Evolutionary stalling and a limit on the power of natural selection to improve a cellular module

Sandep Venkatarama*, Ross Monaskyb, Shohreh H. Sikaroodd1, Sergey Kryazhimskiyb,2,3, and Betul Kacaerb,c,2,3

*Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093; bDepartment of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721; and cLunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721

Edited by Harmit S. Malik, Fred Hutchinson Cancer Research Center, Seattle, WA, and approved June 20, 2020 (received for review December 17, 2019)

Cells consist of molecular modules which perform vital biological functions. Cellular modules are key units of adaptive evolution because organismal fitness depends on their performance. Theory shows that in rapidly evolving populations, such as those of many microbes, adaptation is driven primarily by common beneficial mutations with large effects, while other mutations behave as if they are effectively neutral. As a consequence, if a module can be improved only by rare and/or weak beneficial mutations, its adaptive evolution would stall. However, such evolutionary stalling has not been empirically demonstrated, and it is unclear to what extent stalling may limit the power of natural selection to improve modules. Here we empirically characterize how natural selection improves the translation machinery (TM), an essential cellular module. We experimentally evolved populations of Escherichia coli with genetically perturbed TMs for 1,000 generations. Populations with severe TM defects initially adapted via mutations in the TM, but TM adaptation stalled within about 300 generations. We estimate that the genetic load in our populations incurred by residual TM defects ranges from 0.5 to 19%. Finally, we found evidence that both epistasis and the depletion of the pool of beneficial mutations contributed to evolutionary stalling. Our results suggest that cellular modules may not be fully optimized by natural selection despite the availability of adaptive mutations.

B)

Author contributions: S.V., S.K., and B.K. designed research; S.V., R.M., S.H.S., S.K., and B.K. performed research; S.V., S.K., and B.K. analyzed data; and S.V., S.K., and B.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

Data deposition: The script, modified reference genomes, and the raw data (except for raw sequencing data) used for analysis can be found at GitHub (https://github.com/BetulKacar/ModEvoStalling). Raw sequencing data for this project have been deposited into the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under project PRJNA560965.

*Present address: Molecular Engineering Group, Fate Therapeutics Inc., San Diego, CA 92121.

1S.K. and B.K. contributed equally to this work.

2To whom correspondence may be addressed. Email: betul@arizona.edu or skryazhi@ucsd.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1921881117/-/DCSupplemental.
modules. In particular, improvable modules accumulate adaptive mutations only if such mutations arise frequently and provide fitness benefits above the emergent neutrality threshold. In contrast, modules that are improvable only by rare mutations, or mutations whose fitness benefits are too small, will not adapt (29). We refer to a failure of an otherwise improvable module to accumulate adaptive mutations as “evolutionary stalling.”

Evolutionary stalling can theoretically limit the power of natural selection to improve a module, but whether stalling occurs in biological systems is unclear. There are plausible scenarios in which evolutionary stalling would not occur. For example, if many modules are improvable by mutations with similar rates and fitness effects, all of them would improve simultaneously, without exhibiting evolutionary stalling. Some evolution experiments appear to support this possibility (18, 39–43). However, in some conditions, natural selection focuses on improving a single module or a handful of modules (40, 42, 44–53). For example, resistance mutations in bacteria often occur in the protein complex whose function is inhibited by an antibiotic (12, 13, 54). Similarly, compensatory mutations after a severe genetic perturbation are often concentrated in a handful of pathways with functional relationships to the perturbation (40, 42, 44, 48, 49, 52, 53). The fact that adaptive mutations are not observed in some of the modules can be interpreted in two ways. Either these modules exist only in an evolutionary stalling trajectory not improvable in the first place. Previous studies have not attempted to distinguish between these two possibilities.

Whenever there are modules whose adaptation is stalled, we would generically expect the focus of natural selection to shift over time from some modules to others, even in a constant selective environment. Such shifts should occur because the emergent neutrality threshold that determines which modules adapt and which ones stall is a dynamic quantity. As mentioned above, this threshold depends on the supply and the effects of all adaptive mutations in the genome, which themselves change over time by at least two genetic mechanisms. First, as a population adapts, the supply of available adaptive mutations is being depleted with large-effect mutations being depleted first (41). Second, genetic interactions (epistasis) can modulate the fitness benefits of mutations (39, 55–57) or even open up and close down large pools of adaptive mutations (10, 40, 58, 59). A recent study of coevolution of bacterial translation machinery in Escherichia coli found evidence that both of these genetic mechanisms cause temporal shifts in the statistical distribution of mutations among genes and operons (41). However, this study did not identify the specific modules in which adaptive evolution stalled or resumed, nor did it quantify whether evolutionary stalling imposes any limit on the power of selection to improve any specific module.

