{7}$$

If we return to the original reference growth rate $r_{\mathrm{WT}}^{m_2}$, the equation reduces to the original fitness definition, such that $\tilde{f}_i^{m_2} = f_i^{m_2}$.

We chose the measurements at 30°C to be our reference. For this calculation, we used the wild type growth rates under the four conditions as reported in [54]: $r_{\mathrm{WT}}^{23\mathrm{C}} = 0.25$, $r_{\mathrm{WT}}^{30\mathrm{C}} = 0.5$, $r_{\mathrm{WT}}^{\mathrm{DMSO}} = 0.45$, $r_{\mathrm{WT}}^{37\mathrm{C}} = 0.43$. The growth rate units are Δ OD/hour.

Steepness calculation

Our basic definition of a genotype i steepness is the average fitness difference (absolute-value) between a genotype and its single mutants $N_1(i)$:

$$s_i = \frac{1}{|N_1(i)|} \sum_{j \in N_1(i)} |f_i - f_j|, \tag{8}$$

where f_j are the fitness values of these single mutants and $|N_1(i)|$ is their number. However, our dataset contains fitness values of only a small subset of the tRNA gene fitness landscape with non-uniform sampling of the genotype space: dense close to the wild type and sparser further away. Thus, only for the wild type, we have nearly full coverage of its single mutants, whereas for most other genotypes, only a few of their single mutants were measured. This non-uniform sampling could potentially bias the steepness calculation.

To handle this, we applied an alternative calculation which facilitates only pairwise comparisons between genotypes. To compare the steepness of a genotype G with that of another genotype

C, we took the single mutants of each $N_G \in N_1(G)$ and $N_C \in N_1(C)$. We defined $d(G_1, G_2)$ as the Hamming distance between the genotypes G_1 and G_2 and then imposed symmetry of the two neighbor sets, as follows. For every genotype $N_G \in N_1(G)$ we calculated its distance from C, $d(N_G, C)$. We then searched for a 1-neighbor of C that had equal distance from G, namely, $d(N_C, G) = d(N_G, C)$. If no such neighbor of C existed, we discarded N_G. If there were one, we included both genotypes in the neighbor dataset which would be used for steepness calculation. If multiple appropriate 1-neighbors existed, we randomly selected one of them. We repeated this procedure, until either C or G had no more 1-neighbors left. This procedure creates symmetrical mutational neighborhoods of G and C with equal numbers of neighbors and symmetric arrangement of the neighbors of each genotype with respect to the other. For genotypes with very few neighbors, the neighbor set might be empty and then the steepness calculation becomes unfeasible. We used this procedure to compare the steepness of the wild type with that of other genotypes in Fig. 4a. Since the wild type has a much larger number of neighbors compared to most genotypes, there are often multiple ways to sample its mutational neighborhood for the steepness calculation. Hence, we randomly drew the set of wild type neighbors 100 times. We then obtained a distribution of wild type steepness values that we subtracted from the other genotype steepness. Fig. 4a illustrates different statistical measures of this distribution as representative values of the wild type distribution: mean, median, 75% percentile, maximum and minimum. The choice of maximum steepness value as wild type representative is of course the strictest one. Only then, a few genotypes have lower steepness than the wild type's. We included in the calculation only genotypes that had at least 5 nearest-neighbors.

Simulated fitness landscape using the NK model

We simulated a correlated fitness landscape using the NK model [8]. A genotype in this model is represented by a binary string of length N, so the fitness landscape consists of 2^N genotypes. The parameter K is used to tune the ruggedness of the fitness landscape, such that for $K = 1$ it is fully additive and smooth and for $K = N$ it is the most rugged. The fitness of each genotype is defined as the average of N contributions of its N positions. Each contribution is determined by a particular position and its $K - 1$ neighbors. We drew 2^K random numbers either from a uniform distribution and assigned these values to be the fitness contributions of the possible binary strings of length K, $f(s_0, \ldots, s_{K-1})$, $s_i \in [0, 1]$. The fitness of each genotype encoded by a binary string of length N is defined as the average of the N fitness contributions of the length-K strings it contains (cyclically): $F(s_0, s_1, \ldots, s_{2^N-1}) = \frac{1}{N} \sum_j f(s_j, s_{j+1}, \ldots, s_{(j+K-1) \bmod N})$.

