The Onset of Evolutionary Stalling and the Limit on the Power of Natural Selection to Improve a Cellular Module

Sandeep Venkataram[1], Ross Monasky[2], Shohreh H Sikaroodi[1,3], Sergey Kryazhimskiy[1,*,†], Betül Kaçar[2,*,†]

[1]Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093;
[2]Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721;
[3]Current address: Molecular engineering group, Fate Therapeutics Inc.
[*]Equal contribution
[†]Corresponding authors: betul@arizona.edu, skryazhi@ucsd.edu

Abstract

Cells consist of molecular modules which perform vital biological functions. Modules are key units of adaptive evolution because organismal fitness depends on their performance. Yet, our understanding of adaptive evolution at the level of modules is limited. Theory predicts that in rapidly evolving populations, such as those of many microbes, natural selection focuses on improving one or a few modules at a time and its focus shifts to other modules as adaptation continues. Such shifts have never been directly observed, their timescale is unknown and the extent to which they limit the power of natural selection to improve any particular module is unclear. Here, we empirically characterize how natural selection improves the translation machinery (TM), one of the most essential cellular modules. To this end, we experimentally evolved populations of *Escherichia coli* with genetically perturbed TMs for 1,000 generations. Populations with different TMs embarked on statistically distinct adaptive trajectories. Yet, in all genetic backgrounds, the focus of natural selection shifted away from the TM before its performance was fully restored. Our results show that shifts in the focus of selection can occur on time scales comparable to those of environmental fluctuations. Variability in selection pressures can delay the resumption of adaptation in stalled modules, which would make it difficult for evolution to fully optimize even essential modules.

Introduction

Biological systems are organized hierarchically, from molecules to cells, organisms and populations [1–5]. At the lowest level, macromolecules form cellular modules, such as the translation machinery, or other metabolic pathways [4,6–9]. Different modules perform different cellular functions, which together determine the fitness of the organism. Populations adapt by accumulating beneficial mutations that modify module function, which in turn can profoundly change the physiology of organisms, allowing them to consume new resources [10,11] or become resistant to drugs [12–14]. The dynamics of evolution can be extremely complex, particularly in large populations with limited recombination [14–20]. Much progress has been made recently in our theoretical and empirical characterization of these dynamics at the genetic level [15,18,19,21–24], but it is a major challenge to understand how evolution unfolds at the level of functional cellular modules [25].

The classical population genetics models predict that the speed of module evolution depends on the supply and the fitness effects of mutations in that module alone [26], and modules evolve independently of each other. Natural selection cannot improve a module if the fitness benefits of all beneficial mutations in it are below ~1/N, the inverse of the population size [27–31]. More recent work suggests that adaptive evolution of a module may reach a steady state where the fixation of beneficial mutations is counterbalanced by the accumulation of deleterious mutations [32,33]. Regardless of the nature of the limit, some cellular modules may be fundamentally not improvable by natural selection. However, classical models predict that all improvable modules would adapt, albeit at module-specific rates.

In many populations, particularly in microbes, beneficial mutations are common and recombination is rare, which violates the assumptions of classical models [14,16,17,20]. Then, multiple new mutations affecting different modules may simultaneously arise, they compete with each other and prevent each other from fixing [21–23,34,35]. In this so-called "clonal interference" regime, small-effect mutations are usually outcompeted. Instead, adaptation is driven by mutations that provide fitness benefits above a certain "clonal interference" threshold, which depends on the supply and the fitness effects of all adaptive mutations in the genome [23,34,36]. Thus, modules do not evolve independently, and the rate of adaptation in any one module depends on the supplies and effects of beneficial mutations in all modules.

The dynamics of module evolution in the clonal interference regime can be qualitatively different from those in the absence of clonal interference. At any given time, modules where many mutations provide fitness benefits above the current clonal interference threshold will accumulate adaptive mutations. All other modules would not adapt, that is modules both modules that are not improvable and modules that are improvable but only by mutations with fitness effects below the clonal interference threshold. We refer to this phenomenon as "evolutionary stalling".

Evolutionary stalling limits the power of natural selection to improve a module. However, in contrast to the hard limits imposed by the drift barrier or by the balance between beneficial and deleterious mutations, evolutionary stalling can be overcome. Once the supply of the adaptive mutations with the largest effects is diminished, the clonal interference threshold would drop. As a result, previously stalled modules may come into the focus of natural selection, while those that were previously adapting may in turn stall. Such shifts in the focus of natural selection are expected to occur as long as the supplies and the fitness effects of adaptive mutations are sufficiently

variable between modules. However, because these quantities are unknown, it is unclear whether the theoretically predicted shifts can occur on faster time scales than environmental fluctuations and to what extent they limit the power of natural selection to improve any particular module.

Numerous previous studies have demonstrated that selection is often focused on improving a relatively small number of cellular modules, in both natural and experimental populations [15,37–54], suggesting that adaptive evolution in other modules is stalled. However, these studies do not attempt to identify the time scale on which the focus of natural selection shifts between modules nor do they inform us whether these shifts slow down module improvements.

A shift in the focus of selection is associated with the onset of evolutionary stalling in one or multiple modules and can be detected in two ways. If we can directly measure the physiological performance of a module over time, an abrupt reduction in the rate of its phenotypic improvement despite steady increases in fitness would indicate the onset of stalling. However, it is often unclear which aspects of a module's performance are relevant for fitness. Alternatively, if we know all the genes that encode a module, we could infer the onset of stalling from an abrupt reduction in the rate of accumulation of mutations in such genes despite continued accumulation of beneficial mutations elsewhere in the genome. A recent study of a 60,000 generation long evolution experiment in *Escherichia coli* (*E. coli*) used such genomic approach and found that the statistical distribution of mutations among genes and operons changes over time [51]. This observation implies that the focus of natural selection shifts from some genes and operons to others. However, because most cellulars modules are insufficiently well annotated, one cannot rule out the possibility that all evolving genes encode the same set of modules, such that the shifts in the focus of natural selection occur within but not between modules.

To overcome these difficulties and characterize the onset of evolutionary stalling, we experimentally examine the evolution of the translation machinery (TM) in *E. coli*. There are two reasons for this choice. First, TM is an essential component of every living cell, and TM performance is a major component of fitness. Second, TM is encoded by an extremely well annotated set of genes [8,9,55–57], which allows us to use the genomic approach for detecting evolutionary stalling. We disrupted the TM by replacing the native Elongation Factor Tu (EF-Tu) in *E. coli* with several of its orthologs [58–60] and evolved these strains in rich media where rapid and accurate translation is required for fast growth [61,62]. We hypothesized that more severe disruptions of the TM will increase the supply of large-effect beneficial mutations in the TM [63]. Thus, we expect that natural selection would focus on improving the TM in at least some strains with disrupted TMs, but not in the control strain which carries the native *E. coli* EF-Tu. We then set out to characterize the onset of evolutionary stalling in the TM in two ways. First, we determine which strains acquire substitutions in the known TM genes and how much fitness these strains gain. This allows us to quantify whether natural selection is able to fully restore TM performance before its focus shifts to other cellular modules. Second, we observe how the rate of accumulation of mutations in the TM changes over time, which provides us with direct evidence for the onset of evolutionary stalling on short time scales relevant for the evolution in natural populations.

Results

We previously replaced the native EF-Tu in *E. coli* with its orthologs from *Salmonella typhimurium*, *Yersinia enterocolitica*, *Vibrio cholerae* and *Pseudomonas aeruginosa* and one reconstructed ancestral variant [58] (Table 1). EF-Tu is encoded in *E. coli* by two paralogous genes, *tufA* and *tufB*, with the majority of the EF-Tu molecules being expressed from *tufA* [64]. To replace all EF-Tu molecules in the cells, the *tufB* gene was deleted and the foreign orthologs were integrated into the *tufA* locus [58]. We also included the control strain in which the *tufB* gene was deleted and the original *E. coli tufA* was left intact. We refer to the engineered "founder" *E. coli* strains as E, S, Y, V, A and P by the first letter of the origin of their *tuf* genes (Table 1).

Strain	EF-Tu origin species	Number of amino acid differences from *E. coli* EF-Tu (percent identity)	Fitness ± SEM, % per generation
E	*Escherichia coli* (control)	0 (100)	0 ± 0.7
S	*Salmonella typhimurium*	1 (99.75)	+0.49 ± 0.09
Y	*Yersinia enterocolitica*	24 (93.91)	–3.02 ± 0.03
V	*Vibrio cholerae*	51 (87.06)	–19.0 ± 1.1
A	Reconstructed ancestor	21 (94.67)	–34.4 ± 0.7
P	*Pseudomonas aeruginosa*	62 (84.38)	–35.0 ± 0.2

Table 1. Founders used for the evolution experiment. Strains with foreign EF-Tu orthologs are ordered by their fitness relative to the control E strain. SEM stands for standard error of the mean.

We first quantified the TM defects in our founder strains. Kaçar et al. showed that EF-Tu replacements lead to declines in the *E. coli* protein synthesis rate and proportional losses in growth rate in the rich laboratory medium LB [58]. In our subsequent evolution experiment, natural selection will favor genotypes with higher competitive fitness, which may have other components in addition to growth rate [65–69]. We confirmed that EF-Tu replacements caused changes in competitive fitness relative to the control E strain (Table 1), and that competitive fitness and growth rate were highly correlated (Figure S1). We conclude that the competitive fitness of our founders in our environment reflects their TM performance. Further, we found that the fitness of the S and Y founders were similar to that of the control E strain (≤ 3% fitness change) indicating that their TMs were not substantially perturbed. In contrast, the fitness of the V, A and P founders were dramatically lower (≥ 19% fitness decline; Table 1) indicating that their TMs were severely perturbed.

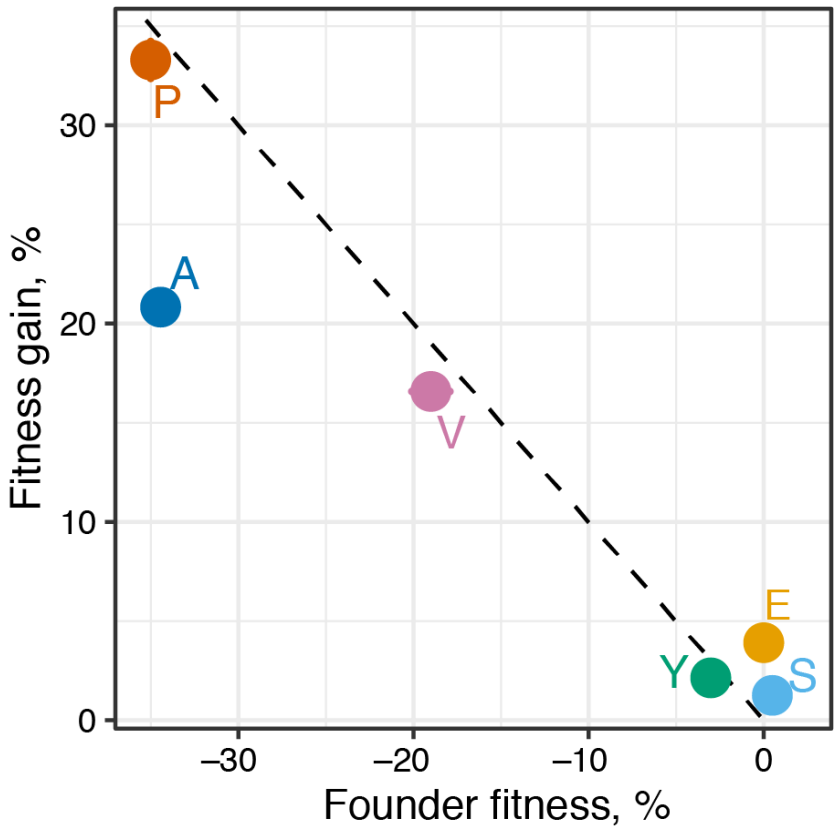


Figure 1. Competitive fitness of founder and evolved populations. The competitive fitness gain after evolution relative to the unevolved founder averaged across replicate populations (y axis) is plotted against the competitive fitness of the founder relative to the E strain (x axis). Fitness is measured in % per generation. Dashed black line is $y = -x$. Populations above (below) this line are more (less) fit than the control E strain, under the assumption that fitness is transitive. Error bars showing ±1 SEM are masked by the symbols (see Table 1 and Figure S2).

Clonal interference slows down TM adaptation

To determine whether natural selection focuses on restoring defective TMs, we instantiated 10 replicate populations from each of our six founders (60 populations total) and evolved them in LB for 1,000 generations (Methods) with daily 1:10^4 dilutions and the bottleneck population size $N = 5\times10^5$ cells. We then measured the competitive fitness of the evolved populations relative to their respective founders. Fitness in all but one population increased significantly (t-test $P < 0.05$ after Benjamini-Hochberg correction; Figure S2), and the average fitness increase of a population correlated negatively with the initial fitness of its founder (Figure 1). These results show that even substantial fitness defects caused by reductions in TM performance can be largely compensated in a short bout of adaptive evolution.