Here we use experimental populations of the bacterium E. coli to explicitly demonstrate the onset of stalling in the evolution of the translation machinery and to quantify the fitness cost imposed by it. In E. coli, the core macromolecular complexes that carry out translation are encoded by a well-annotated set of about 200 genes (8, 9, 60). We operationally define this set as the translation machinery (TM) module. To detect evolutionary stalling, we first disrupt the TM by replacing the native elongation factor Tu (EF-Tu) in E. coli with several of its orthologs (61, 62). We then evolve these strains in rich media where rapid and accurate translation is required for fast growth (63). We thus expect natural selection to favor adaptive mutations in the TM. In addition to such TM-specific mutations, mutations that improve the performance of other cellular modules could also be adaptive. We refer to such mutations as “generic.” We show that the TM-specific and generic mutations compete against each other in our populations. We then demonstrate that TM adaptation stalls in at least some populations and estimate the genetic load caused by stalling. Finally, we characterize the genetic mechanisms that contribute to evolutionary stalling in our 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 (61) (Table 1 and Dataset S1). 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 cell, the tufB gene was deleted, and the foreign orthologs were integrated into the tufA locus (61). 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). We first quantified the TM defects in our founder strains (Materials and Methods). Kacar 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 Luria–Bertani (LB) medium (61). In our subsequent evolution experiments, 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 (SI Appendix, Fig. S1). We conclude that the competitive fitness of our founders in our environment reflects their TM performance. Further, we found that the fitnesses 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. Note that the E strain itself carries a 4% fitness cost compared to the wild-type E. coli strain (Materials and Methods and Dataset S2).

Clonal Interference Slows Down TM Adaptation. We instantiated 10 replicate populations from each of our six founders (60 populations total) and evolved them in LB for 1,000 generations with daily 1:104 dilutions and the bottleneck population size N = 5 × 107 cells (Materials and Methods). The competitive fitness of all but one of the populations against their respective founders significantly increased during evolution (P < 0.05, t test after Benjamini–Hochberg correction; Fig. 4A and Dataset S3), with significantly larger increases observed in the V, A, and P populations compared to E, S, and Y populations (ANOVA, P < 10−16). Furthermore, 24 out of 30 E, S, and Y populations and 3 of the 30 V, A, and P populations became more fit than the control E strain with 95% confidence (Fig. 4B and Dataset S4).

Adaptation in these populations can be driven not only by mutations in the TM but also by mutations in other modules. However, evolutionary stalling in the TM can occur only if mutations improving the TM compete against other types of mutations within the same population. To determine whether both types of mutations occur in our populations, we conducted whole-genome sequencing at multiple time points throughout the evolution experiment. This sequencing strategy allows us to directly observe competition dynamics between mutations in different modules (18, 41, 70, 71).

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 SNPs and small insertion/deletion events in this dataset (Materials and Methods). Then, we called a mutation adaptive if it satisfied two criteria: 1) its frequency changed by more than 20% in a population and 2) it occurred in a multihit gene, i.e., a gene in which two independent mutations passed the
first criterion. We augmented our pipeline with the manual identification of large-scale copy-number variants which could only be reliably detected after they reached high frequency in a population (Materials and Methods and SI Appendix, Fig. S2). We successfully validated 43/45 tested variants via Sanger sequencing (Materials and Methods and Dataset S5).

We first identified and broadly classified the targets of adaptation and documented their competition dynamics across all populations, irrespectively of their founder. Our filtering procedure yielded a total of 167 new putatively adaptive mutations in 28 multihit 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 (Materials and Methods, SI Appendix, Figs. S2 and S3, and Dataset S6). We classified putatively adaptive mutations as TM-specific if the genes where they occurred are annotated as translation-related (Materials and Methods). We classified all other mutations as generic. We found that 38 out of 178 (21%) putatively adaptive mutations in 6 out of 28 multihit genes were TM-specific (Dataset S6). This is significantly more than expected by chance ($P < 10^{-4}$, randomization test) since the 215 genes annotated as translation-related compose only 4.3% 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 (SI Appendix, Fig. S4).

TM-specific mutations occurred in 17 out of 36 sequenced populations. Generic mutations were also observed in all of these populations (SI Appendix, Fig. S3). Thus, whenever TM-specific mutations occurred, generic mutations also occurred, such that the fate of TM-specific mutations likely depended on the outcome of competition between mutations within the same modules (Fig. 2). As a result of this competition, only 14 out of 27 (52%) TM-specific mutations that arose (excluding 11 tufA amplifications) went to fixation, while the remaining 13 (48%) succumbed to clonal interference (Fig. 2 and SI Appendix, Fig. S3). In at least 2 of these 13 cases, a TM-specific mutation was outcompeted by expanding clones driven by generic mutations: in population V6, a TM-specific mutation in fisA was outcompeted by a clone carrying generic mutations in fimD, fisL, and hslO (Fig. 2), and in population P3, a TM-specific mutation in tufA was outcompeted by a clone carrying generic mutations in amiC and trkH (Fig. 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.

**Evolutionary Stalling in the Translational Machinery.** Phenotypic evolution of an organism is often interpreted through Fisher’s geometric model (72). In this model, the fitness of a genotype is a decreasing concave function of its distance to a phenotypic optimum. A key property of this model is that genotypes that are farther from the optimum have access to more beneficial mutations with large effects (72). Motivated by Fisher’s model, we hypothesized that TM adaptation is more likely to stall in founders with initially less severe TM defects because they are presumably closer to their performance optimum. To test this hypothesis, we examined how the TM-specific mutations are distributed among populations derived from different founders.

