

RESEARCH ARTICLE

High mutation rates limit evolutionary adaptation in *Escherichia coli*

Kathleen Sprouffske^{1,2}, José Aguilar-Rodríguez^{1,2*}, Paul Sniegowski³,
Andreas Wagner^{1,2,4*}

1 Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland, **2** The Swiss Institute of Bioinformatics, Lausanne, Switzerland, **3** Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, **4** The Santa Fe Institute, Santa Fe, New Mexico, United States of America

* Current address: Department of Biology, Stanford University, Stanford, California, United States of America; Department of Chemical and Systems Biology, Stanford University School of Medicine Stanford, California, United States of America

* andreas.wagner@ieu.uzh.ch



OPEN ACCESS

Citation: Sprouffske K, Aguilar-Rodríguez J, Sniegowski P, Wagner A (2018) High mutation rates limit evolutionary adaptation in *Escherichia coli*. PLoS Genet 14(4): e1007324. <https://doi.org/10.1371/journal.pgen.1007324>

Editor: Ivan Matic, Université Paris Descartes, INSERM U1001, FRANCE

Received: July 7, 2017

Accepted: March 21, 2018

Published: April 27, 2018

Copyright: © 2018 Sprouffske et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All files are available from the Dryad database (doi:[10.5061/dryad.mh206](https://doi.org/10.5061/dryad.mh206)).

Funding: KS acknowledges support through the Forschungskredit program of the University of Zurich, grant K-74301-03-0. AW acknowledges support by ERC Advanced Grant 739874, by Swiss National Science Foundation grant 31003A_146137, by an EpiphysX RTD grant from SystemsX.ch, as well as by the University Priority Research Program in Evolutionary Biology at the University of Zurich. The funders had no role in

Abstract

Mutation is fundamental to evolution, because it generates the genetic variation on which selection can act. In nature, genetic changes often increase the mutation rate in systems that range from viruses and bacteria to human tumors. Such an increase promotes the accumulation of frequent deleterious or neutral alleles, but it can also increase the chances that a population acquires rare beneficial alleles. Here, we study how up to 100-fold increases in *Escherichia coli*'s genomic mutation rate affect adaptive evolution. To do so, we evolved multiple replicate populations of asexual *E. coli* strains engineered to have four different mutation rates for 3000 generations in the laboratory. We measured the ability of evolved populations to grow in their original environment and in more than 90 novel chemical environments. In addition, we subjected the populations to whole genome population sequencing. Although populations with higher mutation rates accumulated greater genetic diversity, this diversity conveyed benefits only for modestly increased mutation rates, where populations adapted faster and also thrived better than their ancestors in some novel environments. In contrast, some populations at the highest mutation rates showed reduced adaptation during evolution, and failed to thrive in all of the 90 alternative environments. In addition, they experienced a dramatic decrease in mutation rate. Our work demonstrates that the mutation rate changes the global balance between deleterious and beneficial mutational effects on fitness. In contrast to most theoretical models, our experiments suggest that this tipping point already occurs at the modest mutation rates that are found in the wild.

Author summary

Mutation is of central importance in biology. It creates genetic variation, the raw material of evolution by natural selection, It can improve traits and organisms, but can also lead to phenomena like cancerous cells and antibiotic resistant pathogens. Increasing the mutation rate can accelerate evolutionary adaptation, even over many thousands of generations

study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

in a constant environment. Our study describes the laboratory evolution of asexual *Escherichia coli* strains with a range of mutation rates at levels found in the wild (from wild type to strong mutator). Unexpectedly, evolutionary adaptation was most limited in the populations with the highest mutation rate. Our work suggests that deleterious mutations can begin to limit adaptation at lower mutation rates than previously thought.

Introduction

Mutation is fundamental to evolution. Without it, evolution cannot occur, because mutation provides the genetic variation necessary for selection and genetic drift. Each new mutation in an individual can increase its fitness, decrease its fitness, or have no effect on its fitness. Unfortunately, most mutations with fitness effects are deleterious, and fitness-increasing beneficial mutations constitute only a small fraction of all possible mutations [1]. The mutation rate can itself evolve, because it is subject to genetic change in the "mutation rate genome", the part of a genome encoding DNA replication and repair systems [2,3]. Here, we characterize the long-term effects of a range of mutation rates on adaptation, as well as the evolution of the mutation rate itself, by evolving multiple replicate populations of asexual *Escherichia coli* in a minimal medium in the laboratory.