The pattern of "declining adaptability" in Figure 1 has been frequently observed in previous microbial evolution studies [45,46,50,70–74]. It could arise if adaptation is driven either by mutations only in the TM, by mutations only in other modules, or by mutations in the TM and in other modules. For evolutionary stalling in the TM to occur, mutations improving the TM must compete against other types of mutations within the same population. To determine whether both types of mutations occur in our populations, we conducted whole-population whole-genome

sequencing at multiple timepoints throughout the evolution experiment. This sequencing strategy allows us to directly observe competition dynamics between mutations in different modules [15,51,75,76].

We selected replicate populations 1 through 6 descended from each founder (a total of 36 populations), sampled each of them at 100-generation intervals (a total of 11 time points per population) and sequenced the total genomic DNA extracted from these samples. We developed a bioinformatics pipeline to identify de novo mutations in this data set (Methods). Then, we called a mutation adaptive if it satisfied two criteria: (i) its frequency changed by more than 20% in a population; and (ii) it occured in a "multi-hit" gene, i.e., a gene in which two independent mutations passed the first criterion. Reliably tracking the frequencies of some types of mutations (e.g., large copy-number variants) is impossible with our sequencing approach. Therefore, we augmented our pipeline with the manual identification of copy-number variants which could only be reliably detected after they reached high frequency in a population (Methods and Figure S3).

This procedure yielded 167 new putatively adaptive mutations in 28 multi-hit genes, with the expected false discovery rate of 13.6%, along with an additional 11 manually-identified chromosomal amplifications, all of which span the *tufA* locus (Methods and Table S1, Figure S4). We classified each putatively adaptive mutation as "TM-specific" if the gene where it occurred is annotated as translation-related (Methods). We classified mutations in all other genes as "generic". We found that 38 out of 178 (21%) putatively adaptive mutations in 6 out of 28 multi-hit genes were TM-specific (Table S1). This is significantly more mutations than expected by chance ($P < 10^{-4}$, randomization test) since the 215 genes annotated as translation-related comprise only 4.0% of the *E. coli* genome. All of the TM-specific mutations occurred in genes whose only known function is translation-related, such as *rpsF* and *rpsG*, suggesting these mutations arose in response to the initial defects in the TM. The set of TM-specific mutations is robust with respect to our filtering criteria (Figure S5).

TM-specific mutations occurred in 17 out of 36 sequenced populations. Generic mutations were also observed in all of these populations (Figure S4). Thus, whenever TM-specific mutations occurred, generic mutations also occurred, such that the fate of TM-specific mutations must have depended on the outcome of clonal interference between mutations within and between modules (Figure 2). As a result of this competition, only 14 out of 27 (52%) TM-specific mutations that arose (excluding 11 *tufA* ampliciations) went to fixation, while the remaining 13 (48%) succumbed to clonal interference (Figures 2, S4). In at least two of these 13 cases a TM-specific mutation was outcompeted by expanding clones likely driven by generic mutations: in population V6, a TM-specific mutation in *fusA* was outcompeted by a clone carrying generic mutations in *fimD*, *ftsI* and *hslO* (Figure 2); in population P3, a TM-specific mutation in *tufA* was outcompeted by a clone carrying generic mutations in *amiC* and *trkH* (Figure 2). We conclude that, while TM-specific beneficial mutations are sufficiently common and their fitness effects are at least sometimes large enough to successfully compete against generic mutations, clonal interference reduces the power of natural selection to recover TM performance.

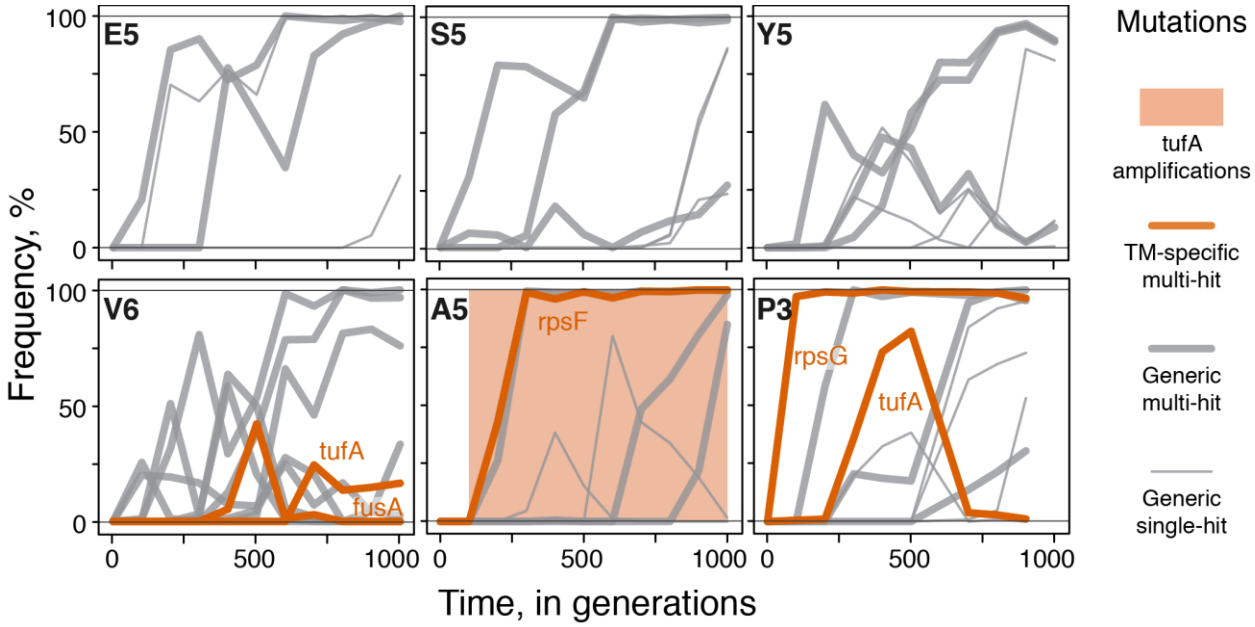


Figure 2. Mutational trajectories in evolving populations. Mutation frequency trajectories for one representative replicate population per founder is shown (complete data for all sequenced populations can be found in Figure S4). Each line represents the frequency trajectory of a single mutation. Shading indicates the range of timepoints in which a *tufA* amplification was detected.

The onset evolutionary stalling and a limit on the power of natural selection to improve the TM

Competition between adaptive mutations in different modules is necessary but not sufficient for evolutionary stalling to occur in any one module. Therefore, we sought direct evidence of evolutionary stalling in the TM. To this end, we examined the distribution of TM-specific mutations among populations derived from different founders.

In the A and P founders, defects in the TM incur a fitness cost of about 35% (Table 1), and we observed a total of 30 TM-specific mutations in these populations (Figure 3A), including eight *tufA* ampliciations (2.5 mutations on average per population). At least one TM-specific mutation fixed in each of these populations (Figure S4). Thus, when the TM defect is large, highly beneficial mutations in the TM are available, and natural selection is focused on improving this module.

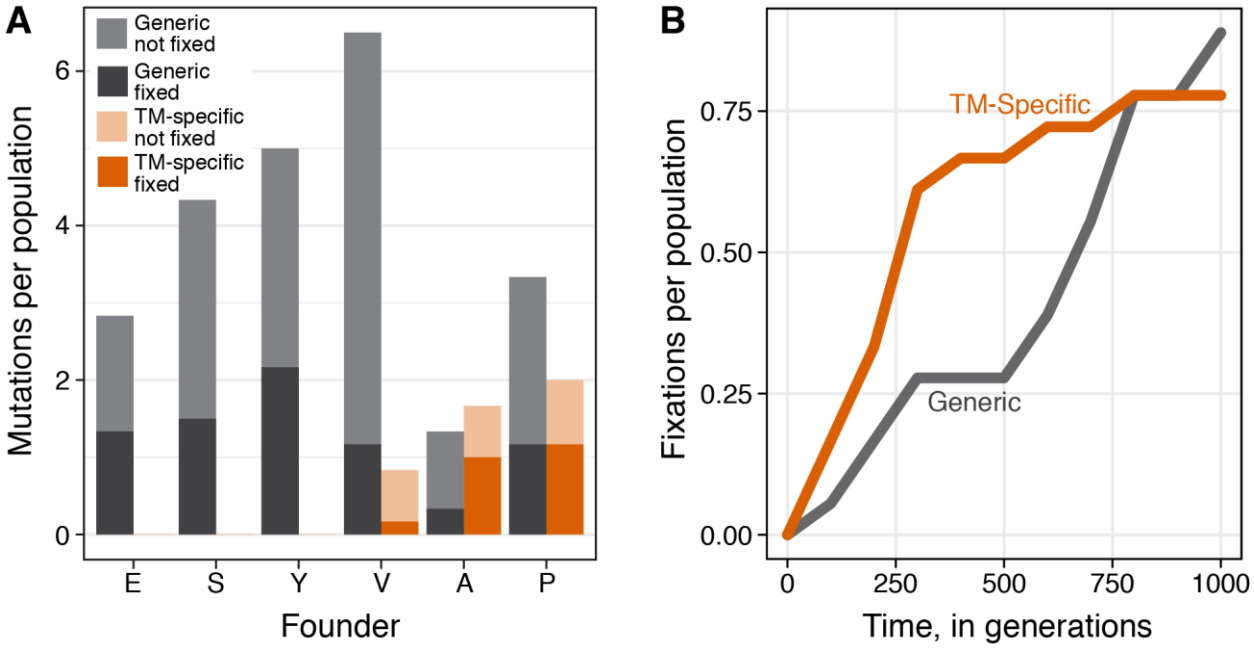


Figure 3. Evidence for evolutionary stalling in the adaptation of the TM. A. The average number of fixed and not fixed adaptive TM-specific and generic mutations across the six sequenced populations derived from each founder. **B.** Cumulative number of fixed TM-specific or generic mutations per population derived from the V, A and P founders.

In the V founder, a defect in the TM incurs a fitness cost of about 19% (Table 1), and we observed a total of 8 TM-specific mutations across all V populations, including three *tufA* amplifications (1.3 mutations on average per population). However, only 2 out of 6 populations fixed at least one TM-specific mutation (Figure S4). Thus, TM defects with a fitness cost of 19% are improvable by mutations, but these mutations provide benefits that are barely enough to survive clonal interference. We conclude that when a TM defect incurs a fitness cost of around 19%, the focus of natural selection begins to shift to other cellular modules, and the adaptation of the TM begins to stall.

Given that the TM adaptation is beginning to stall in the V populations, we predict that no TM-specific mutations would fix in the E, S and Y populations, as their TM defects are much smaller. Consistent with this prediction, we do not observe any TM-specific mutations in the E, S and Y populations (Figure 3A). Interestingly, the Y populations gained on average 2.1% in fitness by fixing on average 2.2 generic mutations. This implies that natural selection in the Y populations is improving modules other than the TM, even though their TM incurs a ~3% fitness cost (Table 1).

We also expect that as the V, A, and P populations accumulate TM-specific mutations and TM performance improves, the focus of natural selection should shift away from the TM to other cellular modules. We examined the distribution of TM-specific and generic adaptive mutations across evolutionary time in these populations. Out of the 14 TM-specific mutations that eventually fixed, 12 (86%) did so in the first selective sweep (this excludes 11 *tufA* amplifications). In contrast, out of the 16 generic mutations that fixed, only seven (44%) did so in the first selective sweep. As a result, an average TM-specific beneficial mutation reached fixation after only 300 ± 52 generations, compared to 600 ± 72 generations for an average generic mutation (Figure 3B, S4). Only one (7%) TM-specific beneficial mutation reached

fixation after generation 600, in comparison to nine (56%) generic beneficial mutations. Thus, by the end of the experiment, the fixation of TM-specific mutations has essentially ceased. Yet, these populations remained on average ~5.3% less fit than the control E strain, assuming fitness is transitive (Figure 1). Even if we conservatively attribute all these fitness gains to improvements in the TM, this observation suggests that the evolved TMs have not reached the level of performance of the TM in the control E founder. As with the Y founder, we conclude that the TMs in these evolved V, A and P populations has stalled before TM performance was fully recovered.

For adaptive mutations in the TM to fix, they must provide fitness benefits above the clonal interference threshold. To determine where this threshold is in P and A populations which fixed most of the TM-specific mutations, we selected two TM-specific mutations that arose in these populations, genetically reconstructed them in their respective founders and directly measured their fitness benefits. We found that the TM-specific mutation A74G in the *rpsF* gene, which arose in population A5, provides an 8.2 ± 1.0% fitness benefit in the A founder. And the TM-specific mutation G331A in gene *rpsG*, which arose in populations P2, P3 and P5, provides a 6.5 ± 1.2% fitness benefit in the P founder. This result suggests that the clonal interference threshold in the A and P populations is below 8.2% and 6.5%, respectively. Similarly, by reconstructing mutation T193D in the *ybeD* gene, we estimated that the clonal interference threshold in the V populations was below 5.9%. Finally, the upper bound on the clonal interference threshold in the Y, S, and E populations can be estimated from the total fitness gains in these populations (Figure 1), and it is below 2.1%, 1.3% and 3.9%, respectively.