We found a total of 5, 10, and 12 TM-specific mutations in V, A, and P populations, respectively (0.8, 1.7, and 2.0 mutations per population on average; Fig. 3A). We also observed three, six, and two chromosomal amplifications spanning the tufA locus in these populations. In contrast, we did not observe any TM-specific

---

**Table 1. Founders used for the evolution experiment**

| Strain | EF-Tu origin species | Number of amino acid differences from E. coli EF-Tu (percent identity) | Fitness relative to E strain ± SEM, % per generation |
|--------|----------------------|------------------------------------------------|-------------------------------------------------|
| E      | E. coli (control)    | 0 (100)                                      | 0 ± 0.4                                         |
| S      | S. typhimurium       | 1 (99.75)                                    | +0.49 ± 0.05                                    |
| Y      | Y. enterocolitica    | 24 (93.91)                                   | −3.02 ± 0.02                                    |
| V      | V. cholerae          | 51 (87.06)                                   | −19.0 ± 0.7                                     |
| A      | Reconstructed ancestor | 21 (94.67)                                 | −34.4 ± 0.4                                    |
| P      | P. aeruginosa        | 62 (84.38)                                   | −35.0 ± 0.1                                    |

---

 Venkataram et al.  
 PNAS Latest Articles  | 3 of 9
mutations in the E, S, and Y populations (Fig. 3A). The differences in the observed rates of TM-specific mutations between founders are statistically significant (P = 4 × 10⁻¹¹, χ² test). Consistent with Fisher’s model, the number of TM-specific mutations observed in a founder was negatively correlated with its initial fitness (r = −0.993, P = 7.8 × 10⁻¹⁰). We also found consistent patterns in the distribution of fixed TM-specific mutations: at least one of them fixed in each of the sequenced A and P populations (Fig. 3A and SI Appendix, Fig. S3), but only three out of six sequenced V populations fixed one or more TM-specific mutations, assuming that all tufA amplifications present in the population at generation 1,000 are fixed (Fig. 3A and SI Appendix, Fig. S3). Thus, in the A and P genetic backgrounds, which have the most severe TM defects, TM-specific beneficial mutations are abundant enough and have large enough fitness effects that they drive adaptation. In the V founder, which has an intermediate TM defect, TM-specific mutations also contribute to adaptation but to a lesser degree, suggesting that they are less common and/or provide weaker benefits in this background. In the E, S, and Y founders whose TMs have only weak defects, TM-specific mutations either are not available or, if they are available, are so rare and/or so weak that they fail to cross our detection threshold.

If the supply and/or the fitness benefits of TM-specific mutations indeed increase with the magnitude of the TM defect, the contribution of such mutations to adaptation should shrink as populations accumulate TM-specific mutations. To test this hypothesis, we examined the temporal distribution of fixed adaptive mutations in V, A, and P populations.

Out of the 14 TM-specific mutations that eventually fixed, 12 (86%) did so in the first selective sweep. As a result, an average TM-specific beneficial mutation reached fixation by generation 300 ± 52, and only one (7%) reached fixation after generation 600 (Fig. 3B and SI Appendix, Fig. S5). In contrast, the average fixation time of generic mutations in the V, A, and P populations was 600 ± 72 generations, and nine of them (56%) fixed after the first selective sweep (Fig. 3B and SI Appendix, Fig. S5). As a consequence, TM-specific mutations were overrepresented among mutations that were fixed within the first 300 generations in the V, A, and P populations (P = 0.014, Fisher’s exact test). These numbers exclude 11 tufA amplifications, for which we do not have temporal resolution. We observe similar trends if we consider the average allele frequencies of mutations (Fig. 3C).

We conclude that the accumulation of TM-specific mutations in the A, P, and V populations has essentially ceased by the end of the evolution experiment, while the accumulation of adaptive mutations in other modules continued (Fig. 3B and C). In other words, natural selection initially improved both the TM and other cellular modules, but after approximately one TM-specific mutation was fixed, the focus of selection shifted away from the TM.

Why did TM-specific mutations fix early in evolution? As mentioned above, TM genes occupy only about 4.3% of the E. coli genome. It is therefore unlikely that TM-specific mutations arise at higher rates than generic mutations. Instead, their early fixation suggests that their fitness benefits must have been typically much larger than those of generic mutations. To test this hypothesis, we estimated selection coefficients of mutations from the mutation frequency trajectories. We found that mutations in the A and P populations that are detected by generation 300 provide an average selective advantage of 3.7% (n = 22), while mutations that are detected after generation 300 provide an average selective advantage of 1.6% (n = 28, P = 0.0017, t test). This approach likely underestimates the effects of the strongest mutations because we cannot accurately resolve selective sweeps that take ≥ 100 generations. To complement this approach, we genetically reconstructed two TM-specific mutations in their respective founder backgrounds and directly measured their fitness benefits (Materials and Methods, SI Appendix, Fig. S7, and Dataset S7). We found that mutation A74G in the rpsF gene, which was observed in the population A5, confers an
8.2 ± 0.4% fitness benefit to the A founder. Mutation G331A in *rpsG* gene, which was found in populations P2, P3, and P5, confers a fitness benefit of 6.5 ± 0.5% to the P founder (Dataset S8). Consistent with our hypothesis, these fitness benefits are much larger than the average estimated 1.6% fitness effect of late-arriving mutations (which are mostly generic). Even more strikingly, either of the two reconstructed mutations provides a larger fitness gain to its founder than the total fitness gains achieved by the E, S, and Y populations during the entire evolution experiment (Fig. 1B).