Evolutionary simulations

The quasi-species evolutionary dynamics is given by [45, 46], where we use the notation as in [75]:

$$\dot{x}_i = \sum_{j=1}^{n} x_j q_{ji} f_j - \bar{f} x_i \qquad i = 1 \ldots n, \tag{9}$$

where x_i is the frequency of the i'th genotype, such that $\sum_i x_i = 1$. q_{ji} is the probability that replication of genotype j results in genotype i, such that $\sum_i q_{ji} = 1$. If replication is error-free,

$q_{ii} = 1$ and $q_{ji} = 0$ for $j \neq i$. f_i is the fitness of genotype i and $\bar{f} = \sum_j x_j f_j$ is the population mean fitness. We assume that the mutation rate q_{ji} between genotypes j and i only depends on the Hamming distance h_{ji} between those genotypes:

$$q_{ji} = (1-q)^{L-h_{ji}} q^{h_{ji}}, \tag{10}$$

whereas L is genome length and q is the mutation rate per position per generation.

Rather than solving the differential equations, we used an agent-based stochastic simulation to study the evolutionary dynamics on the experimentally measured fitness landscape. As some genotypes in the dataset have no single mutants, they are unreachable via single-point mutations. Hence, we excluded such genotypes and ran the simulation on a subset of only 15,000 genotypes, which were all, at most 3 mutations away from the wild type. To simplify the implementation, all the measured fitness values were re-scaled to the range $[0, 1]$. To tackle memory and complexity issues, we computed the Hamming distance matrix H, $H_{ij} = h_{ij}$ up front. Each row of H was calculated using sparse matrix multiplication, resulting in a $\sim 10^8$ entries matrix stored on disk in an h5 file type which allows easy slicing access.

We initialized a population of N =10,000 individuals as detailed below. We randomly chose 10% of the possible genotypes in our simulation dataset (approx. 1500 genotypes) as the seed of the initial population and then assigned equal number of copies (6-7 copies) of each genotype to the initial population. The simulation algorithm then followed the Moran model [76]: At each iteration, we pick an individual i. Then either of three things happens: it can either replicate (with or without mutation) or not replicate at all. With probability $1 - f_i$, where f_i is its fitness, it will not replicate. With probability f_i, it will replicate and replace another randomly chosen individual j regardless of its fitness f_j. Replication can be accompanied by mutation of i in a single position with probability $1 - (1-q)^L$. In the latter case, its mutated version $\tilde{i}$ will replace j. We then move on to the next individual. This algorithm ensures a fixed population size of N. The simulation was written using the mesa python package [77]. The simulation algorithm (written in pseudo-code) is detailed here:

Algorithm 1 quasi-species simulation

1. Input:
   (a) $N \leftarrow$ #Population Size
   (b) $q \leftarrow$ #Mutation Rate per base pair per generation
   (c) $g \leftarrow$ #Number of generations to run
2. Randomly initialize the population
3. Do g times:
   (a) Go through the whole population and for each individual i do:
      i. $r' \leftarrow$ sample from a uniform distribution in the range [0,1] (# We normalized the fitness values)
      ii. If $r' < f_i$: (# the individual will reproduce)
         A. $j \leftarrow$ select another individual to die
         B. $r'' \leftarrow$ sample from a uniform distribution in the range [0,1]
         C. if $r'' < 1 - (1-q)^L$: //The chances of at least one mutation to occur while replicating
            - $\tilde{i} \leftarrow$ Choose a 1-neighbor of i
            - Replace individual j by $\tilde{i}$
         D. else: #($r'' \geqslant 1 - (1-q)^L$)
            - Replace individual j by i
   (b) Shuffle the individuals order

The simulation allows only single mutation steps, because the probability for higher-order mutations is very low. The simulation was run for a fixed number of 2500 generations, where in every generation the simulation goes over all individuals in the populations. This number of generations was chosen after we verified that it is sufficient to reach a mutation-selection balance. We repeated the simulation 15 times for each parameter combination, such that each repeat is initialized independently.