Epistasis and historical contingency in TM evolution

We observed that natural selection improved all severely perturbed TMs, but it is unclear whether different TM defects can be alleviated by a common set of mutations or whether repairing each TM defect requires its own unique genetic solution. Previous work has shown that genetic interactions (or "epistasis") between mutations in the TM have been important in the evolutionary divergence of TMs along the tree of life [58,77–80]. We reasoned that genetic interactions might be similarly important in the short bout of evolution observed in our experiment. Specifically, we asked whether different initial TM variants acquired adaptive mutations in the same or in different translation-associated genes.

We found that 4 out of 7 classes of TM-specific mutations arose in a single founder (Figure 4A). For example, we detected six independent mutations in the *rpsG* gene, which encodes the ribosomal protein S7, and all of these mutations occurred in the P founder ($P < 10^{-4}$, randomization test with Benjamini-Hochberg correction, Methods). Similarly, all four mutations in the *rpsF* gene, which encodes the ribosomal protein S6, occurred in the A founder ($P < 10^{-4}$, randomization test with Benjamini-Hochberg correction). To directly measure how the effects of these mutations vary across genetic backgrounds, we attempted to genetically reconstruct mutation A74G in the *rpsF* gene and mutation G331A in *rpsG* gene in all six of our founder strains. We successfully reconstructed both of these mutations in the founder strains in which they arose and confirmed that they were strongly beneficial, as described above (8.2 ± 1.0% and 6.5 ± 1.2% benefit, respectively). In contrast, our multiple reconstruction attempts in all other founders were unsuccessful (Methods), suggesting that these mutations are strongly deleterious in all other genetic backgrounds that we tested.

These results suggest that genetic interactions between different TM components cause initially different TM variants to embark on divergent adaptive trajectories and lead to historical contingency and entrenchment in TM evolution [81–83].

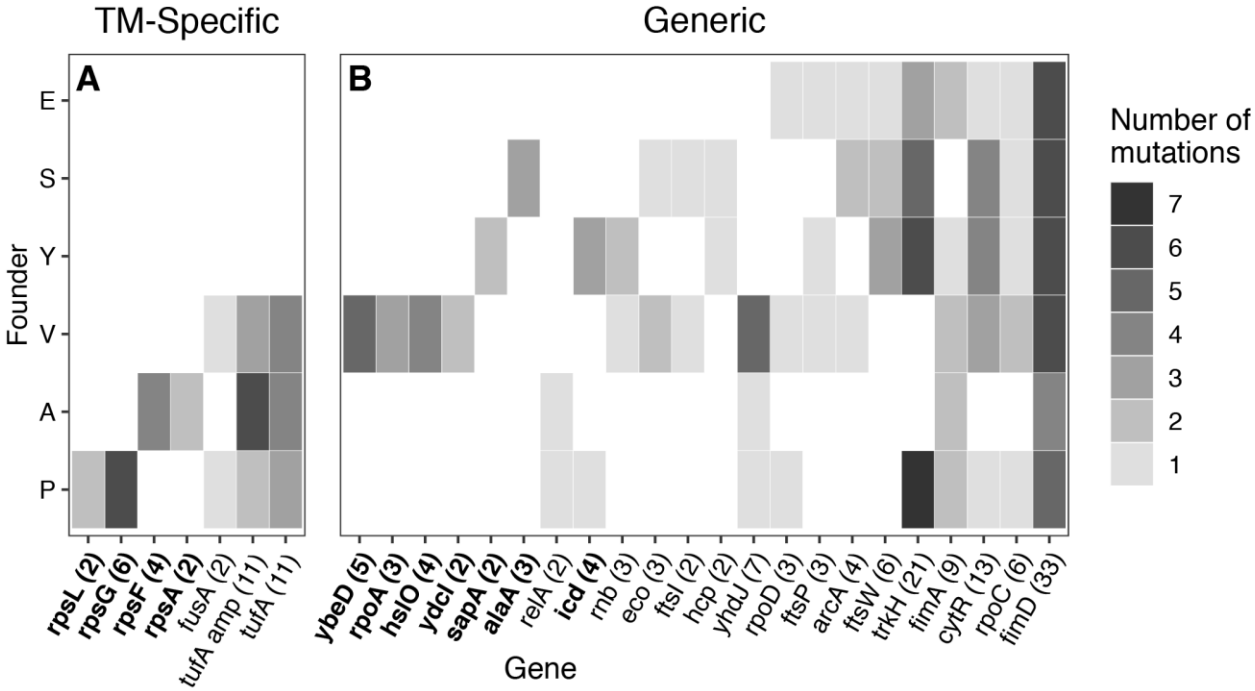


Figure 4. Distribution of putatively adaptive mutations. Heatmap of all putatively adaptive mutations identified via whole-genome sequencing, grouped by founder and by gene. Amplification of the *tufA* locus are counted separately from other mutations in *tufA*. **A.** Translation-associated genes **B.** All other genes. Genes in bold are those where mutations were detected in significantly fewer founders than expected by chance ($P < 0.05$, Benjamini-Hochberg correction, see Methods). Numbers in parentheses indicate the total number of mutations in that gene observed across all sequenced populations.

Genome-wide adaptive responses to TM perturbations

Adaptive evolution of the TM stalls because natural selection acts on multiple cellular modules in *E. coli*, all of which are encoded on a single non-recombining chromosome. However, modules are linked not only physically by the encoding DNA but also functionally in that they all contribute to the fitness of the organism. This functional interdependence implies that mutations in one module may alter the selection pressure on other modules. For example, improvements in translation efficiency may increase the selection pressure to improve efficiency of catabolic reactions, analogously to the "shifting and swaying of selection coefficients" on enzymes in the same metabolic pathway discussed in the classic work by Hartl et al. [27]. Therefore, in addition to intra-module epistasis demonstrated above we might expect inter-module epistasis, such that initially different TM variants could precipitate distinct adaptive responses in the rest of the genome. To test this hypothesis, we examined the distribution of generic mutations among founder genotypes.

We found that generic mutations in 7 out of 22 genes occurred in fewer founders than expected by chance (Figure 4B, Methods). For example, we detected five independent mutations in the *ybeD* gene, which encodes a protein with an unknown function, and all these mutations occurred in the V founder ($P < 10^{-4}$, randomization test with Benjamini-Hochberg correction). Similarly,

all three mutations in the *alaA* gene, which encodes a glutamate-pyruvate aminotransferase, occurred in the A founder ($P < 10^{-4}$, randomization test with Benjamini-Hochberg correction). To corroborate these statistical observations, we reconstructed the T93G mutation in the *ybeD* gene in all six founder strains and directly measured its fitness effects. As expected, this mutation confers a 5.9% fitness benefit in the V founder. In contrast, it is strongly deleterious in the P founder and indistinguishable from neutral in the remaining founders (Figure 5). These results show that at least some genetic perturbations in the TM can have genome-wide repercussions. They can precipitate bouts of genome-wide adaptive evolution that are contingent on the initial perturbations in the TM.

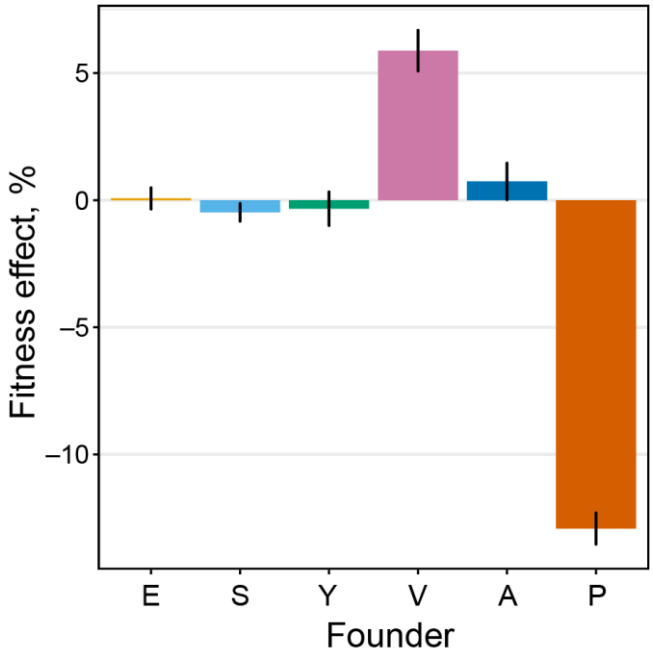


Figure 5. Fitness effect of the *ybeD* T193D mutation in different founders. Fitness effect of the mutation is measured in a direct competition of each founder with the mutation against the founder without the mutation. Error bars show the SEM.

Discussion

The fitness of an organism depends on the performance of many molecular modules inside cells. While natural selection favors genotypes with better-performing modules, it is difficult for evolution to improve multiple modules simultaneously, particularly when recombination rates are low and many adaptive mutations in different modules are available. In this regime, natural selection is expected to focus on those modules where mutations provide large fitness benefits, while the adaptive evolution in other modules stalls. Here we have documented and characterized the onset of evolutionary stalling in the adaptation of the translation machinery (TM) in *E. coli*.

We found that populations whose TMs were initially severely perturbed (incurring ~35% fitness cost) rapidly discovered and fixed TM-specific beneficial mutations. Populations whose TM had a intermediately strong defect (incurring ~19% fitness defect), discovered TM-specific mutations, but only some of these populations were able to fix them due to clonal interference. Thus, adaptive evolution of the TM begins to stall when the TM defect incurs a fitness cost of about 19%. As a consequence, populations whose TMs were initially mildly perturbed (incurring $\lesssim$ 3% fitness cost) adapted by acquiring mutations that did not directly affect the TM. Furthermore, even in populations where TM-specific mutations occurred, their fixations essentially ceased by the end of the experiment, while mutations in other modules continued to accumulate. Our results imply that the focus of natural selection shifted from the TM to other modules in the cell within ~600 generations (about 1 to 2 fixation events), before the performance of the original unperturbed TM was fully recovered.

There could be two explanations for why the focus of natural selection shifts away from the TM, despite it still being defective. First, there may simply be no more TM-specific beneficial mutations available, i.e., natural selection may have reached its limit on TM improvement. Alternatively, TM-specific beneficial mutations are available but provide fitness benefits that are too small to overcome clonal interference. We favor the latter explanation because both theoretical and empirical (albeit limited) evidence suggest that small-effect beneficial mutations should be more common than large-effect mutations [24,84–86]. The fact that we observed TM-specific mutations with effects $\geq$ 5% indicates that the rate of such mutations is relatively high, and we can expect the rate of TM-specific mutations with effects < 5% to be even higher. In fact, if we relax the stringency criteria for detecting beneficial mutations, we find one TM-specific mutation in the gene *rbbA* in the population E5 (Figure S5). This suggests that small-effect TM-specific beneficial mutations do arise and supports the conjecture that adaptation of the TM stalled rather than stopped.

Our results give us a glimpse of the fitness landscape of the TM. This landscape appears to be broadly consistent with Fisher's geometric model [85,87,88] in that more defective TMs have access to beneficial mutations with larger fitness benefits than less defective TMs. However, Fisher's model does not inform us how many distinct genotypes encode highly performing TMs and how they are connected in the genotype space. We observed that the different founders gained distinct TM-specific adaptive mutations. This suggests that the high-performance TMs can be encoded by multiple genotypes that either form a single contiguous neutral network [89] or multiple isolated neutral networks [90]. Moreover, we observed that most of our populations with initially severely perturbed TMs were able to discover TM-specific mutations. This

suggests that genotypes that encode high-performing TMs may be present in the mutational neighborhoods of many genotypes [89,91].

In this work, we identified several TM-specific adaptive mutations, but their biochemical and physiological effects are at this point unknown. However, the fact that 11 chromosomal amplifications and 12 noncoding or synonymous events occurred in the *tufA* operon suggests that some of the TM-specific mutations are beneficial because they adjust EF-Tu abundance in the cell. This would be consistent with previous evolution experiments [42,92,93]. Directly measuring the phenotypic effects of the TM-specific mutations described here is an important avenue for future work.

Our results show that evolutionary stalling limits the ability of natural selection to improve a module, but this limit is not absolute. As a population accumulates beneficial mutations in other modules, their supply will be depleted and their fitness effects will likely decrease due to diminishing returns epistasis [50,63,70,71,94,95]. These changes will in turn increase the chances for small-effect mutations in the focal module to survive clonal interference thereby overcoming evolutionary stalling. While we did not observe resumption of adaptive evolution in the TM in this experiment, we find some evidence for such a transition in one other module. We detected 11 mutations in multi-hit genes that affect cytokinesis (Methods, Figures S6, S7). Most of these mutations reached high frequency in the second half of the experiment, suggesting that adaptation in the cytokinesis module was initially stalled and then resumed.