Our observations so far can be broadly interpreted in two ways. One explanation is that the TMs in E, S, and Y founders are unimprovable by natural selection because they are at their local performance peaks and that TMs in the V, A, and P founders become unimprovable after acquiring approximately one TM-specific mutation. The second interpretation is that natural selection fails to improve the TMs in E, S, and Y founders because TM-specific mutations in these backgrounds are too rare and/or too weak to drive adaptation, i.e., TM adaptation is stalled. At the same time, TM adaptation stalls after the fixation of approximately one TM-specific mutation. TM adaptation stalls when considering the hypothesis that all TMs are at their local performance peaks; instead, TM-specific mutations fail to keep accumulating in these populations because they are outcompeted by stronger beneficial mutations that improve other cellular modules, i.e., the TM exhibits evolutionary stalling.

Evolutionary stalling prevents a module from reaching its local performance peak and thereby imposes a genetic load, i.e., the organism carrying a stalled module suffers a fitness cost relative to an organism whose module performance is optimal. It is difficult to measure these loads in our populations directly because we do not know how far below their local optima they are. However, we can estimate these loads if we assume that strains to an organism whose founder has a wild-type TM (41). If the wild-type TM is improvable, it is likely that the TMs in our E, S, and Y strains—which carry mild defects—are also improvable. It is possible to imagine a structure of epistasis between mutations, such that each of these observations is consistent with the hypothesis that all TMs are at their local performance peaks. However, when considered together, these observations are more parsimoniously explained by the alternative hypothesis: at least some of the TMs in our experiment are not at their local performance peaks; instead, TM-specific mutations fail to keep accumulating in these populations because they are outcompeted by stronger beneficial mutations that improve other cellular modules, i.e., the TM exhibits evolutionary stalling.

Venkataram et al. PNAS Latest Articles | 5 of 9

---

Fig. 3. Variation in TM adaptation across founders and time. (A) Average number of fixed and nonfixed TM-specific and generic mutations per population. (B) Average number of fixed TM-specific and generic mutations per population over time in the populations derived from V, A, and P founders. (C) Average total allele frequency of TM-specific and generic mutations per population over time. Data in all panels do not include *tufA* amplifications.
However, TM adaptation in all of these populations stalled before the end of the evolution experiment. Thus, these populations suffer an average genetic load of at least ∼0.5% due to residual TM defects. In fact, the genetic loads could be much higher if the initial TM defects present in the founders have not been alleviated by mutations outside of the canonically annotated TM. Under this assumption, the fact that we did not observe any TM-specific mutations in population V4 would imply that it still carries an ∼19% genetic load imposed by the original TM defect in the V founder. By the same logic, all Y populations still carry an ∼3% genetic load. These estimates suggest that the power of natural selection to improve an individual module can be severely limited by evolutionary stalling.

Genetic Mechanisms Underlying Evolutionary Stalling in the TM.

Next, we asked what genetic mechanisms contributed to evolutionary stalling in the TM. We begin with the observation that TM adaptation is stalled or stopped in some founders (E, S, and Y) and initially not stalled in others (V, A, and P). This observation indicates that founders have access to different pools of beneficial mutations, i.e., evolutionary stalling is caused by historical contingency epistasis (73). To further characterize the extent of historical contingency on evolution in our experiment, we examined the distribution of beneficial TM-specific and generic mutations across populations descended from different founders.

We found that TM-specific mutations were not uniformly distributed among V, A, and P populations. Specifically, four out of seven classes of TM-specific mutations arose in a single founder (Fig. 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; Materials and Methods). Similarly, all four mutations in the rpsF gene, which encodes the ribosomal protein S6, occurred in the A founder (P < 10⁻⁴). We have shown above that adaptive alleles at these loci confer strong fitness benefits in their respective founder backgrounds. However, our multiple reconstruction attempts in all other founders were unsuccessful (Materials and Methods), which suggests that these mutations are strongly deleterious in other genetic backgrounds and would explain why we did not observe these mutations in other founders. This result indicates that the severity of the defect in the TM is the not the only determinant of the availability of TM-specific beneficial mutations. Instead, defects caused by different foreign tufA alleles open up different adaptive pathways within the TM.