Acknowledgements We thank Chuan Li and Jianzhi Zhang for sharing their experimental data and Amos Tanay for help with the steepness calculation. We thank Yoav Ram and Daniel Weissman for comments on the manuscript. T.F. acknowledges funding from the Hebrew University of Jerusalem. Y. P. acknowledges grant support from the Minerva foundation.

References

[1] Sewall Wright. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. 1932.

[2] Ivan G. Szendro, Martijn F. Schenk, Jasper Franke, Joachim Krug, and J. Arjan G. M. de Visser. Quantitative analyses of empirical fitness landscapes. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(01):P01005, January 2013.

[3] Uri Obolski, Yoav Ram, and Lilach Hadany. Key issues review: evolution on rugged adaptive landscapes. *Reports on Progress in Physics*, 81(1):012602, 2018.

[4] Benjamin P. Roscoe, Kelly M. Thayer, Konstantin B. Zeldovich, David Fushman, and Daniel N. A. Bolon. Analyses of the Effects of All Ubiquitin Point Mutants on Yeast Growth Rate. *Journal of Molecular Biology*, 425(8):1363–1377, April 2013.

[5] Hervé Jacquier, André Birgy, Hervé Le Nagard, Yves Mechulam, Emmanuelle Schmitt, Jérémy Glodt, Beatrice Bercot, Emmanuelle Petit, Julie Poulain, Guilène Barnaud, Pierre-Alexis Gros, and Olivier Tenaillon. Capturing the mutational landscape of the beta-lactamase TEM-1. *Proceedings of the National Academy of Sciences*, 110(32):13067–13072, August 2013. Publisher: National Academy of Sciences Section: Biological Sciences.

[6] Karen S. Sarkisyan, Dmitry A. Bolotin, Margarita V. Meer, Dinara R. Usmanova, Alexander S. Mishin, George V. Sharonov, Dmitry N. Ivankov, Nina G. Bozhanova, Mikhail S. Baranov, Onuralp Soylemez, Natalya S. Bogatyreva, Peter K. Vlasov, Evgeny S. Egorov, Maria D. Logacheva, Alexey S. Kondrashov, Dmitry M. Chudakov, Ekaterina V. Putintseva, Ilgar Z. Mamedov, Dan S. Tawfik, Konstantin A. Lukyanov, and Fyodor A. Kondrashov. Local fitness landscape of the green fluorescent protein. *Nature*, 533(7603):397–401, May 2016.

[7] Olga Puchta, Botond Cseke, Hubert Czaja, David Tollervey, Guido Sanguinetti, and Grzegorz Kudla. Network of epistatic interactions within a yeast snoRNA. *Science*, 352(6287):840–844, May 2016.

[8] Stuart Kauffman and Simon Levin. Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology*, 128(1):11–45, September 1987.

[9] Daniel B. Weissman, Michael M. Desai, Daniel S. Fisher, and Marcus W. Feldman. The Rate at Which Asexual Populations Cross Fitness Valleys. *Theoretical population biology*, 75(4):286–300, June 2009.

[10] Daniel M Weinreich. The Rank Ordering of Genotypic Fitness Values Predicts Genetic Constraint on Natural Selection on Landscapes Lacking Sign Epistasis. *Genetics*, 171(3):1397–1405, November 2005.

[11] David M. McCandlish. On the Findability of Genotypes. *Evolution*, 67(9):2592–2603, 2013. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/evo.12128.

[12] J. F. C. Kingman. A simple model for the balance between selection and mutation. *Journal of Applied Probability*, 15(1):1–12, March 1978.