Although adaptive evolution of a stalled module can resume once large-effect mutations in other modules are depleted, variability in selection pressures can replenish the supply of these competing mutations and thereby leave the focal module stalled for long periods of time. Given that populations in nature typically experience fluctuation in selective pressures, natural selection may not be able to fully optimize any module even over long evolutionary timescales.

Our results imply that it is impossible to fully understand the evolution a cellular module in isolation from the genome where it is encoded and the population-level processes that govern evolution. The ability of natural selection to improve any one module depends on the population size, the rate of recombination, the supply and the fitness effects of all beneficial mutations in the genome and on how these quantities change as the population adapts. Further theoretical work and empirical measurements integrated across multiple levels of biological organization are required for us to understand adaptive evolution of modular biological systems.

Materials and methods

Materials, data and code availability

All strains and plasmids constructed and used in this work are available per request. Raw sequencing data were analyzed with the python-based workflow implemented in Ref. [51] and run on the UCSD TSCC computing cluster via a custom python wrapper script. All analysis and plots reported in this manuscript have been performed using the R computing environment. The script, modified reference genomes and the raw data (except for raw sequencing data) used for analysis can be found at https://github.com/sandeepvenkataram/EvoStalling. Raw sequencing data for this project have been deposited into the NCBI SRA under project PRJNA560969.

Media and culturing conditions

Liquid medium is the Luria-Bertani medium (LB) (per liter, 10 g NaCl, 5 g yeast extract, and 10 g tryptone) and solid medium is LBA (LB with 1.5% agar), unless noted otherwise. All incubations were done at 37°C, and liquid cultures were shaken at 200 rpm for aeration, unless noted otherwise. All media components and chemicals were purchased from Sigma, unless noted otherwise.

Strains and plasmids

All strains in this study were derived from *E. coli* K12 MG1655. Strain genotypes are listed in Table S2. Complete methods for the construction of the E, S, Y, V, A and P strains, which harbor a single *tuf* gene variant replacing *tufA* gene, can be found in Ref. [58]. Strains with engineered *ybeD*, *rpsF* and *rpsG* mutations were constructed using the same method, except the chromosomal *kanR* marker was not removed (Figure S9). For a full list of primer sequences used for *ybeD*, *rpsF* and *rpsG* engineering, see Table S2.

Plasmids pZS1-TnSL and pZS2-TnSL were used in competition assays to provide Ampicillin and Kanamycin resistance, respectively. pZS1-TnSL, derived from pUA66 [96], was kindly provided by Georg Rieckh. pZS2-TnSL was constructed from pZS1-TnSL by replacing the *ampR* cassette with *kanR*.

Evolution experiment

Experimental evolution was performed by serial dilution at 37°C in LB broth. To start the evolution experiment, an initial 5 mL overnight culture was inoculated from a single colony from the frozen stock of each founder strains. 10 replicate populations were started from single colonies derived from these overnight cultures. The replicates were serially transferred every 24h (±1h) as follows: 100 µL of saturated culture were transferred into 10 mL saline solution (145 mM NaCl), 50 µL of these dilutions were then transferred to 5 mL fresh LB (tubes were vigorously vortexed prior to pipetting). This resulted in a bottleneck population size of about 5×10^5 cells. Freezer stocks (200 µL of 20% glycerol + 1 mL saturated culture) were prepared approximately every 100 generations and stored at –80°C.

Competitive fitness assays

To carry out pairwise competition assays, an Ampicillin-resistant and a Kanamycin-resistant versions of the query and reference strains/populations were generated by transforming these strains/populations with plasmids pZS1-TnSL and pZS2-TnSL, using standard methods [97]. Two replicate competition assays were performed for each query-reference pair with reciprocal markers (four assays total per pair), except for allele-replacement mutants (see below). To validate that the resistance-marker plasmids do not differentially impact fitness in any of the six founder genetic backgrounds, we carried out three-way competition assays between the KanR-marked, AmpR-marked and the unmarked versions of the founders (Figure S8). Since the allele-replacement mutants carry a chromosomal *kanR* marker (see above), they were only competed against AmpR reference strains.

To start a competition assay, a query and a reference cultures were scraped from frozen stocks and inoculated into 5 mL LB-Amp or LB-Kan media as appropriate. After about 24 hours, the query and the reference cultures were mixed together in ratio 1:9 and diluted 1:10,000 into 5 mL fresh LB media. After that, the mixed culture was propagated as in the evolution experiment. To determine the relative abundances of the query and reference individuals in the mixed culture, 100 µl of appropriately diluted cultures were plated on both LB-Amp and LB-Kan plates after 24, 48 and 74 hours of competition. For some competitions, where fitness differences were particularly large or small, samples from 0 or 96 hours were also obtained. Plates were photographed after an ~24-hour incubation period (when colonies were easily visible) and colonies were automatically counted with the OpenCFU software [98]. In each competition, we estimated the fitness of the query strain relative to the reference strain by linear regression of the natural logarithm of the ratio of the query to reference strain dilution-adjusted colony counts against time. Variance was also estimated from these regressions. Replicate measurements were combined into the final estimate using the inverse variance weighting method.

Competitions between two reciprocally marked versions of the same strain represent a special case. If the two marker-carrying plasmids impose exactly the same fitness cost, our competition assay between two reciprocally marked versions of the same strain is fully symmetric, which implies that in expectation it must yield a fitness value of exactly zero. Any estimate of fitness from a finite number of measurements even in such idealized fully symmetric case will not zero. However, such deviations from zero would reflect only measurement noise rather than any biologically meaningful fitness difference. In reality, the two marker-carrying plasmids may impose slightly different fitness costs, but because the difference in the cost is detectable (see above), we still interpret deviations from zero in our fitness estimates as noise. Therefore, in competitions of reciprocally marked versions of the same strain, we set our fitness estimate to zero and use the four fitness values obtained from the replicate assays to estimate the noise variance as the average of the squared fitness value.

Growth rate assays

Strains were inoculated from frozen stock into 5 mL LB media in 15 mL culture tubes and grown overnight. After 24 hours of growth, the cultures were diluted 1:100 into fresh 5 mL of media and grown for 4 hours. They were diluted again 1:100 into 200 µl of LB in flat-bottom Costar 96 well microplates (VWR Catalog #25381-056) and grown in a Molecular Devices SpectraMax i3x

Multi-mode microplate reader at 37°C with shaking for 24 hours with absorbance measurements at 600 nm every 15 minutes. Three replicate growth measurements were conducted for each strain. Optical density data were first ln-transformed. A linear regression model was fit to all sets of 5 consecutive data points where OD was below 0.1. Growth rate for the culture was estimated as the maximal slope across all of these 5-point regressions. The mean growth rate and standard error of the mean were calculated from replicate measurements.

Genome sequencing

Whole-genome sequencing was conducted for population samples of 6 replicate populations for each of the 6 founders (36 total populations). Each population was sequenced at 11 timepoints, every 100 generations beginning at generation 0. Four lanes of 100 bp paired end sequencing was conducted at the UCSD IGM Genomics center on an Illumina HiSeq 4000 machine. The average per-base-pair coverage across all samples was 131x. Samples E1_t600, E2_t500, Y3_t600, P3_t600, P2_t800, P2_t1000 and A1_t700 yielded data inconsistent with the rest of the allele frequency trajectories from the same population, likely due to mislabelling during sample preparation. These samples were subsequently removed from our analysis.

DNA extraction and library preparation

To minimize competitive growth during handling, 100 µl of a 1:10,000 dilution of frozen stock from each sample was plated on LB agar plates and incubated at 37°C overnight. The entire plate of colonies was then scraped and used for genomic DNA extraction. DNA extractions were conducted using the Geneaid Presto mini gDNA Bacteria Kit (#GBB300) following the manufacturer's protocol. Library preparation was conducted using a modified Illumina Nextera protocol as described in [99].

Validation of variants with Sanger sequencing

43/45 variants, particularly those in loci previously annotated to be involved with translation, were validated using Sanger sequencing. Briefly, populations and timepoints containing the variant at substantial frequency were identified, and clones isolated for genomic DNA extraction, PCR and Sanger Sequencing using standard protocols. The primers used for this validation are detailed in Table S3. The two mutations that failed to validate were expected to be at relatively low frequency in their populations (17% and 38%), so additional clone sampling may be required to validate these events.

Analysis of sequencing data

Variant calling

Sequenced samples were mapped to the MG1655 reference genome (NCBI accession U00096.3) and variants were called using a custom breseq-based pipeline described in Supplementary text section 4 of Ref. [51] and kindly provided to us by Dr. Benjamin Good. Briefly, this method leverages the fact that each population was sampled multiple times across the evolution

experiment to increase our ability to distinguish real low-frequency variants from sequencing errors and other sources of noise.

The reference genome was modified with the appropriate *tufA* sequence for each genetic background used in the evolution experiment along with the removal of the *tufB* sequence, and annotation coordinates were lifted over to be consistent with the original MG1655 reference sequence using custom scripts. The modified reference genomes and annotation files are included in the github repository. The variants reported in Table S1 have been lifted back to be compatible with the original MG1655 reference genome.

Annotation

Variant annotation was conducted using the software package ANNOVAR[100]. Coding variants were established as normal, while noncoding variants were annotated as being associated with the closest gene (in either strand, in either direction) in the genome, as long as it was less than 1 kb away. As ANNOVAR is not set up to work with *E. coli* by default, the *E. coli* MG1655 nucleotide annotation was downloaded in GFF3 form from NCBI Genbank (U00096.3). Cufflinks[101] gffread tool was used to convert this file to GTF, which was then converted to GenePred by using the UCSC Genome Browser gtfToGenePred tool. The final annotation file was generated using the ANNOVAR retrieve_seq_from_fasta.pl script. The annotation file was lifted over to be compatible with each reference sequence.

Copy number variants were called manually using genome-wide coverage plots generated using samtools[102] "view" command and the R computing environment. As these variants have their frequency confounded with their copy number, only their presence/absence was noted for downstream analysis.

Filtering

We considered single nucleotide polymorphisms, short insertion/deletion and the manually identified copy number variants for further analysis. Chromosomal aberrations were ignored because breseq appears to have a high false positive rate (average of 27 "junction" calls per population across all timepoints). Variants were filtered in three successive steps. (1) Variants not identified in multiple consecutive time points were removed. (2) Variants supported by less than 10 reads across all timepoints in a given population were removed. (3) Since we observed fixation events in every population and since there should be no DNA exchange, all truly segregating variants present in a population at generation 100 must either be fixed or lost in generations 900 and 1000. Thus, we removed variants that failed to do so.

Variants that were present at an average frequency $\geq 95\%$ at generation 100 across at least 18 populations were denoted as ancestral mutations that differentiate the founder from the reference genome ($n = 10$). Variants that were not ancestral but present at $\geq 95\%$ on average across all populations derived from one founder were denoted as founder mutations ($n = 11$). These mutations were likely introduced as a byproduct of the strain engineering process. Multiallelic variants (two or more derived alleles present in a single population at the same site) were also removed as likely mapping artifacts. Finally, variants that were present at generation 100 in 11+ populations (of 36 total sequenced populations) are either mapping artifacts or pre-existing variants and were not considered further ($n = 169$, including the 10 ancestral mutations identified earlier).

Identification of adaptive mutations

The putatively adaptive mutations were identified as follows. We first identified mutations that reached at least 10% frequency, were present in at least two consecutive time points and whose frequency changed by at least 20% throughout the evolution. We then merged together such mutations within 10 bp of each other as likely being derived from a single event. This resulted in a set of candidate adaptive mutations. To identify likely adaptive mutations in this candidate set, we considered only mutations in "multihit" genes, i.e., genes with 2 or more candidate adaptive mutations.

Identification of modules in the genome

The 215 genes annotated as being associated with translation were identified using the Gene Ontology database at http://geneontology.org/ by searching for all *E. coli* K12 genes that were identified in a search for "translation OR ribosom". Similarly, the 45 genes associated with cytokinesis were identified using a search for "cytokinesis".

Statistical analyses

The expected number of mutations in multihit genes was calculated via multinomial sampling. Mutations were randomly redistributed across all genes in the *E. coli* genome controlling for variation in gene length. The average of 10,000 such randomizations was used to calculate an empirical FDR. A similar procedure was used to estimate the probability of observed as many or more TM-specific mutations by chance as we actually observed in this study.

To test whether mutations in the 7 TM-specific multi-hit loci were distributed uniformly across the six founders we first estimated the entropy of the distribution of mutations across founders for each gene. Mutations in that gene were then randomly redistributed across six founders 10,000 times, weighted by the total number of TM-specific mutations observed in each founder. An empirical *P*-value was calculated as the fraction trials with smaller than observed entropy value. These *P*-values were then corrected for multiple testing across the 7 TM-specific loci using the Benjamini-Hochberg procedure. We used the same procedure to test for significant deviations in the distributions of generic mutations across founders.