The distribution of some generic mutations among founders was also nonuniform. We found that 7 out of 22 classes of generic mutations occurred in fewer founders than expected by chance (Fig. 4B and Materials and 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). Similarly, all three mutations in the alaA gene, which encodes a glutamate-pyruvate aminotransferase, occurred in the A founder (P < 10⁻⁴). 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 (Dataset S8). As expected, this mutation confers a 5.9 ± 0.4% fitness benefit in the V founder. In contrast, it is strongly deleterious in the P founder and indistinguishable from neutral in the remaining founders (SI Appendix, Fig. S6). Thus, historical contingency epistasis not only opens up or closes down certain adaptive pathways within the TM but also changes the availability of at least some beneficial mutations outside of it.

We next turn our attention to the genetic mechanisms for the onset of evolutionary stalling in the V, A, and P founders. Historical contingency could play a role here too if mutations that accumulate early in evolution alter the identities and fitness effects of subsequent beneficial mutations (39, 41). Additionally, another mechanism, known as “modular epistasis” (39) or “coupon collecting” (41), could also operate if fixation of a mutation depletes the pool of further beneficial mutations in the same module, without changing the effects of mutations in other modules (for example, one beneficial loss-of-function substitution eliminates the benefits of other loss-of-function mutations in the same gene). Modular epistasis could lead to evolutionary stalling in the TM because the most frequent TM-specific mutations with the largest fitness benefits tend to fix earlier thereby enriching the pool of available TM-specific mutations for rare and weak mutations which are more likely to fall below the emergent neutrality threshold (41). If repeated mutations in a given gene occur unusually often in the same population (underdispersion), this signals historical contingency epistasis with unknown mutations present in that population (41). On the other hand, unexpectedly even distribution of mutations in a given gene across different populations signals the presence of modular epistasis (41).

![Figure 4](https://example.com/figure4.png)

**Figure 4.** Distribution of putatively adaptive mutations across genetic backgrounds. Heat map of all putatively adaptive mutations identified via whole-genome sequencing, grouped by founder and by gene. (A) Translation-associated genes. Amplifications of the tufA locus are counted separately from other mutations in tufA. (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; Materials and Methods). Numbers in parentheses indicate the total number of mutations in each gene observed across all sequenced populations.

6 of 9 | www.pnas.org/cgi/doi/10.1073/pnas.1921881117 Venkataram et al.
To detect historical contingency and modular epistasis, we examined distribution of mutations across populations (tufA amplifications were excluded from this analysis; Materials and Methods). We find that TM-specific mutations collectively are overdispersed ($P = 0.0018$). That is, TM-specific mutations at any given locus usually arise only once per population. The distribution of generic mutations collectively is consistent with the random expectation. However, mutations in some individual genes show opposing tendencies. Specifically, mutations in genes trkH and fimD are overdispersed (corrected $P < 10^{-4}$), while mutations in icd and yedC are underdispersed (corrected $P = 0.004$ and $<10^{-5}$). Taken together, these results show that both historical contingency and modular epistasis have both likely contributed to evolutionary stalling in our populations.

**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 may be difficult for it 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 large effect mutations occur frequently, while improvements in other modules are expected to stall. Here we have documented such evolutionary stalling in the TM module during laboratory adaptation of *E. coli*.

We found that populations whose TMs were initially mildly perturbed (incurred ≤3% fitness cost) adapted by acquiring mutations that did not directly affect the TM. Populations whose TM had a moderately severe defect (incurred ~19% fitness cost) discovered TM-specific mutations, but clonal interference often prevented their fixation. Populations whose TMs were initially severely perturbed (incurred ~35% fitness cost) rapidly discovered and fixed TM-specific beneficial mutations. However, after approximately one TM-specific mutation fixed, further accumulation of mutations in the TM essentially stalled, while mutations in other modules continued to accumulate despite the fact that the TMs remained improvable in at least some of our populations.

We estimated that the genetic load imposed by evolutionary stalling in the TM ranges from 0.5 to 19% in our populations. These estimates are inherently uncertain because we do not know the exact fitness effects of all fixed mutations. The true genetic loads may be higher than 0.5% because some of the fitness gains in the A and P populations are probably caused by improvements in other cellular modules and because the optimal TM probably provides fitness higher than that of the E strain (since the E strain itself has a TM defect). On the other hand, true genetic loads may be lower than 19% because TM performance may be improved by some mutations in genes not annotated as translation-related.

In our analyses, we defined the TM module operationally as a set of 215 genes in the *E. coli* genome that are annotated as translation-related. However, there may be other genes that contribute to translation that are not annotated as such. A more fundamental problem is that while the core of the translation machinery is well defined, its periphery may in reality be quite amorphous. In other words, a model of a cell where modules are discrete and clearly delineated is an abstraction. For example, a recent study found that a genetic defect in the DNA replication module can be alleviated by mutations outside of it (53). Similarly, we found that populations with different initial tufA defects acquire mutations in statistically distinct sets of genes that are not annotated as translation-related, suggesting that some of these genes may nevertheless be important for translation. Nevertheless, our results show that the focus of natural selection can transition sharply from one set of genes—those encoding the core of the translation machinery where the initial defect occurred—to functionally more distant parts of the cell such as genes trkH and fimD that are involved with potassium transport and fimbiae production, respectively.