[13] Stuart A. Kauffman and Edward D. Weinberger. The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology*, 141(2):211–245, November 1989.

[14] Su-Chan Park and Joachim Krug. Evolution in random fitness landscapes: the infinite sites model. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(04):P04014, April 2008. Publisher: IOP Publishing.

[15] David M. McCandlish. Long-term evolution on complex fitness landscapes when mutation is weak. *Heredity*, 121(5):449–465, November 2018.

[16] Sergey Kryazhimskiy, Gašper Tkačik, and Joshua B. Plotkin. The dynamics of adaptation on correlated fitness landscapes. *Proceedings of the National Academy of Sciences*, 106(44):18638–18643, November 2009.

[17] Walter Fontana and Peter Schuster. Shaping Space: the Possible and the Attainable in RNA Genotype–phenotype Mapping. *Journal of Theoretical Biology*, 194(4):491–515, October 1998.

[18] M. E. J. Newman and Robin Engelhardt. Effects of selective neutrality on the evolution of molecular species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1403):1333–1338, July 1998. Publisher: Royal Society.

[19] Matthew C. Cowperthwaite and Lauren Ancel Meyers. How Mutational Networks Shape Evolution: Lessons from RNA Models. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):203–230, 2007. _eprint: https://doi.org/10.1146/annurev.ecolsys.38.091206.095507.

[20] Ryan K. Shultzaberger, Daniel S. Malashock, Jack F. Kirsch, and Michael B. Eisen. The Fitness Landscapes of cis-Acting Binding Sites in Different Promoter and Environmental Contexts. *PLoS Genet*, 6(7):e1001042, July 2010.

[21] Marjon G. J. de Vos, Alexandre Dawid, Vanda Sunderlikova, and Sander J. Tans. Breaking evolutionary constraint with a tradeoff ratchet. *Proceedings of the National Academy of Sciences*, 112(48):14906–14911, December 2015.

[22] José Aguilar-Rodríguez, Joshua L. Payne, and Andreas Wagner. A thousand empirical adaptive landscapes and their navigability. *Nature Ecology & Evolution*, 1:0045, January 2017.

[23] A. S. Perelson and C. A. Macken. Protein evolution on partially correlated landscapes. *Proceedings of the National Academy of Sciences*, 92(21):9657–9661, October 1995. Publisher: National Academy of Sciences Section: Research Article.

[24] Tamar Friedlander, Roshan Prizak, Nicholas H. Barton, and Gašper Tkačik. Evolution of new regulatory functions on biophysically realistic fitness landscapes. *Nature Communications*, 8(1):216, August 2017.

[25] Andrea I. Collins-Hed and David H. Ardell. Match fitness landscapes for macromolecular interaction networks: Selection for translational accuracy and rate can displace tRNA-binding interfaces of aminoacyl-tRNA synthetases. *Theoretical Population Biology*, April 2019.

[26] Wilfred Ndifon, Joshua B. Plotkin, and Jonathan Dushoff. On the Accessibility of Adaptive Phenotypes of a Bacterial Metabolic Network. *PLOS Computational Biology*, 5(8):e1000472, August 2009. Publisher: Public Library of Science.

[27] Fernanda Pinheiro, Omar Warsi, Dan I. Andersson, and Michael Lässig. Metabolic fitness landscapes predict the evolution of antibiotic resistance. *Nature Ecology & Evolution*, pages 1–11, March 2021. Publisher: Nature Publishing Group.

[28] Peter F. Stadler. Landscapes and their correlation functions. *Journal of Mathematical chemistry*, 20(1):1–45, 1996. Publisher: Springer.

[29] Juannan Zhou and David Martin McCandlish. Minimum epistasis interpolation for sequence-function relationships. *bioRxiv*, page 657841, June 2019.

[30] Claudia Bank, Sebastian Matuszewski, Ryan T. Hietpas, and Jeffrey D. Jensen. On the (un)predictability of a large intragenic fitness landscape. *Proceedings of the National Academy of Sciences*, page 201612676, November 2016.