Evolutionary adaptation under increased mutation pressure in large non-recombining populations like ours has been explored in past work (all mutations that occur in our *E. coli* laboratory strain's genome are linked). The joint effects of mutation and linkage on selection (and the related topics of diversity and the evolution of sex) have been much studied since Fisher [4] and Muller [5] ([6–10], recently reviewed in [11–14]). Under increased mutation pressure, multiple clones within a population may acquire new mutations, and then compete with each other for fixation. While relevant studies show that the speed of adaptation can increase with the genomic mutation rate [10,15–18], they leave open the possibility that extremely high mutation rates could hinder adaptation. This possibility is raised by a variety of models that predict declining fitness in populations with extreme mutation rates. An early, influential, but simple model predicted that a population's fitness will decrease when the rate of mutation increases beyond a critical "error threshold" [19] whose value depends on model details. Other models of populations evolving at high mutation rates are more realistic and take into account phenomena like beneficial mutations and demography. However, they also predict that adaptation can be slowed and eventually reversed at sufficiently high mutation rates by the effects of deleterious mutations [20–24].

Many studies have documented the evolution of increased mutation rates [25–31], which can evolve in certain conditions. For example, after a recent environmental change that creates opportunities for novel adaptations and new beneficial mutations [32,33], a cell with a mutator allele is more likely to produce large-effect beneficial mutations than a cell with a wild-type mutation rate. Because of their improved fitness, cell lineages with newly acquired beneficial alleles (and their linked mutator alleles) can increase in frequency in the population. Thus, hypermutation can readily evolve when mutator alleles hitchhike to fixation with beneficial mutations [34–37].

In the long term, however, hypermutation can be detrimental, because most non-neutral mutations have deleterious consequences [1]. Thus, an individual with a higher mutation rate may accumulate more deleterious mutations overall, which can result in lower fitness. For this reason, selection has been predicted to reduce mutation rates [38]. However, there are several potential reasons why mutation rates may not decline all the way to zero. One of them is that

the physiological mechanisms required to improve replication fidelity and DNA repair carry a fitness cost [39–42]. Another is that the power of selection to reduce the mutation rate is limited by population size via the so-called drift-barrier [43,44]. Experimental observations of evolved reductions in the mutation rate have been reported, but are relatively infrequent [27,31,45–50] (reviewed in [51]).

While some previous experiments explored the adaptive responses and mutation rate changes that can take place under increased mutational pressure [46–48,50], they focused on one or two mutation rates, and did not include genomic analyses (except [50]). Here, we sought to provide a uniquely comprehensive empirical data set across a range of mutation rates, including whole genome population sequencing data, mutation rate data, and fitness measurements in a number of environments. To do so, we engineered four isogenic *E. coli* K12 MG1655 derivative strains with increased mutation rates and evolved eight replicate populations of each strain for 3000 generations in a serial-transfer experiment. Genomic mutation rates differed more than a hundred-fold among these strains and ranged from $U = 0.00034$ to $U = 0.036$ point mutations per genome per generation by one method of estimation. During evolution, we periodically characterized the growth rate and stationary population density of each population. We also assayed the fitness of evolved populations in a variety of stressful environments. High-throughput population sequencing allowed us to characterize how far our populations spread through sequence space, and to study the mutations occurring in each population.

Results

The experimental design is summarized in Fig 1. We evolved eight independent replicate populations for each of four isogenic *E. coli* strains with increasing mutation rates, and did so for 175 days (about 3000 generations) in minimal medium. The mutation rates, as estimated by mutation to rifampicin resistance, differed 139-fold between the ancestor strains with the lowest and highest mutation rates. We called these strains MR^S , MR^M , MR^L , and MR^{XL} for strains