As long as the environment remains constant, the supply of beneficial mutations in an adapting population is gradually being depleted, and their fitness effects typically decrease (39, 55–57, 74, 75), thereby lowering the effective neutrality threshold. These changes should in turn allow for less frequent mutations with smaller effects to contribute to adaptation, and adaptation in previously stalled modules may resume. While we did not observe resumption of adaptive evolution in the TM during the duration of this experiment, we find evidence for a transition from stalling to adaptation in trkH and fimD genes. Mutations in these two genes appear to be beneficial in all our genetic backgrounds (Fig. 4). These mutations are among the earliest to arise and fix in E, S, and Y populations where the TM does not adapt (SI Appendix, Figs. S3 and S5). In contrast, mutations in trkH and fimD arise in A and P populations much later, typically following fixations of TM-specific mutations (SI Appendix, Figs. S3 and S5). In other words, natural selection in these populations is initially largely focused on improving the TM, while adaptation in trkH and fimD is stalled. After a TM-specific mutation is fixed, the focus of natural selection shifts away from the TM to other modules, including trkH and fimD.

As the external environment changes, new large-effect adaptive mutations may become available in some modules, which would increase the effective neutrality threshold. This may in turn lead to the onset of evolutionary stalling or prolong the period of stalling in other modules. If environmental fluctuations are sufficiently frequent, e.g., due to seasonality or ecological interactions, and if these fluctuations preferentially open up new adaptive mutations in a subset of cellular modules (e.g., in stress response but not in housekeeping modules), some modules may remain stalled for long periods of time despite being improvable, at least in the absence of recombination.

In addition to evolutionary stalling, adaptive evolution in the concurrent mutations regime can have other important consequences for the biology of the organism. For example, Held et al. recently showed that fixation of deleterious hitchhiker mutations imposes a limit on the complexity of the organism (36). They also showed that recombination would alleviate this limit. Similarly, recombination is expected to reduce the genetic loads imposed by evolutionary stalling, provided that modules are encoded by tightly linked genes.

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 (72, 76, 77) 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 form either a single contiguous neutral network (78) or multiple isolated neutral networks (79). 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 (78, 80).

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 (which consists of tufA, fusA, rpsG, and rpsL) 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 (46, 81, 82). Directly measuring the
phenotypic effects of the TM-specific mutations described here is an important avenue for future work.

Overall, our results highlight the fact that it is impossible to fully understand the evolution of 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 populations adapt. 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

Isogenic strains were reconstructed using previously described methods (61). Laboratory evolution experiments were conducted with a standard batch culture protocol with daily 1:10,000 dilution, in LB medium, at 37 °C and shaken at 200 rpm on a platform incubator–shaker. Competitive fitness assays were conducted using evolved and ancestral strains engineered with Kanamycin and Ampicillin resistance cassettes and followed established methods (27). Growth was conducted using evolved and ancestral strains engineered with Kanamycin and Ampicillin resistance cassettes and followed established methods (27). Growth rates were estimated by growing strains in a microplate reader with temperature control and measuring optical density at 600 nm. Whole-genome sequencing was conducted using the Illumina HiSeq 4000 platform and was analyzed using established bioinformatic pipelines (41). Details of the experimental model, media and growth conditions, fitness and growth measurements, genome sequencing, and all statistical analyses are provided in SI Appendix.

Data Availability. Raw sequencing data were analyzed with the Python-based workflow implemented in ref. 58 and run on the University of California San Diego Triton Shared Computing Cluster via a custom Python wrapper script. All analyses and plots reported in this manuscript have been performed using the R computing environment. The script, modified reference genomes, and raw data (except for raw sequencing data) used for analysis can be found on GitHub at https://github.com/kandeepenjakatara/EvoStalling (83). Raw sequencing data for this project have been deposited into the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under project PRJNA560969 (84).

ACKNOWLEDGMENTS. We thank members of the Kryazhimskiy and Kacar groups, Joanna Masel, Ryan Gutenkunz, Suparna Sanyal, Grant Kinsler, Justin Meyer, and Lin Chao for input and feedback. We thank Alex Pilca, Divya K. Rau, Emily Peñaherrera, Kevin Longoria, Ledy Villarreal, Alena Martsul, and Sarah Ardell for laboratory assistance. We thank Huanyu Kuo for the analysis of growth curve data. We thank Eva Garmendia for providing the recombining plasmids and Georg Rieck 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 UCSF Institute for Genomic Medicine for sequencing services and the San Diego Supercomputing Center for providing the computational environment. B.K. acknowledges the support by the John Templeton Foundation (grants 58562 and 61239), the NASA Exobiology and Evolutionary Biology Program (grant H006201406), and the NASA Astrobiology Institute (grant NNA17BB05A). S.K. acknowledges the support by Burroughs Wellcome Fund Career Award at the Scientific Interface (grant 101079.01), the Alfred P. Sloan Foundation (grant FG-2017-9227), and the Hellman Foundation.

1. H. A. Simon, The architecture of complexity. Proc. Am. Philos. Soc. 106, 467–482 (1962).

2. J. W. Valentine, C. L. May, Hierarchies in biology and paleontology. Paleobiology 22, 23–33 (1996).

3. D. W. McShea, The hierarchical structure of organisms: A scale and documentation of a trend in the maximum. Paleobiology 27, 405–423 (2001).