[31] J. Arjan G. M. de Visser and Joachim Krug. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics*, 15(7):480–490, July 2014.

[32] Takuyo Aita, Yuuki Hayashi, Hitoshi Toyota, Yuzuru Husimi, Itaru Urabe, and Tetsuya Yomo. Extracting characteristic properties of fitness landscape from in vitro molecular evolution: A case study on infectivity of fd phage to E.coli. *Journal of Theoretical Biology*, 246(3):538–550, June 2007.

[33] H. Allen Orr. The Population Genetics of Adaptation: The Adaptation of Dna Sequences. *Evolution*, 56(7):1317–1330, 2002. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0014-3820.2002.tb01446.x.

[34] Tamar Friedlander, Roshan Prizak, Călin C. Guet, Nicholas H. Barton, and Gašper Tkačik. Intrinsic limits to gene regulation by global crosstalk. *Nature Communications*, 7:12307, August 2016.

[35] Rok Grah and Tamar Friedlander. The relation between crosstalk and gene regulation form revisited. *PLOS Computational Biology*, 16(2):e1007642, February 2020. Publisher: Public Library of Science.

[36] Tamar Friedlander, Avraham E. Mayo, Tsvi Tlusty, and Uri Alon. Mutation Rules and the Evolution of Sparseness and Modularity in Biological Systems. *PLoS ONE*, 8(8), August 2013.

[37] T. Friedlander and N. Brenner. Adaptive response and enlargement of dynamic range. *Mathematical Biosciences and Engineering*, 8(2):515–528, 2011.

[38] Erik van Nimwegen, James P. Crutchfield, and Martijn Huynen. Neutral evolution of mutational robustness. *Proceedings of the National Academy of Sciences*, 96(17):9716–9720, August 1999.

[39] Stephen R. Proulx and Patrick C. Phillips. The Opportunity for Canalization and the Evolution of Genetic Networks. *The American Naturalist*, 165(2):147–162, February 2005. Publisher: The University of Chicago Press.

[40] Joanna Masel and Mark L. Siegal. Robustness: mechanisms and consequences. *Trends in Genetics*, 25(9):395–403, September 2009.

[41] Robert J. Woods, Jeffrey E. Barrick, Tim F. Cooper, Utpala Shrestha, Mark R. Kauth, and Richard E. Lenski. Second-Order Selection for Evolvability in a Large Escherichia coli Population. *Science*, 331(6023):1433–1436, March 2011. Publisher: American Association for the Advancement of Science Section: Report.

[42] Lee Altenberg. Fundamental Properties of the Evolution of Mutational Robustness. *arXiv:1508.07866 [math, q-bio]*, August 2015. arXiv: 1508.07866.

[43] Milo S. Johnson, Alena Martsul, Sergey Kryazhimskiy, and Michael M. Desai. Higher-fitness yeast genotypes are less robust to deleterious mutations. *Science*, 366(6464):490–493, October 2019.

[44] Matteo Smerlak. Quasi-species evolution maximizes genotypic reproductive value (not fitness or flatness). *Journal of Theoretical Biology*, 522:110699, August 2021.

[45] Manfred Eigen. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58(10):465–523, October 1971.

[46] Manfred Eigen and Peter Schuster. A principle of natural self-organization. *Naturwissenschaften*, 64(11):541–565, November 1977.

[47] Peter Schuster. Phase transitions, strong quasispecies, and neutrality. page 98.

[48] Claus O. Wilke. Quasispecies theory in the context of population genetics. *BMC Evolutionary Biology*, 5(1):44, August 2005.

[49] Jörg Swetina and Peter Schuster. Self-replication with errors: A model for polvnucleotide replication22This paper is considered as part II of Model Studies on RNA replication. Part I is the Gassner and Schuster [14]. *Biophysical Chemistry*, 16(4):329–345, December 1982.