Acknowledgements

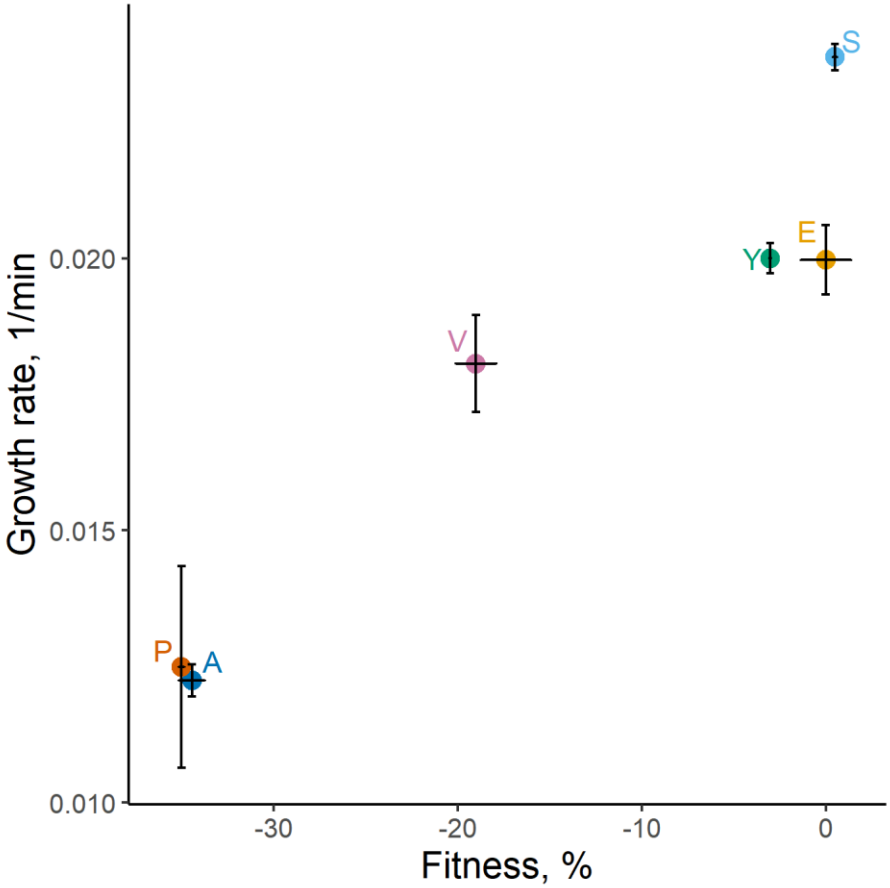
We thank members of the Kryazhimskiy and Kaçar groups, Joanna Masel, Ryan Gutenkunst, Suparna Sanyal, Grant Kinsler, Justin Meyer, and Lin Chao for input and feedback. We thank Alex Pleşa, Divjot Kaur, Emily Peñaherrera, Kevin Longoria and Lesly Villarejo for laboratory assistance. We thank Huanyu Kuo for the analysis of growth-curve data. We thank Eva Garmendia for providing the recombineering plasmids and Georg Rieckh for providing the resistance marker plasmids. We thank Benjamin Good for help with his genome sequencing data analysis pipeline. We thank Kristen Jepsen and the UCSD Institute for Genomic Medicine for sequencing services and the San Diego Supercomputing Center for providing the computational environment. BK acknowledges the support by the John Templeton Foundation (#58562 and #61239); the NASA Exobiology and Evolutionary Biology Program (#H006201406) and the NASA Astrobiology Institute (#NNA17BB05A). SK acknowledges the support by BWF Career

598 Award at the Scientific Interface (#1010719.01), the Alfred P. Sloan Foundation (#FG-2017-
599 9227) and the Hellman Foundation.

600 Supplemental Figures and Tables

601 Figure S1

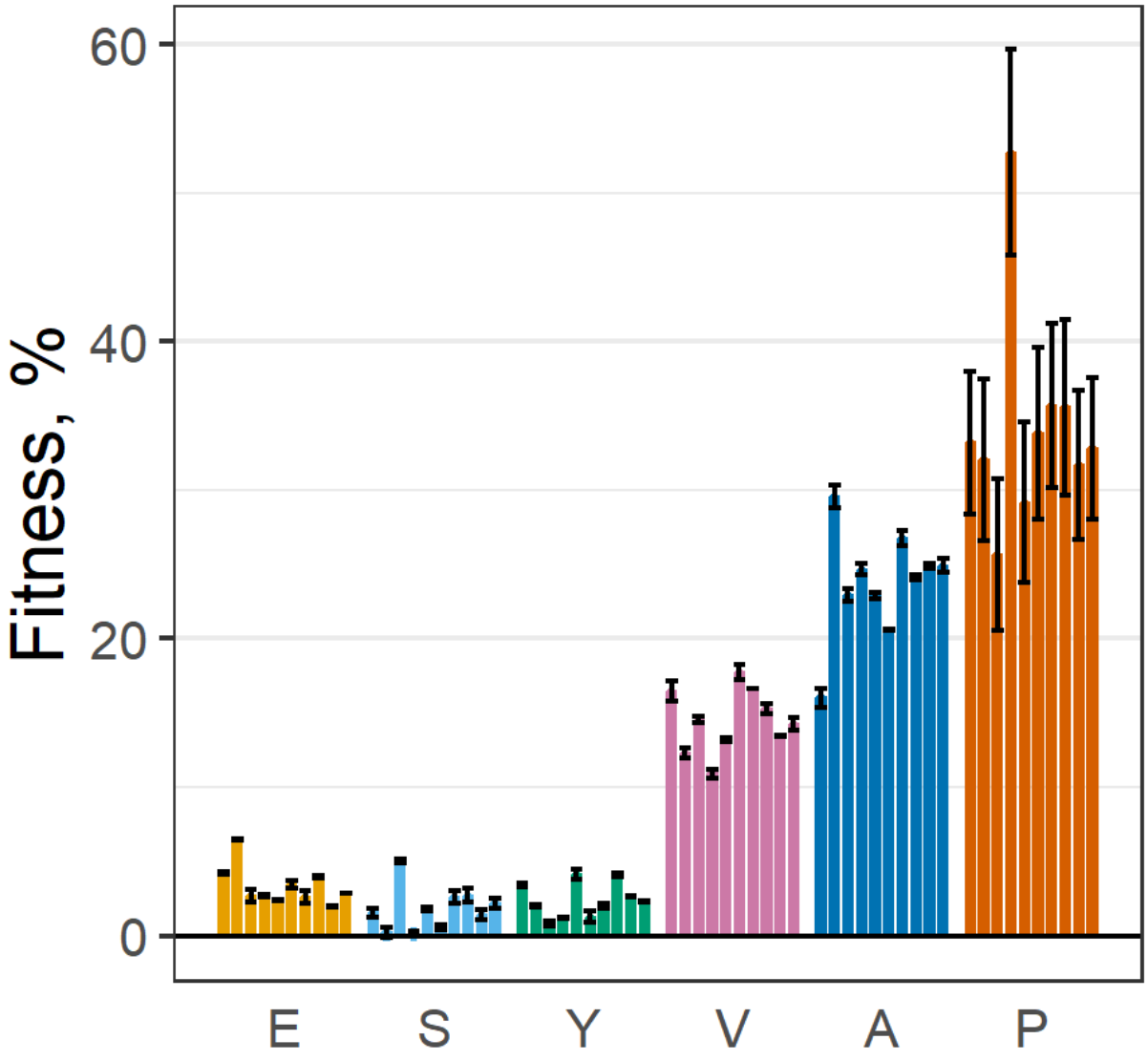
602 Fitness of founders relative to the E strain vs. growth rate. Related to Figure 1.



603
604

605 Figure S2

606 Competitive fitness of the 60 evolved populations relative to their founders, with error bars
607 showing ±1 SEM. Related to figure 1.



608
609

Figure S3

Genome-wide coverage of 11 populations with evidence of *tufA* amplifications shown at one timepoint where the amplification is detected. The red vertical dashed lines denote the genomic coordinates ~50kb upstream and downstream of the *tufA* locus. Related to figure 2.

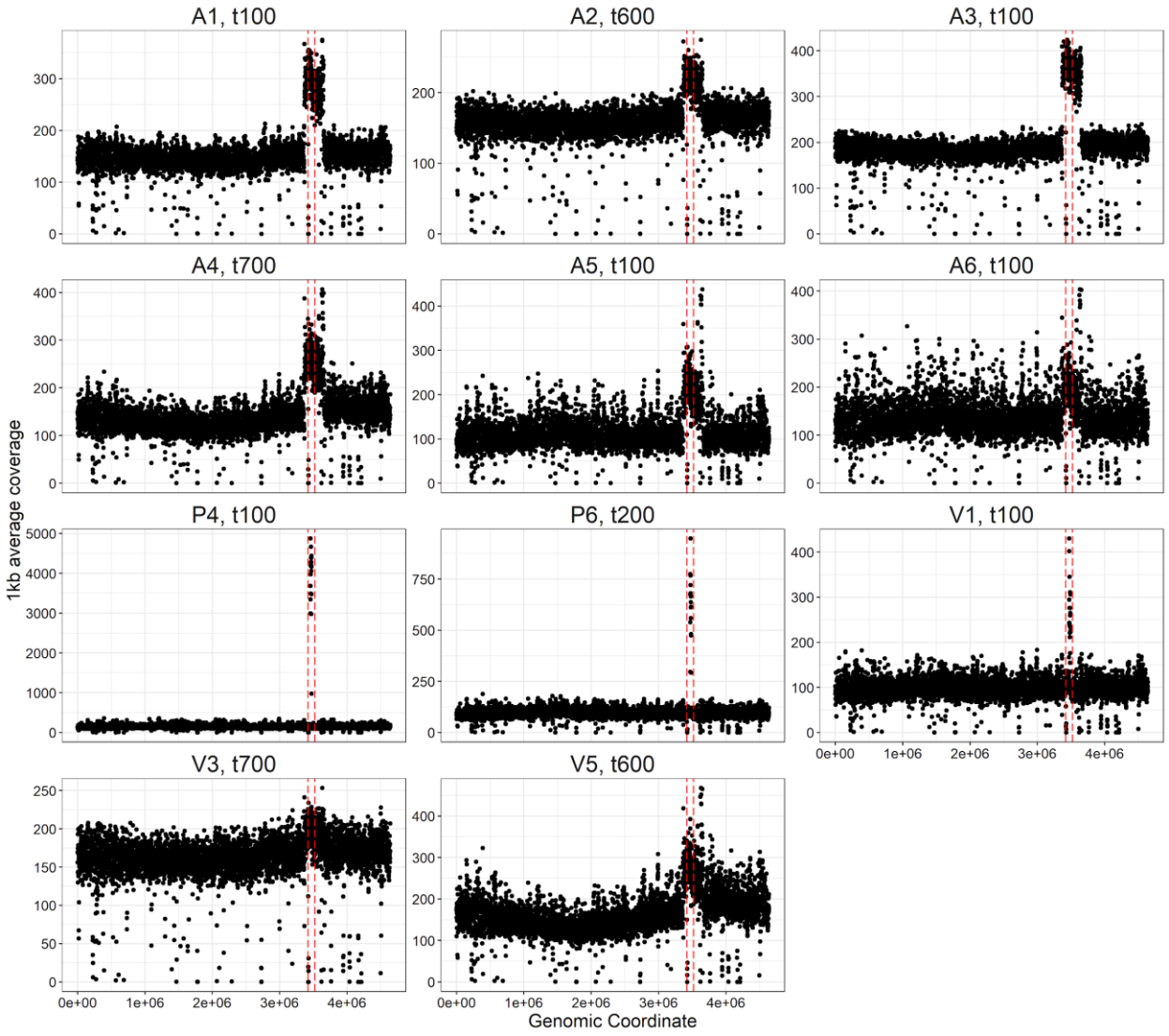


Figure S4

Mutational trajectories for all 36 sequenced populations. Each line represents a single mutation that changed in frequency by > 20%. Thin lines represent mutations in loci that were only mutated once across the entire dataset, while thick lines are loci that were multiply mutated. Generic mutation are shown in grey and TM-specific mutations are shown orange. Orange shading indicates timepoints in which a tufA amplification was detected in that population. Related to figure 2.