4. E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, A. L. Barabási, Hierarchical organization of metabolic networks. Science 297, 1551–1555 (2002).

5. H. Mengistu, J. Huizinga, J. B. Mount, J. Clune, The evolutionary origins of hierarchy. PLoS Comput. Biol. 12, e1004829 (2016).

6. L. H. Hartwell, J. J. Hopfield, S. Leibler, A. W. Murray, From molecular to modular cell biology. Nature 402 (suppl. 6761), C47–C52 (1999).

7. G. P. Wagner, M. Pavlicev, J. M. Cheverud, The road to modularity. Nat. Rev. Genet. 8, 921–931 (2007).

8. C. Woese, Molecular mechanics of translation: A reciprocating ratchet mechanism. Nature 226, 817–820 (1970).

9. L. A. Simon, Evolution in glib. J. Theor. Biol. 187, 481–498 (2001).

10. L. A. Simon, Evolutionary advantages of modularity. J. Theor. Biol. 187, 481–498 (2001).

11. S. Gleizer, Evolution in glib. J. Theor. Biol. 187, 481–498 (2001).

12. M. M. Desai, D. S. Fisher, Beneficial mutation selection balance and the effect of linkage on positive selection. Genetics 176, 1759–1798 (2007).

13. R. A. Neher, B. I. Shraiman, Competition between recombination and epistasis can cause a transition from allele to genotype selection. Proc. Natl. Acad. Sci. U.S.A. 106, 6866–6871 (2009).

14. B. I. Shraiman, I. M. Rouzine, J. D. Balick, O. Hallatschek, M. M. Desai, Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proc. Natl. Acad. Sci. U.S.A. 109, 4950–4955 (2012).

15. S. F. Levy et al., Quantitative evolutionary dynamics using high-resolution lineaging. Proc. Natl. Acad. Sci. U.S.A. 115, 16990–16994 (2018).

16. B. I. Shraiman, I. M. Rouzine, J. D. Balick, O. Hallatschek, M. M. Desai, Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proc. Natl. Acad. Sci. U.S.A. 109, 4950–4955 (2012).

17. D. L. Hartl, D. E. Dykhuisen, A. M. Dean, Limits of adaptation: The evolution of selective neutrality. Genetics 111, 655–674 (1985).

18. D. L. Hartl, C. H. Taubes, Towards a theory of evolutionary adaptation. Genetics 102–103, 525–533 (1998).

19. M. Lynch, Evolutionary and the limits to cellular perfection. Proc. Natl. Acad. Sci. U.S.A. 109, 18851–18856 (2012).

20. A. Nourmohammad, S. Schiffels, M. Lüssig, Evolution of molecular phenotypes under stabilizing selection. J. Stat. Mech: Theory Exp. 2013, P01012 (2013).

21. S. Goyal et al., Dynamic mutation-selection balance as an evolutionary attractor. Genetics 191, 1309–1319 (2012).

22. D. P. Rice, B. H. Good, M. M. Desai, The evolutionarily stable distribution of fitness effects. Genetics 200, 321–329 (2015).

23. T. Held, D. Klemmer, M. Lüssig, Survival of the simplest in microbial evolution. Nat. Commun. 10, 2472 (2019).

24. S. Schiffels, G. J. Sziklai, V. Mustonen, M. Lüssig, Emergent neutrality in adaptive evolution. Genetics 189, 1361–1375 (2015).

25. G. I. Lang, D. Bostein, M. M. Desai, Genetic variation and the fate of beneficial mutations in asexual populations. Genetics 188, 647–661 (2011).

26. S. Kryazhimskiy, D. P. Rice, E. R. Jirson, M. M. Desai, Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science 326, 1519–1522 (2014).

27. E. R. Jirson et al., Genetic variation in adaptability and pleiotropy in budding yeast. eLife 6, e27167 (2017).

28. B. H. Good, M. J. McDonald, J. E. Barrick, R. E. Lenski, M. M. Desai, The dynamics of molecular evolution over 60,000 generations. Nature 511, 45–50 (2014).

29. J. I. Rojas Echenique, S. Kryazhimskiy, A. N. Nguyen Ba, M. M. Desai, Modular epistasis and the compensatory evolution of gene deletion mutants. PLoS Genet. 15, e1007958 (2019).

30. E. R. Jirson et al., Genetic variation in adaptability and pleiotropy in budding yeast. eLife 6, e27167 (2017).

31. S. Kryazhimskiy, D. P. Rice, E. R. Jirson, M. M. Desai, Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science 326, 1519–1522 (2014).

32. E. R. Jirson et al., Genetic variation in adaptability and pleiotropy in budding yeast. eLife 6, e27167 (2017).

33. S. Kryazhimskiy, D. P. Rice, E. R. Jirson, M. M. Desai, Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science 326, 1519–1522 (2014).

34. R. Coudago, S. Chen, Y. Shamo, In vivo molecular evolution reveals biophysical origins of organismal fitness. Mol. Cell 22, 441–449 (2006).
54. Y. T. Tamer et al., Growth phase-specific evolutionary benefits of natural trans-
formation in Acinetobacter baylyi. ISME J. 9, e1003972 (2015).