[50] Josep Sardanyés, Santiago F. Elena, and Ricard V. Solé. Simple quasispecies models for the survival-of-the-flattest effect: The role of space. *Journal of Theoretical Biology*, 250(3):560–568, February 2008.

[51] Claus O. Wilke, Jia Lan Wang, Charles Ofria, Richard E. Lenski, and Christoph Adami. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, 412(6844):331–333, July 2001.

[52] Francisco M. Codoñer, José-Antonio Darós, Ricard V. Solé, and Santiago F. Elena. The Fittest versus the Flattest: Experimental Confirmation of the Quasispecies Effect with Subviral Pathogens. *PLOS Pathogens*, 2(12):e136, December 2006. Publisher: Public Library of Science.

[53] Rafael Sanjuán, José M. Cuevas, Victoria Furió, Edward C. Holmes, and Andrés Moya. Selection for Robustness in Mutagenized RNA Viruses. *PLOS Genetics*, 3(6):e93, June 2007. Publisher: Public Library of Science.

[54] Chuan Li, Wenfeng Qian, Calum J. Maclean, and Jianzhi Zhang. The fitness landscape of a tRNA gene. *Science*, 352(6287):837–840, May 2016.

[55] Chuan Li and Jianzhi Zhang. Multi-environment fitness landscapes of a tRNA gene. *Nature Ecology & Evolution*, 2(6):1025, June 2018.

[56] John H. Gillespie. *Population Genetics: A Concise Guide*. The Johns Hopkins University Press, 2nd edition, July 2004.

[57] Sasha F. Levy, Jamie R. Blundell, Sandeep Venkataram, Dmitri A. Petrov, Daniel S. Fisher, and Gavin Sherlock. Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*, 519(7542):181–186, March 2015. Number: 7542 Publisher: Nature Publishing Group.

[58] SGD-Saccharomyces cerevisiae Genome, Saccharomyces cerevisiae Genome Snapshot, database link retrieved June 22nd 2017. Bionumbers - BNID 100237, https://bionumbers.hms.harvard.edu/

[59] Joseph Schacherer, Joshua A. Shapiro, Douglas M. Ruderfer, and Leonid Kruglyak. Comprehensive polymorphism survey elucidates population structure of Saccharomyces cerevisiae. *Nature*, 458(7236):342–345, March 2009. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7236 Primary_atype: Research Publisher: Nature Publishing Group.

[60] Bryan P. Thornlow, Josh Hough, Jacquelyn M. Roger, Henry Gong, Todd M. Lowe, and Russell B. Corbett-Detig. Transfer RNA genes experience exceptionally elevated mutation rates. *Proceedings of the National Academy of Sciences*, 115(36):8996–9001, September 2018. Publisher: National Academy of Sciences Section: Biological Sciences.

[61] Natalie Saini, Steven A. Roberts, Joan F. Sterling, Ewa P. Malc, Piotr A. Mieczkowski, and Dmitry A. Gordenin. APOBEC3B cytidine deaminase targets the non-transcribed strand of tRNA genes in yeast. *DNA Repair*, 53:4–14, May 2017.

[62] Grant Kinsler, Kerry Geiler-Samerotte, and Dmitri A Petrov. Fitness variation across subtle environmental perturbations reveals local modularity and global pleiotropy of adaptation. *eLife*, 9:e61271, December 2020. Publisher: eLife Sciences Publications, Ltd.

[63] Gregory I. Lang and Andrew W. Murray. Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae. *Genetics*, 178(1):67–82, January 2008.

[64] Pengyao Jiang, Anja R. Ollodart, Vidha Sudhesh, Alan J. Herr, Maitreya J. Dunham, and Kelley Harris. A modified fluctuation assay reveals a natural mutator phenotype that drives mutation spectrum variation within Saccharomyces cerevisiae. *bioRxiv*, page 2021.01.11.425955, January 2021. Publisher: Cold Spring Harbor Laboratory Section: New Results.