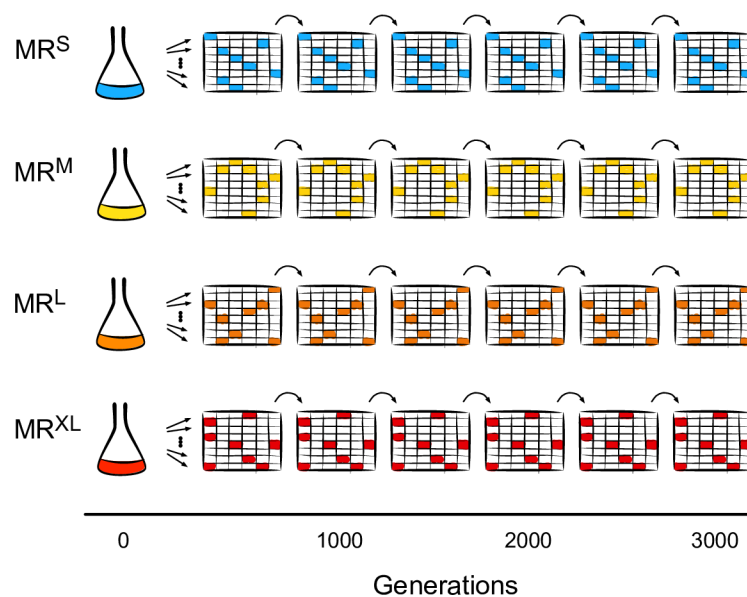


Fig 1. Experimental design. We evolved eight replicate populations for each of four *E. coli* strains with increasing mutation rates for nearly 3000 generations, and sequenced all evolving populations approximately every 1000 generations. All eight replicate populations with the same mutation rate shared a single, common ancestor.

<https://doi.org/10.1371/journal.pgen.1007324.g001>

with small, medium, large, and extra-large mutations rates (MR). We evolved all 32 replicate populations at 37°C in 2 mL minimal medium, and diluted them daily 100,000-fold into fresh medium. Because the population density at carrying capacity differed among strains (S1 Fig), the notional effective population size of the evolving replicates also differed, ranging between 42,500 and 92,800 (S2A Fig). Linkage is known to reduce the effective population size [52], so we also estimated its effect on this size (S2B Fig) [53–55]. We found that populations with higher mutation rates generally had lower effective population sizes. We sequenced samples of the evolving populations at generations 0, 1000, 2000, and 3000 to high genome-wide coverage. We measured the mutation rates and tested growth in stressful environments at generations 0 and 3000. We also periodically measured maximum growth rates and growth curves as a proxy for fitness.

Very high mutation rates can lead to reduced adaptation

We used a population’s maximum growth rate during exponential growth as a proxy for fitness, and we refer to relative fitness as the difference between an evolved and a reference population, usually the ancestral population. We first measured the fitness of the ancestral strains relative to the *E. coli* K12 MG1655 strain and found that the ancestral strains had similar fitness values, with the exception of the MR^L strain, which had much lower fitness than the other strains (S3A Fig, linear mixed effects analysis, $X^2(4) = 248$; MR^S: 0.34 ± 0.02 ; MR^M: 0.32 ± 0.02 ; MR^L: 0.11 ± 0.02 ; MR^{XL}: 0.25 ± 0.02 , relative fitness \pm s.e.m., $p < 2 \times 10^{-16}$). We next measured the fitness of each evolving replicate population relative to its ancestor at several time points during the experiment (Fig 2, S3B Fig, S4A Fig). At generation 3000, replicate populations with higher mutation rates showed a greater increase in fitness, except for the MR^{XL} strain, which