55. M. Fumasoni, A. W. Murray, The evolutionary plasticity of chromosome metabolism
allows adaptation to constitutive DNA replication stress. eLife 9, e51963 (2020).

56. A. I. Khan, D. M. Dinh, D. Schneider, R. E. Lenski, T. F. Cooper, Negative epistasis
between beneficial mutations in an evolving bacterial population. Science 335, 1193–1196.e22 (2011).

57. A. Wünsche et al., Diminishing-returns epistasis decreases adaptability along an
evolutionary trajectory. Nat. Ecol. Evol. 1, 61 (2017).

58. O. Tenaillon et al., The molecular diversity of adaptive convergence. Science 335, 457–461 (2012).

59. D. J. Kvitek, G. Sherlock, Reciprocal sign epistasis between frequently experimentally
adapted mutations causes a rugged fitness landscape. PLoS Genet. 7, e1002056 (2011).

60. V. Ramakrishnan, Ribosome structure and the mechanism of translation. Cell 108, 557–572 (2002).

61. B. Kacar, E. Garmedina, N. Tuncbag, D. I. Andersson, D. Hughes, Functional constraints on
replacing an essential gene with its ancient and modern homologs. mBio 8, e01276-17 (2017).

62. J. Nilson, P. Nissen, Elongation factors on the ribosome. Curr. Opin. Struct. Biol. 15, 349–354 (2005).

63. M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, T. Hwa, Interdependence of cell
growth and gene expression: Origins and consequences. Science 330, 1099–1102 (2010).

64. R. Schnell, F. Abdualkarim, M. Kälmän, L. A. Isaksson, Functional EF-Tu with large
terminal extensions in an E. coli strain with a precise deletion of both chromosomal
tuf genes. FEBS Lett. 538, 139–144 (2003).

65. R. E. Lenski et al., Evolution of competitive fitness in experimental populations of
E. coli. What makes one genotype a better competitor than another? Antonie van
Leeuwenhoek 73, 35–47 (1998).

66. Y. Li et al., Hidden complexity of yeast adaptation under simple evolutionary con-
ditions. Curr. Biol. 28, 515–525.e6 (2018).

67. M. Manhart, B. V. Adkar, E. I. Shakhnovich, Trade-offs between microbial growth
phases lead to frequency-dependent and non-transitive selection. Proc. Biol. Sci. 285, 20172459 (2018).

68. F. Vasi, M. Travisano, R. E. Lenski, Long-term experimental evolution in Escherichia
coli. II. Changes in life-history traits during adaptation to a seasonal environment. Am.
Nat. 164, 432–456 (1999).

69. A. L. G. Utes et al., Growth phase-specific evolutionary benefits of natural trans-
formation in Acinetobacter baylyi. ISME J. 9, 2221–2231 (2015).

70. M. D. Herron, M. Doebeli, Parallel evolutionary dynamics of adaptive diversification in
Escherichia coli. PLoS Biol. 11, e1001490 (2013).

71. D. J. Kvitek, G. Sherlock, Whole genome, whole population sequencing reveals that
loss of signaling networks is the major adaptive strategy in a constant environment. PLoS Genet.
9, e1003972 (2013).

72. O. Tenaillon, The utility of Fisher’s geometric model in evolutionary genetics. Annu.
Evol. Ecol. Res. 45, 179–201 (2014).

73. Z. D. Blount, R. E. Lenski, J. B. Losos, Contingency and determinism in evolution:
Replaying life’s tape. Science 362, eaam5979 (2018).

74. R. C. Maclean, G. G. Perron, A. Gardner, Diminishing returns from beneficial muta-
tions and pervasive epistasis shape the fitness landscape for rifampicin resistance in
Pseudomonas aeruginosa. Genetics 186, 1345–1354 (2010).

75. X. Wei, J. Zhang, Patterns and mechanisms of diminishing returns from beneficial
mutations. Mol. Biol. Evol. 36, 1008–1021 (2019).

76. R. Fisher, The Genetical Theory of Natural Selection (Oxford University Press, 1930).

77. H. A. Orr, The distribution of fitness effects among beneficial mutations in Fisher’s
genetic model of adaptation. J. Theor. Biol. 238, 279–285 (2006).

78. A. Wagner, Neutralism and selectionism: A network-based reconciliation. Nat. Rev.
Genet. 9, 965–974 (2008).

79. S. Schaper, I. G. Johnston, A. A. Louis, Epistasis can lead to fragmented neutral spaces
in natural populations. Annu. Rev. Ecol. Evol. Syst. 45, 1–38 (2014).

80. Tenaillon et al., Hidden complexity of yeast adaptation under simple evolutionary con-
ditions. Curr. Biol. 28, 515–525.e6 (2018).

81. S. Venkataram, R. Monasky, S. H. Sikaroodi, S. Kryazhimskiy, B. Kacar, EvoStalling.
GitHub. https://github.com/sandeepvenkataram/EvoStalling. Deposited 20 November
2019.