[65] Mato Lagator, Srdjan Sarikas, Magdalena Steinrück, David Toledo-Aparicio, Jonathan P. Bollback, Gasper Tkacik, and Calin C. Guet. Structure and Evolution of Constitutive Bacterial Promoters. preprint, Evolutionary Biology, May 2020.

[66] Thomas A. Hopf, John B. Ingraham, Frank J. Poelwijk, Charlotta P. I. Schärfe, Michael Springer, Chris Sander, and Debora S. Marks. Mutation effects predicted from sequence co-variation. *Nature Biotechnology*, 35(2):128–135, February 2017. Number: 2 Publisher: Nature Publishing Group.

[67] Matteo Smerlak. Effective potential reveals evolutionary trajectories in complex fitness landscapes. *arXiv:1912.05890 [cond-mat, q-bio]*, December 2019. arXiv: 1912.05890.

[68] Shimon Bershtein, Adrian W. R. Serohijos, Sanchari Bhattacharyya, Michael Manhart, Jeong-Mo Choi, Wanmeng Mu, Jingwen Zhou, and Eugene I. Shakhnovich. Protein Homeostasis Imposes a Barrier on Functional Integration of Horizontally Transferred Genes in Bacteria. *PLOS Genetics*, 11(10):e1005612, October 2015. Publisher: Public Library of Science.

[69] Aisha I. Khan, Duy M. Dinh, Dominique Schneider, Richard E. Lenski, and Tim F. Cooper. Negative Epistasis Between Beneficial Mutations in an Evolving Bacterial Population. *Science*, 332(6034):1193–1196, June 2011. Publisher: American Association for the Advancement of Science Section: Report.

[70] Gautam Reddy and Michael M. Desai. Global epistasis emerges from a generic model of a complex trait. *bioRxiv*, page 2020.06.14.150946, June 2020. Publisher: Cold Spring Harbor Laboratory Section: New Results.

[71] Daniel R. Schrider, David Houle, Michael Lynch, and Matthew W. Hahn. Rates and Genomic Consequences of Spontaneous Mutational Events in Drosophila melanogaster. *Genetics*, 194(4):937–954, August 2013. Publisher: Genetics Section: Investigations.

[72] Yuan O. Zhu, Mark L. Siegal, David W. Hall, and Dmitri A. Petrov. Precise estimates of mutation rate and spectrum in yeast. *Proceedings of the National Academy of Sciences*, 111(22):E2310–E2318, June 2014.

[73] Rob W. Ness, Andrew D. Morgan, Radhakrishnan B. Vasanthakrishnan, Nick Colegrave, and Peter D. Keightley. Extensive de novo mutation rate variation between individuals and across the genome of Chlamydomonas reinhardtii. *Genome Research*, 25(11):1739–1749, November 2015. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab.

[74] Jacob J. Michaelson, Yujian Shi, Madhusudan Gujral, Hancheng Zheng, Dheeraj Malhotra, Xin Jin, Minghan Jian, Guangming Liu, Douglas Greer, Abhishek Bhandari, Wenting Wu, Roser Corominas, Aine Peoples, Amnon Koren, Athurva Gore, Shuli Kang, Guan Ning Lin, Jasper Estabillo, Therese Gadomski, Balvindar Singh, Kun Zhang, Natacha Akshoomoff, Christina Corsello, Steven McCarroll, Lilia M. Iakoucheva, Yingrui Li, Jun Wang, and Jonathan Sebat. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell*, 151(7):1431–1442, December 2012.

[75] KAREN M. Page and MARTIN A. Nowak. Unifying Evolutionary Dynamics. *Journal of Theoretical Biology*, 219(1):93–98, November 2002.

[76] Sean H. Rice. *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates is an imprint of Oxford University Press, Sunderland, Mass., USA, 1 edition edition, June 2004.

[77] David Masad and Jacqueline Kazil. MESA: an agent-based modeling framework. In *14th PYTHON in Science Conference*, pages 53–60, 2015.