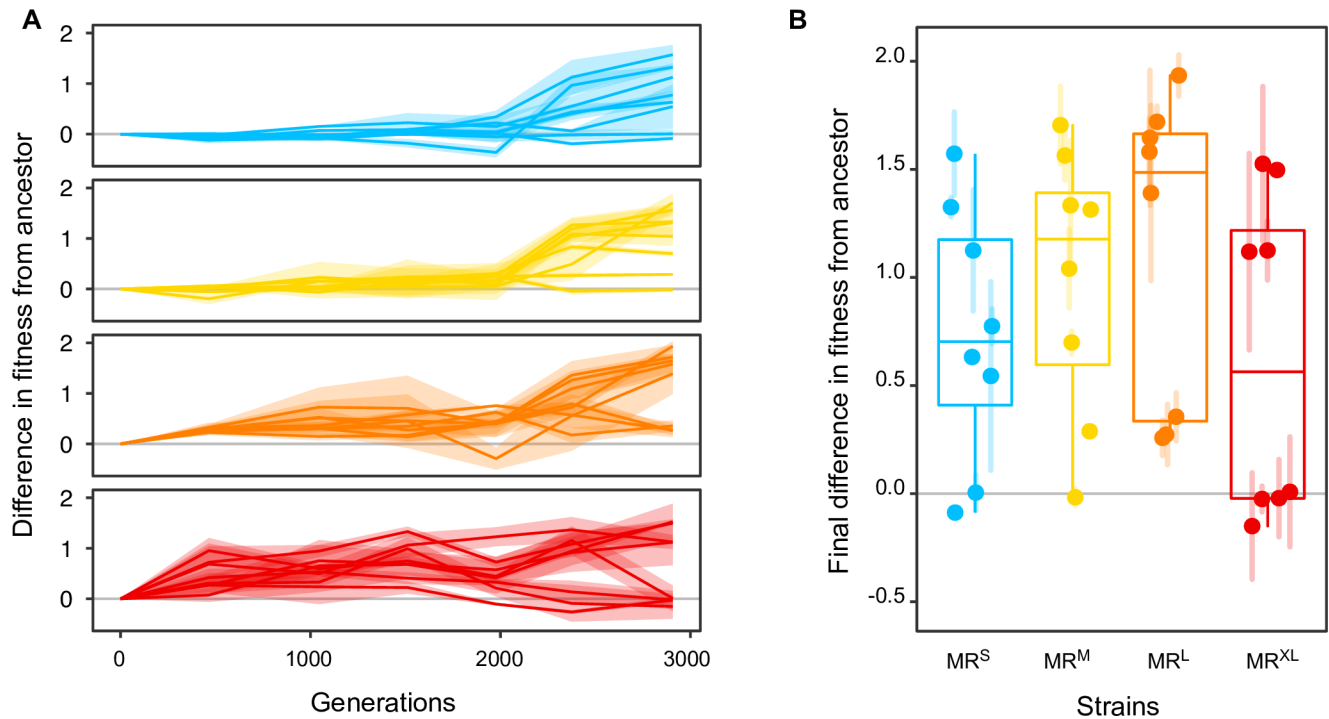


Fig 2. Fitness of the evolving replicate populations relative to their ancestors (A) over time, and (B) at the end of the experiment. A relative fitness value greater than zero indicates that the evolved population has higher fitness than its ancestor. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains. Shaded areas (A) or bars (B) indicate s.e.m.

<https://doi.org/10.1371/journal.pgen.1007324.g002>

had the smallest fitness increase of all strains ($p = 0.002$, linear mixed effects analysis, $X^2(4) = 16$; simultaneous tests for general linear hypotheses that relative fitness is unchanged at generation 3000: $MR^S: 0.7 \pm 0.2$, $p = 0.004$; $MR^M: 1.0 \pm 0.2$, $p < 4 \times 10^{-5}$; $MR^L: 1.1 \pm 0.2$, $p < 1 \times 10^{-6}$; $MR^{XL}: 0.6 \pm 0.2$, $p = 0.02$, evolved fitness difference from ancestor \pm s.e.m., significance). In replicate populations with lower mutation rates (MR^S and MR^M) fitness only began to rise substantially after 1000 generations, while in replicate populations with higher mutation rates (MR^L , MR^{XL}) this fitness increase began earlier (Fig 2A; linear mixed effects analysis, $X^2(4) = 17$, $p = 0.002$; simultaneous tests for general linear hypotheses that relative fitness is unchanged at generation 1000: $MR^S: -0.01 \pm 0.07$, $p = 1.0$; $MR^M: 0.07 \pm 0.07$, $p = 0.76$; $MR^L: 0.40 \pm 0.07$, $p < 0.0001$; $MR^{XL}: 0.57 \pm 0.07$, $p < 0.0001$, evolved fitness difference from ancestor \pm s.e.m., significance). A second growth curve metric that integrates information about the lag phase, growth rate, and carrying capacity yielded similar results (S1 Text). The delays in fitness gains in populations with lower mutation rates are also reflected in reduced fitness variation among replicate populations at any one point in time (S4B Fig). MR^L and MR^{XL} populations seem to form two clusters with either high or low (ancestral-like) fitness at the end of the experiment (Fig 2B). In principle, the fitness of populations with lower fitness might either not have increased at all, or it might have increased at first and subsequently decreased again. Examination of individual fitness trajectories (S4A Fig) shows that most MR^{XL} replicate populations with low fitness at generation 3000 gained and then lost fitness again.