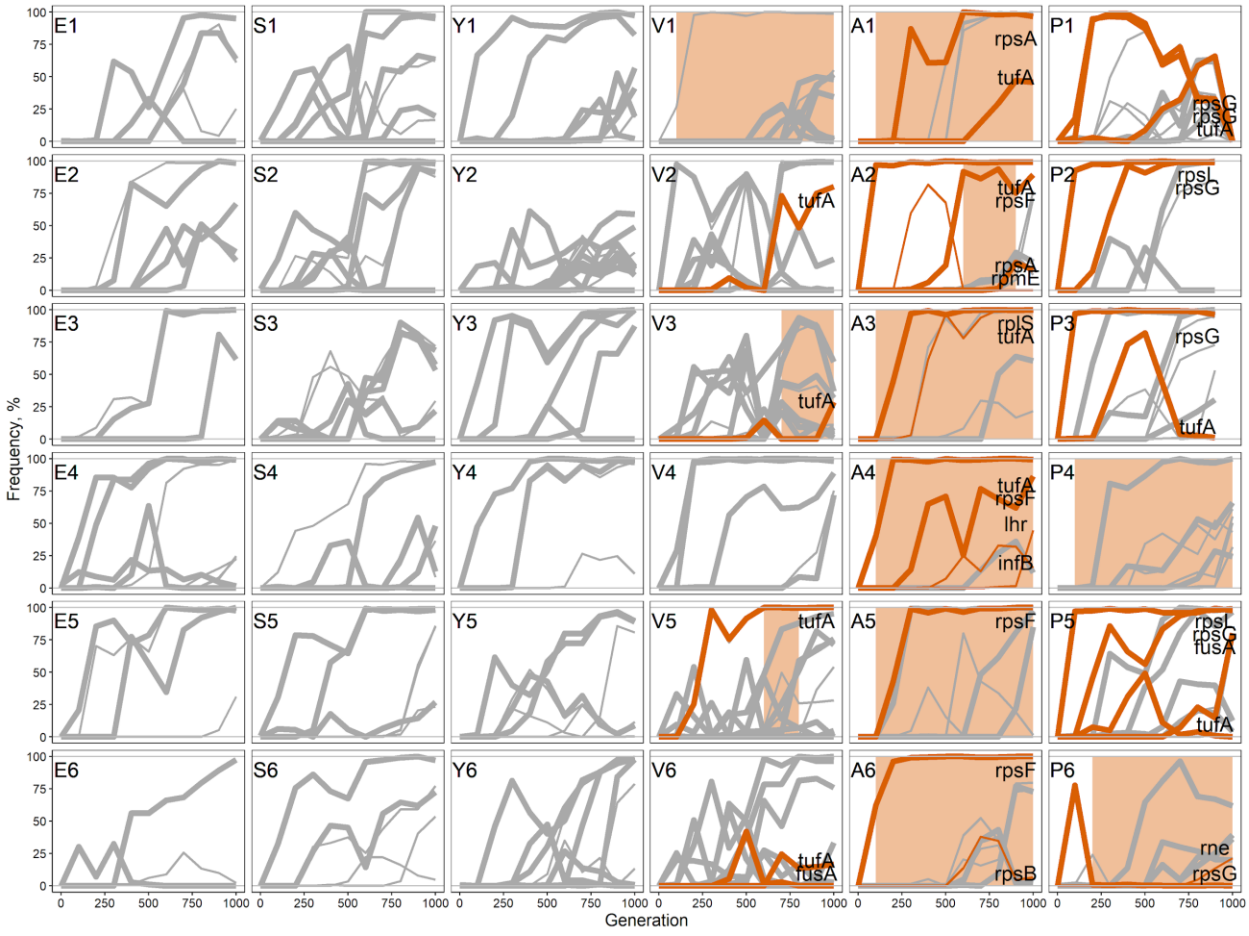


Figure S5

Similar to Figure S4 but for mutations with the relaxed criterion for frequency changes (> 10% instead of > 20%). This relaxation of our filtering criteria increases the total number of detected mutations from 288 to 423 and the number of putatively adaptive mutations in multi-hit genes from 178 to 263. The FDR for putatively adaptive mutations in multi-hit genes is 21.9%. This relaxed filtering yields only one additional TM-specific mutation in gene *rbbA* in population E5. Related to figure 2.

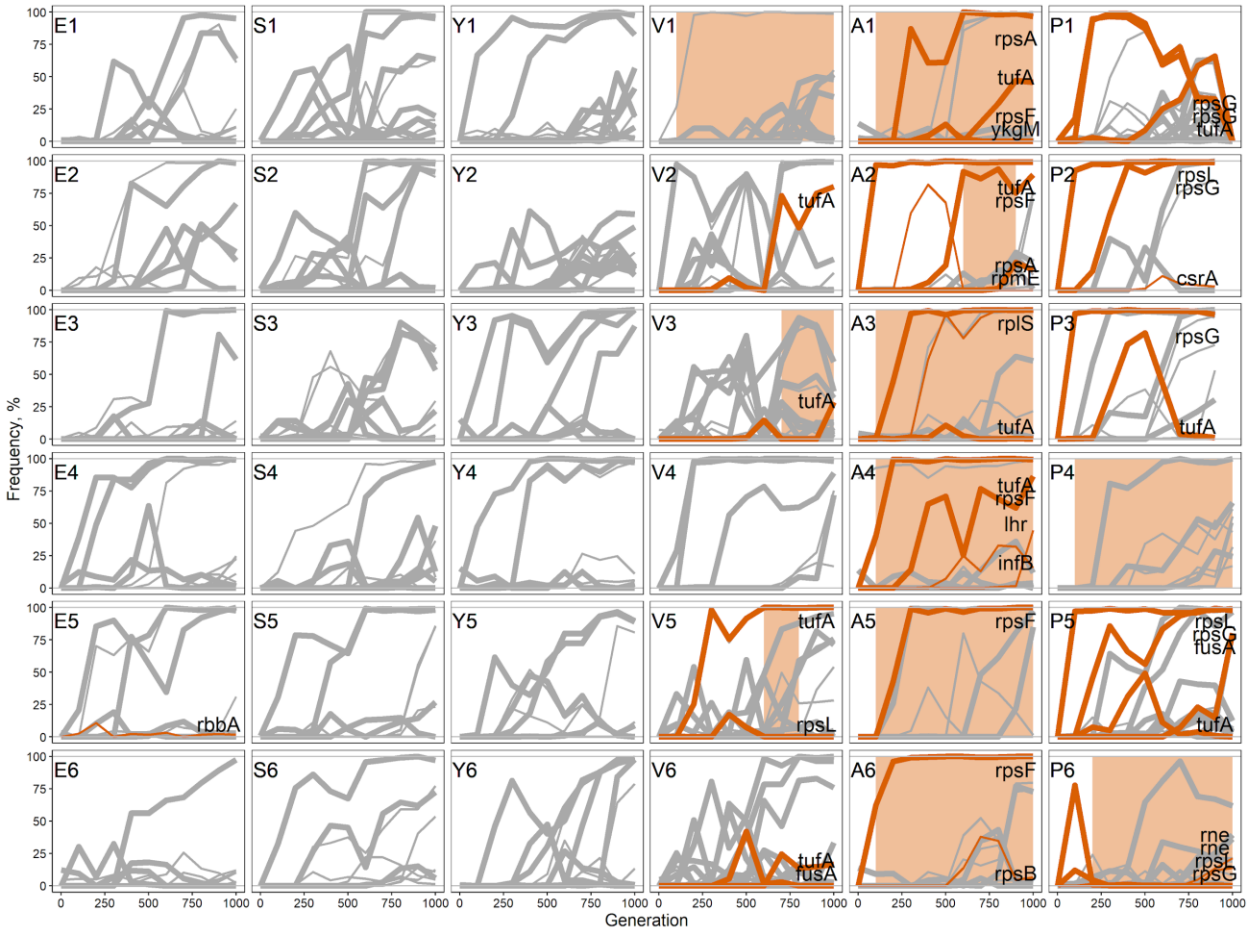


Figure S6

Similar to Figure S4 but highlighting genes annotated as being involved with cytokinesis (blue-green) along with translation genes (orange). Gene names are only shown for cytokinesis mutations. Related to figure 2.

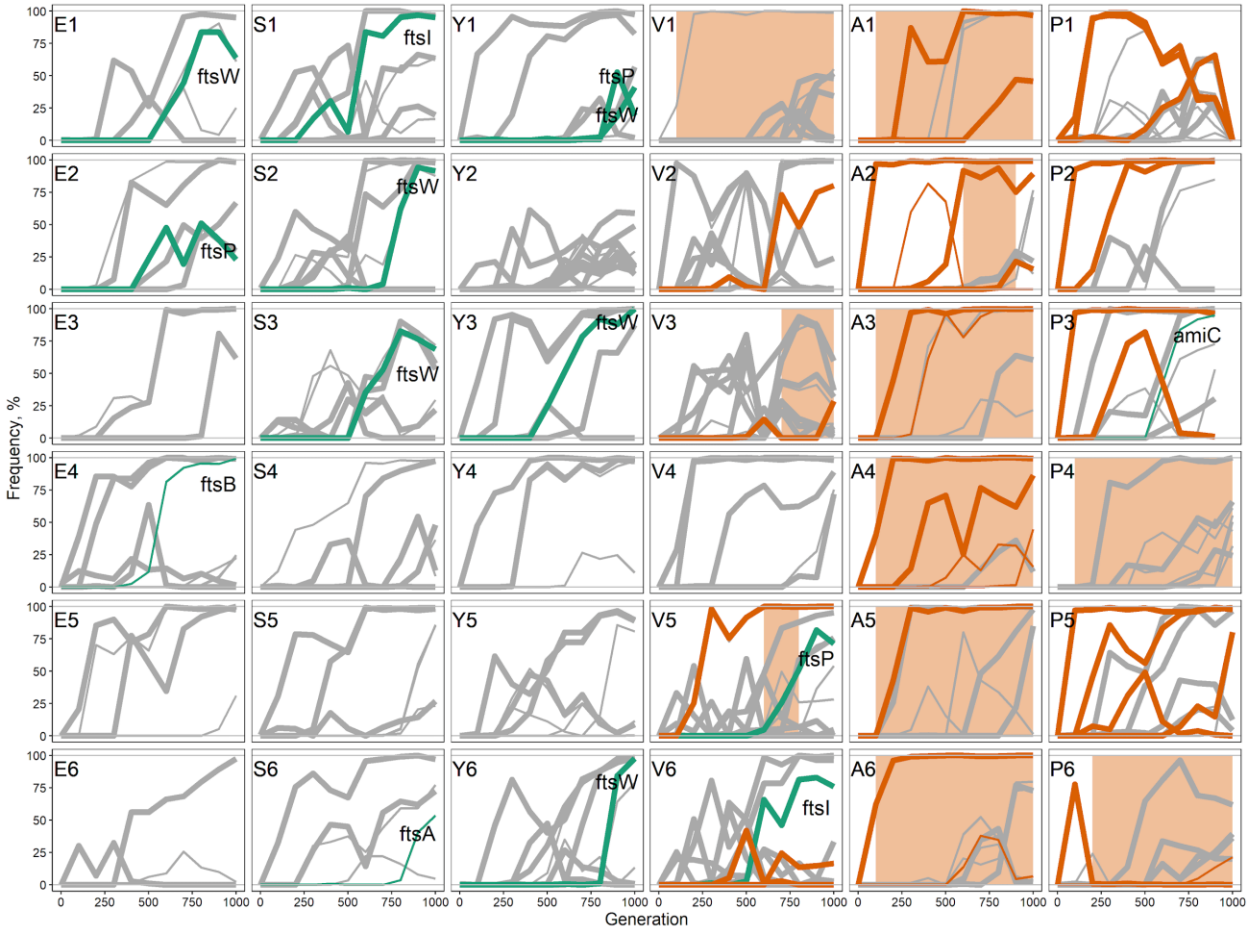


Figure S7

Similar to Figure 3, but considering mutations (*n*=11) in cytokinesis-specific rather than TM-specific genes in all founders. "Generic" in this case is all non-cytokinesis-associated genes, including TM-specific genes. Related to figure 3.