Population sequencing at regular intervals

We sequenced a sample of each heterogeneous evolving population rather than a clone isolated from each population, so that we could estimate the genetic diversity within each sequenced population. We sequenced 100 populations in total: the four ancestor populations with different mutation rates, and eight replicates evolved from each ancestor at generations 1000, 2000, and 3000. Specifically, we sequenced the four ancestor populations after one day of growth just before having split them into their replicate populations. The mean sequence coverage for populations was 364-fold (standard deviation 98), with 99% (83%) of the samples having at least 100-fold (200-fold) coverage across at least 95% of the genome. In virtually all sequenced samples less than 1% of the genome had no sequence coverage (S5 Fig). We identified the frequency of SNPs in each sequenced population and their annotations using breseq, which has been widely used in microbial studies and has been optimized for bacterial data [56] (S6 Fig). We discovered several SNPs at non-zero frequency in these sequenced ancestral populations (1, 3, 2, and 64 loci for MR^S , MR^M , MR^L , and MR^{XL} , respectively; S7 Fig), some of which may have been transferred to the evolving populations. Previous studies had suggested that there may be biases in the mutational spectra caused by the reduced efficacy of the *mutL* and *dnaQ* gene products [57–59], but our sequence data shows that any such bias is weak or absent in our strains (S8 Fig).

Higher mutation rates lead to larger mutant clouds and more high frequency derived alleles

One can view an evolving population as a cloud of mutant individuals in sequence space. We suspected that this mutant cloud would be spread out further in genotype space—indicating greater standing diversity—for populations with a higher mutation rate. To test whether this was indeed the case, we first defined the center of an evolving population as its consensus sequence and then computed the average distance of each population to this consensus. A strain's mutation rate affected the size of its mutant cloud (linear mixed effects analysis, cube root of diversity taken to ensure homoscedasticity, $X^2(4) = 79$; $p = 3 \times 10^{-16}$; see Methods), such