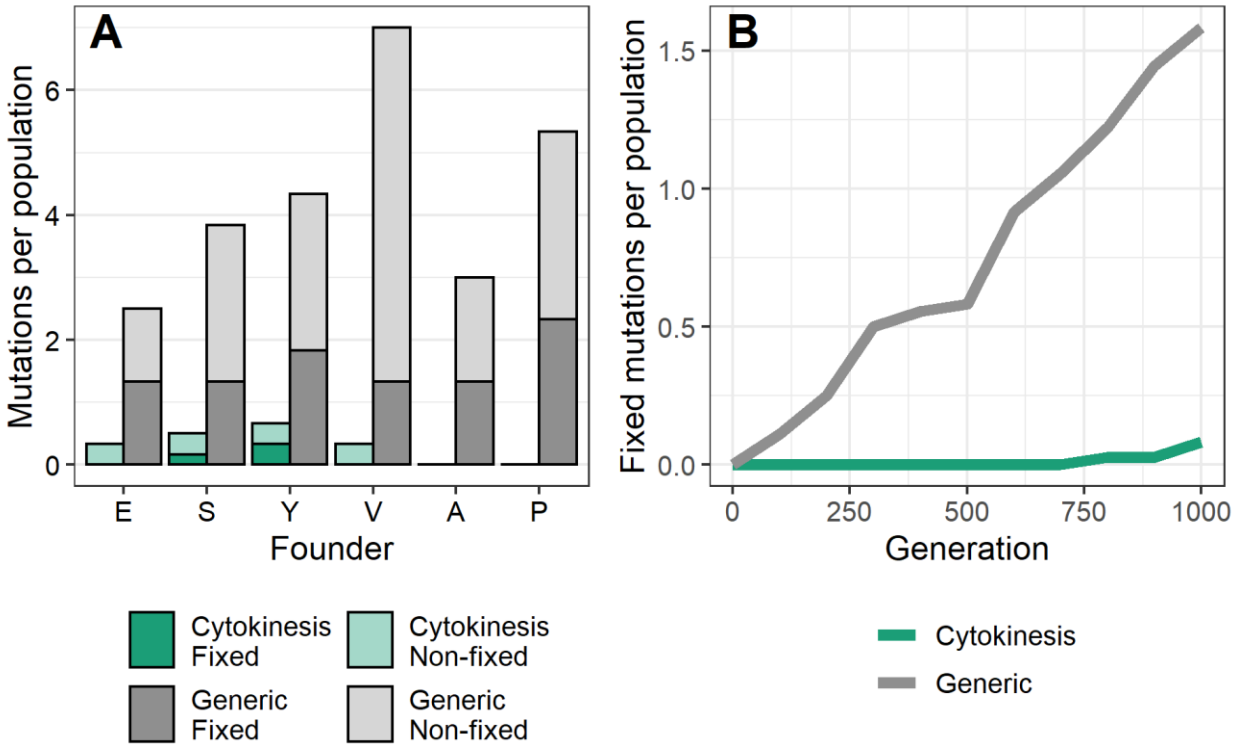
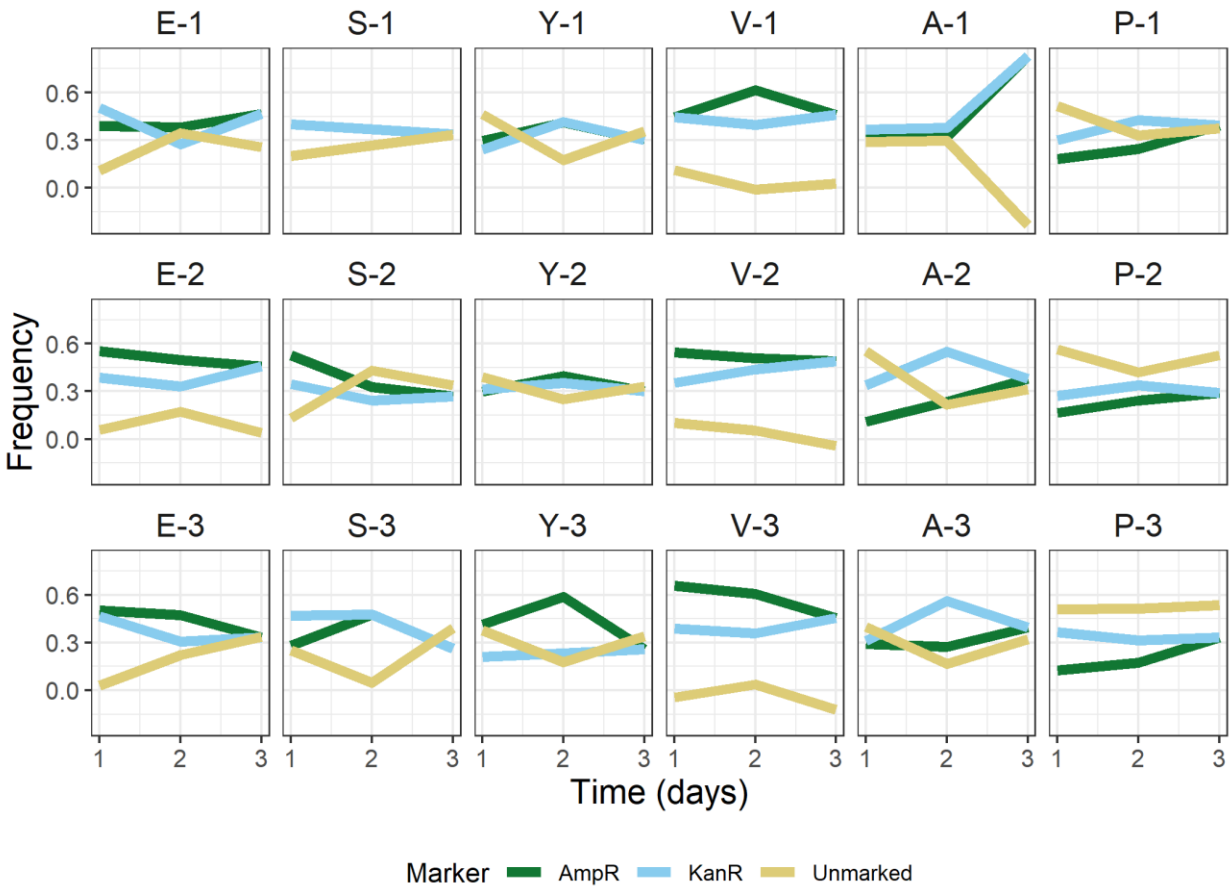


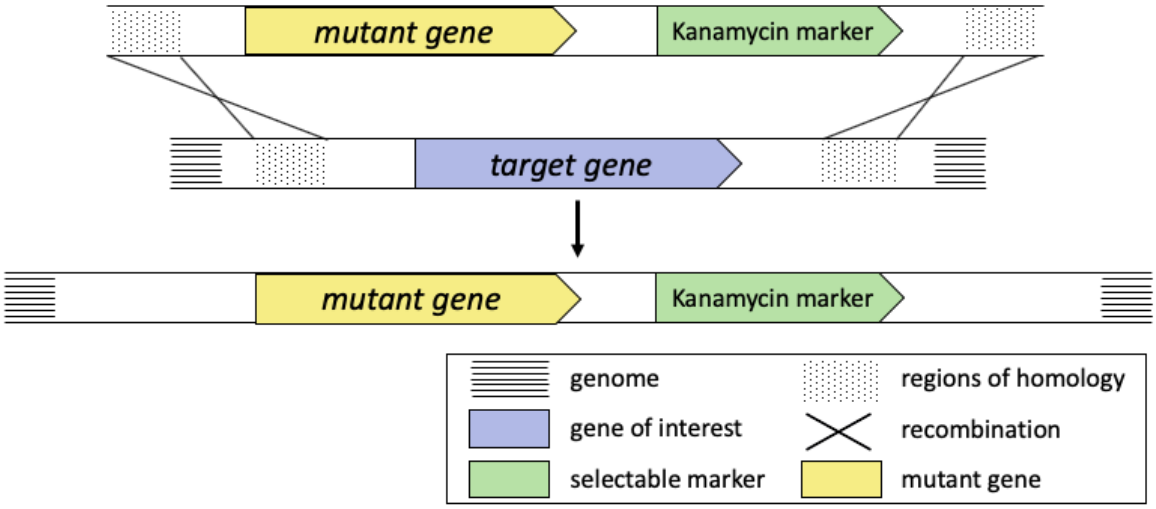
Figure S8

3-way competition between Kanamycin-resistant, Ampicillin-resistant and unmarked strains for each founder to test for fitness differences between markers. Three replicates were conducted for each assay. Related to table 1.



652 # Figure S9

653 Methods for integrating mutations into the genome. Related to figure 5.



654

655

Table S1

List of selected (>20% frequency change) variants. Putatively adaptive (locus repeatedly mutated) variants, which were used for all analyses after Figure 2, are annotated in the table.

Table S2

Primer names and sequences used for strain construction and validation.

Table S3

Primer names and sequences used for Sanger sequencing validation of 43/45 tested variants.

References

1. Simon HA. The Architecture of Complexity. Proc Am Phil Soc. 1962;106: 467–482.

2. Valentine JW, May CL. Hierarchies in Biology and Paleontology. Paleobiology. 1996;22: 23–33.

3. McShea DW. The hierarchical structure of organisms: a scale and documentation of a trend in the maximum. Paleobiology. 2001;27: 405–423.

4. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási AL. Hierarchical organization of modularity in metabolic networks. Science. 2002;297: 1551–1555.

5. Mengistu H, Huizinga J, Mouret J-B, Clune J. The Evolutionary Origins of Hierarchy. PLoS Comput Biol. 2016;12: e1004829.

6. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. Nature. 1999;402: C47–52.

7. Wagner GP, Pavlicev M, Cheverud JM. The road to modularity. Nat Rev Genet. 2007;8: 921–931.

8. Woese C. Molecular mechanics of translation: a reciprocating ratchet mechanism. Nature. 1970;226: 817–820.

9. Frank J. The ribosome--a macromolecular machine par excellence. Chem Biol. 2000;7: R133–41.

10. Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli. Proc Natl Acad Sci U S A. 2008;105: 7899–7906.

11. Gleizer S, Ben-Nissan R, Bar-On YM, Antonovsky N, Noor E, Zohar Y, et al. Conversion of Escherichia coli to Generate All Biomass Carbon from CO2. Cell. 2019;179: 1255–1263.e12.

12. Toprak E, Veres A, Michel J-B, Chait R, Hartl DL, Kishony R. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. Nat Genet. 2012;44: 101–105.

13. Gomez JE, Kaufmann-Malaga BB, Wivagg CN, Kim PB, Silvis MR, Renedo N, et al. Ribosomal mutations promote the evolution of antibiotic resistance in a multidrug environment. Elife. 2017;6. doi:10.7554/eLife.20420

14. Lieberman TD, Michel J-B, Aingaran M, Potter-Bynoe G, Roux D, Davis MR Jr, et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. Nat Genet. 2011;43: 1275–1280.

15. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, et al. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature. 2013;500: 571–574.

16. Barroso-Batista J, Sousa A, Lourenço M, Bergman M-L, Sobral D, Demengeot J, et al. The first steps of adaptation of Escherichia coli to the gut are dominated by soft sweeps. PLoS Genet. 2014;10: e1004182.

17. Lourenço M, Ramiro RS, Güleresi D, Barroso-Batista J, Xavier KB, Gordo I, et al. A Mutational Hotspot and Strong Selection Contribute to the Order of Mutations Selected for during Escherichia coli Adaptation to the Gut. PLoS Genet. 2016;12: e1006420.

18. Blundell JR, Schwartz K, Francois D, Fisher DS, Sherlock G, Levy SF. The dynamics of adaptive genetic diversity during the early stages of clonal evolution. Nature Ecology & Evolution. 2019;3: 293–301.

19. Nguyen Ba AN, Cvijović I, Rojas Echenique JI, Lawrence KR, Rego-Costa A, Liu X, et al. High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. Nature. 2019;575: 494–499.

20. Garud NR, Good BH, Hallatschek O, Pollard KS. Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. PLoS Biol. 2019;17: e3000102.

21. Desai MM, Fisher DS. Beneficial mutation–selection balance and the effect of linkage on positive selection. Genetics. 2007;176: 1759–1798.

22. Neher RA, Shraiman BI. Competition between recombination and epistasis can cause a transition from allele to genotype selection. Proc Natl Acad Sci U S A. 2009;106: 6866–6871.

23. Good BH, Rouzine IM, Balick DJ, Hallatschek O, Desai MM. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proc Natl Acad Sci U S A. 2012;109: 4950–4955.

24. Levy SF, Blundell JR, Venkataram S, Petrov DA, Fisher DS, Sherlock G. Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature. 2015;519: 181.

25. Lynch M, Field MC, Goodson HV, Malik HS, Pereira-Leal JB, Roos DS, et al. Evolutionary cell biology: two origins, one objective. Proc Natl Acad Sci U S A. 2014;111: 16990–16994.

26. Kryazhimskiy S, Tkacik G, Plotkin JB. The dynamics of adaptation on correlated fitness landscapes. Proc Natl Acad Sci U S A. 2009;106: 18638–18643.

27. Hartl DL, Dykhuizen DE, Dean AM. Limits of adaptation: the evolution of selective neutrality. Genetics. 1985;111: 655–674.

28. Hartl DL, Taubes CH. Towards a theory of evolutionary adaptation. Genetica. 1998;102-103: 525–533.

29. Lynch M. Evolutionary layering and the limits to cellular perfection. Proc Natl Acad Sci U S A. 2012.

30. Nourmohammad A, Schiffels S, Lässig M. Evolution of molecular phenotypes under stabilizing selection. J Stat Mech: Theory Exp. 2013;2013: P01012.

31. Lynch M. Evolutionary layering and the limits to cellular perfection. Proceedings of the National Academy of. 2012.

32. Goyal S, Balick DJ, Jerison ER, Neher RA, Shraiman BI, Desai MM. Dynamic Mutation–Selection Balance as an Evolutionary Attractor. Genetics. 2012;191: 1309–1319.

33. Rice DP, Good BH, Desai MM. The evolutionarily stable distribution of fitness effects. Genetics. 2015;200: 321–329.

34. Barton NH. Linkage and the limits to natural selection. Genetics. 1995;140: 821–841.

735 35. Gerrish PJ, Lenski RE. The fate of competing beneficial mutations in an asexual population.
736 Genetica. 1998;102-103: 127–144.

737 36. Schiffels S, Szöllosi GJ, Mustonen V, Lässig M. Emergent neutrality in adaptive asexual evolution.
738 Genetics. 2011;189: 1361–1375.

739 37. Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, et al. The repertoire and dynamics
740 of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet.
741 2008;4: e1000303.

742 38. Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, Farkas Z, et al. The genomic landscape of
743 compensatory evolution. PLoS Biol. 2014;12: e1001935.

744 39. Hong J, Gresham D. Molecular specificity, convergence and constraint shape adaptive evolution in
745 nutrient-poor environments. PLoS Genet. 2014;10: e1004041.

746 40. Payen C, Di Rienzi SC, Ong GT, Pogachar JL, Sanchez JC, Sunshine AB, et al. The dynamics of
747 diverse segmental amplifications in populations of Saccharomyces cerevisiae adapting to strong
748 selection. G3 . 2014;4: 399–409.

749 41. Couñago R, Chen S, Shamoo Y. In vivo molecular evolution reveals biophysical origins of
750 organismal fitness. Mol Cell. 2006;22: 441–449.

751 42. Chou H-H, Delaney NF, Draghi JA, Marx CJ. Mapping the fitness landscape of gene expression
752 uncovers the cause of antagonism and sign epistasis between adaptive mutations. PLoS Genet.
753 2014;10: e1004149.

754 43. Michener JK, Neves AAC, Vuilleumier S, Bringel F, Marx CJ. Effective use of a horizontally-
755 transferred pathway for dichloromethane catabolism requires post--transfer refinement. Elife.
756 2014;3: e04279.

757 44. Venkataram S, Dunn B, Li Y, Agarwala A, Chang J, Ebel ER, et al. Development of a
758 Comprehensive Genotype-to-Fitness Map of Adaptation-Driving Mutations in Yeast. Cell.
759 2016;166: 1585–1596.e22.

760 45. Jerison ER, Kryazhimskiy S, Mitchell JK, Bloom JS, Kruglyak L, Desai MM. Genetic variation in
761 adaptability and pleiotropy in budding yeast. Elife. 2017;6.

762 46. Rojas Echenique JI, Kryazhimskiy S, Nguyen Ba AN, Desai MM. Modular epistasis and the
763 compensatory evolution of gene deletion mutants. PLoS Genet. 2019;15: e1007958.

764 47. Rodrigues JV, Shakhnovich EI. Adaptation to mutational inactivation of an essential gene converges
765 to an accessible suboptimal fitness peak. Elife. 2019;8. doi:10.7554/eLife.50509

766 48. Laan L, Koschwanez JH, Murray AW. Evolutionary adaptation after crippling cell polarization
767 follows reproducible trajectories. Elife. 2015;4. doi:10.7554/eLife.09638

768 49. Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, et al. The
769 molecular diversity of adaptive convergence. Science. 2012;335: 457–461.

770 50. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. Global epistasis makes adaptation predictable
771 despite sequence-level stochasticity. Science. 2014;344: 1519–1522.

772 51. Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. The dynamics of molecular evolution
773 over 60,000 generations. Nature. 2017;551: 45–50.

774 52. Castoe TA, de Koning APJ, Kim H-M, Gu W, Noonan BP, Naylor G, et al. Evidence for an ancient
775 adaptive episode of convergent molecular evolution. Proc Natl Acad Sci U S A. 2009;106: 8986–
776 8991.

777 53. Chen L, DeVries AL, Cheng CH. Convergent evolution of antifreeze glycoproteins in Antarctic
778 notothenioid fish and Arctic cod. Proc Natl Acad Sci U S A. 1997;94: 3817–3822.

779 54. Rosenblum EB, Römpler H, Schöneberg T, Hoekstra HE. Molecular and functional basis of
780 phenotypic convergence in white lizards at White Sands. Proc Natl Acad Sci U S A. 2010;107:
781 2113–2117.

782 55. Ramakrishnan V. Ribosome structure and the mechanism of translation. Cell. 2002;108: 557–572.

783 56. Williamson JR. The ribosome at atomic resolution. Cell. 2009;139: 1041–1043.

784 57. Melnikov S, Ben-Shem A, Garreau de Loubresse N, Jenner L, Yusupova G, Yusupov M. One core,
785 two shells: bacterial and eukaryotic ribosomes. Nat Struct Mol Biol. 2012;19: 560–567.

786 58. Kaçar B, Garmendia E, Tuncbag N, Andersson DI, Hughes D. Functional Constraints on Replacing
787 an Essential Gene with Its Ancient and Modern Homologs. MBio. 2017;8. doi:10.1128/mBio.01276-
788 17

789 59. Miller DL, Weissbach H. Factors involved in the transfer of aminoacyl-tRNA to the ribosome. In:
790 Petska S, Weissbach H, editors. Molecular mechanisms of protein biosynthesis. Academic Press,
791 London; 1977. pp. 323–373.

792 60. Nilsson J, Nissen P. Elongation factors on the ribosome. Curr Opin Struct Biol. 2005;15: 349–354.

793 61. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. Interdependence of Cell Growth and
794 Gene Expression: Origins and Consequences. Science. 2010;330: 1099–1102.

795 62. Tubulekas I, Hughes D. Suppression of rpsL phenotypes by tuf mutations reveals a unique
796 relationship between translation elongation and growth rate. Mol Microbiol. 1993;7: 275–284.

797 63. MacLean RC, Perron GG, Gardner A. Diminishing returns from beneficial mutations and pervasive
798 epistasis shape the fitness landscape for rifampicin resistance in Pseudomonas aeruginosa. Genetics.
799 2010;186: 1345–1354.

800 64. Schnell R, Abdulkarim F, Kálmán M, Isaksson LA. Functional EF-Tu with large C-terminal
801 extensions in an E. coli strain with a precise deletion of both chromosomal tuf genes. FEBS Lett.
802 2003;538: 139–144.

803 65. Lenski RE, Mongold JA, Sniegowski PD, Travisano M, Vasi F, Gerrish PJ, et al. Evolution of
804 competitive fitness in experimental populations of E. coli: what makes one genotype a better
805 competitor than another? Antonie Van Leeuwenhoek. 1998;73: 35–47.

806 66. Li Y, Venkataram S, Agarwala A, Dunn B, Petrov DA, Sherlock G, et al. Hidden Complexity of
807 Yeast Adaptation under Simple Evolutionary Conditions. Curr Biol. 2018;28: 515–525.e6.

808 67. Manhart M, Adkar BV, Shakhnovich EI. Trade-offs between microbial growth phases lead to

809 frequency-dependent and non-transitive selection. Proc Biol Sci. 2018;285.

810 68. Vasi F, Travisano M, Lenski RE. Long-Term Experimental Evolution in Escherichia coli. II.
811 Changes in Life-History Traits During Adaptation to a Seasonal Environment. Am Nat. 1994;144:
812 432–456.

813 69. Utnes ALG, Sørum V, Hülter N, Primicerio R, Hegstad J, Kloos J, et al. Growth phase-specific
814 evolutionary benefits of natural transformation in Acinetobacter baylyi. ISME J. 2015;9: 2221–2231.

815 70. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative epistasis between beneficial
816 mutations in an evolving bacterial population. Science. 2011;332: 1193–1196.

817 71. Wünsche A, Dinh DM, Satterwhite RS, Arenas CD, Stoebel DM, Cooper TF. Diminishing-returns
818 epistasis decreases adaptability along an evolutionary trajectory. Nat Ecol Evol. 2017;1: 61.

819 72. Wiser MJ, Ribeck N, Lenski RE. Long-term dynamics of adaptation in asexual populations. Science.
820 2013;342: 1364–1367.

821 73. Schoustra SE, Bataillon T, Gifford DR, Kassen R. The properties of adaptive walks in evolving
822 populations of fungus. PLoS Biol. 2009;7: e1000250.

823 74. Couce A, Tenaillon OA. The rule of declining adaptability in microbial evolution experiments. Front
824 Genet. 2015;6: 99.

825 75. Herron MD, Doebeli M. Parallel evolutionary dynamics of adaptive diversification in Escherichia
826 coli. PLoS Biol. 2013;11: e1001490.

827 76. Kvitek DJ, Sherlock G. Whole genome, whole population sequencing reveals that loss of signaling
828 networks is the major adaptive strategy in a constant environment. PLoS Genet. 2013;9: e1003972.

829 77. Asai T, Zaporojets D, Squires C, Squires CL. An Escherichia coli strain with all chromosomal rRNA
830 operons inactivated: complete exchange of rRNA genes between bacteria. Proc. Natl. Acad. Sci. U.
831 S. A. 1999. pp. 1971–1976.

832 78. Dutheil JY, Jossinet F, Westhof E. Base pairing constraints drive structural epistasis in ribosomal
833 RNA sequences. Mol Biol Evol. 2010;27: 1868–1876.

834 79. Barreto FS, Burton RS. Evidence for compensatory evolution of ribosomal proteins in response to
835 rapid divergence of mitochondrial rRNA. Mol Biol Evol. 2013;30: 310–314.

836 80. Sloan DB, Triant DA, Wu M, Taylor DR. Cytonuclear interactions and relaxed selection accelerate
837 sequence evolution in organelle ribosomes. Mol Biol Evol. 2014;31: 673–682.

838 81. Shah P, McCandlish DM, Plotkin JB. Contingency and entrenchment in protein evolution under
839 purifying selection. Proc Natl Acad Sci U S A. 2015;112: E3226–35.

840 82. Starr TN, Flynn JM, Mishra P, Bolon DNA, Thornton JW. Pervasive contingency and entrenchment
841 in a billion years of Hsp90 evolution. Proc Natl Acad Sci U S A. 2018;115: 4453–4458.

842 83. Blount ZD, Lenski RE, Losos JB. Contingency and determinism in evolution: Replaying life's tape.
843 Science. 2018;362. Available: https://www.ncbi.nlm.nih.gov/pubmed/30409860

844 84. Orr HA. The distribution of fitness effects among beneficial mutations. Genetics. 2003;163: 1519–

1526.

85. Orr HA. The distribution of fitness effects among beneficial mutations in Fisher's geometric model of adaptation. J Theor Biol. 2006;238: 279–285.

86. Martin G, Lenormand T. The distribution of beneficial and fixed mutation fitness effects close to an optimum. Genetics. 2008;179: 907–916.

87. Fisher R. The genetical theory of natural selection. Oxford University Press; 1930.

88. Tenaillon O. The Utility of Fisher's Geometric Model in Evolutionary Genetics. Annu Rev Ecol Evol Syst. 2014;45: 179–201.

89. Wagner A. Neutralism and selectionism: a network-based reconciliation. Nat Rev Genet. 2008;9: 965–974.

90. Schaper S, Johnston IG, Louis AA. Epistasis can lead to fragmented neutral spaces and contingency in evolution. Proc Biol Sci. 2012;279: 1777–1783.

91. Newman MEJ, Engelhardt R. Effects of selective neutrality on the evolution of molecular species. Proc Roy Soc B: Biological Sciences. 1998;265: 1333–1338.

92. Lind PA, Tobin C, Berg OG, Kurland CG, Andersson DI. Compensatory gene amplification restores fitness after inter-species gene replacements. Mol Microbiol. 2010;75: 1078–1089.

93. Kacar B, Ge X, Sanyal S, Gaucher EA. Experimental Evolution of Escherichia coli Harboring an Ancient Translation Protein. J Mol Evol. 2017;84: 69–84.

94. Chou H-H, Chiu H-C, Delaney NF, Segrè D, Marx CJ. Diminishing returns epistasis among beneficial mutations decelerates adaptation. Science. 2011;332: 1190–1192.

95. Wei X, Zhang J. Patterns and mechanisms of diminishing returns from beneficial mutations. Mol Biol Evol. 2019.

96. Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, et al. Just-in-time transcription program in metabolic pathways. Nat Genet. 2004;36: 486–491.

97. Chung CT, Niemela SL, Miller RH. One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci U S A. 1989;86: 2172–2175.

98. Geissmann Q. OpenCFU, a new free and open-source software to count cell colonies and other circular objects. PLoS One. 2013;8: e54072.

99. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. Inexpensive multiplexed library preparation for megabase-sized genomes. PLoS One. 2015;10: e0128036.

100. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38: e164.

101. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7: 562–578.

881 102. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence
882 Alignment/Map format and SAMtools. Bioinformatics. 2009;25: 2078–2079.

883