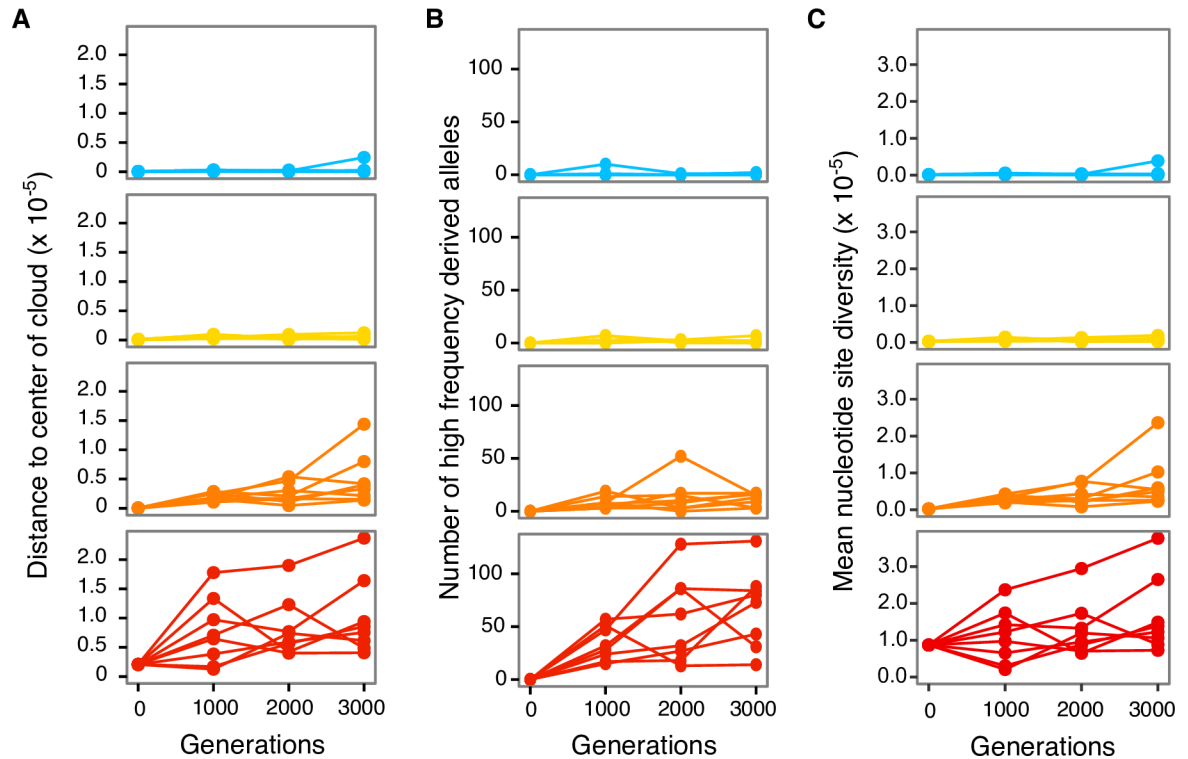


Fig 3. Replicate populations with higher mutation rates have increased genetic diversity and more high frequency derived alleles. Each circle represents (A) the distance of a population to its center (vertical axis), (B) the number of sites with derived alleles at a frequency exceeding 50% (vertical axis), or (C) the mean nucleotide site diversity (vertical axis) in an evolving replicate population over the course of the evolution experiment (horizontal axes). Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.

<https://doi.org/10.1371/journal.pgen.1007324.g003>

that higher mutation rates led to a larger cloud (Fig 3A; MR^S:0.005±0.001; MR^M:0.007±0.001; MR^L:0.011±0.001; MR^{XL}:1.018±0.001, cube root of diversity ± s.e.m.). Similarly, higher mutation rates led to higher levels of mean nucleotide site diversity (Fig 3C).

We also expected evolving replicate populations with higher mutation rates to accumulate more high frequency derived alleles than those with lower mutation rates because populations with more variation are expected to adapt faster. Here, we defined a high frequency derived allele at a given site as a derived allele found in more than 50% of the population. Again, strain identity affected the number of high frequency derived alleles (linear mixed effects analysis, $X^2(4) = 69$, $p = 4 \times 10^{-14}$; see Methods), such that strains with higher mutation rates accumulated more high frequency derived alleles (Fig 3B; MR^S:0.8±0.3; MR^M:1.8±0.8; MR^L:12±2; MR^{XL}:68±13, number of high frequency derived alleles ± s.e.m. at generation 3000).

Beneficial mutations

Most new mutations are thought to be effectively neutral or deleterious, and only a small fraction are beneficial in a given environment [1]. To identify putatively beneficial mutations in our replicate populations, we developed a statistical test that identifies genes in which more replicate populations contain high frequency derived alleles of any one gene than one would expect by chance alone (Methods, S9 Fig). In addition to identifying beneficial mutations, this approach can also identify artifacts such as mutational hotspots or the violation of independence across samples. Our test identified 20 genes with putatively beneficial mutations.

Statistical analysis places the highest confidence for the existence of beneficial mutations in eight of these genes (*clpB*: 6/32 populations, $p < 6 \times 10^{-6}$, *cspC*: 3/32 populations, $p = 3 \times 10^{-5}$, *mreB*: 5/32 populations, $p < 6 \times 10^{-6}$, *pykF*: 16/32 populations, $p < 6 \times 10^{-6}$, *rnb*: 7/32 populations, $p < 6 \times 10^{-6}$, *rpoC*: 6/32 populations, $p < 6 \times 10^{-6}$, *topA*: 9/32 populations, $p < 6 \times 10^{-6}$, and *ygeN*: 4/32 populations, $p = 4 \times 10^{-5}$ (S9A Fig, S10 Fig). Most of the mutations are nonsynonymous or nonsense mutations, and are thus likely to affect gene function (S9A Fig). Furthermore, between 70 and 100% of the observed mutations in any one gene occurred at different sites in different replicate populations (*clpB*: 100%, *cspC*: 100%, *mreB*: 91%, *pykF*: 95%, *rnb*: 78%, *rpoC*: 100%, *topA*: 70%, and *ygeN*: 100%), indicating that they occurred *de novo* and independently from each other. The two most commonly mutated genes were *pykF* and *topA*, which encode pyruvate kinase and topoisomerase A, respectively. Pyruvate kinase is a key enzyme in glycolysis, and topoisomerase A can affect the superhelicity of DNA. Both genes have repeatedly acquired beneficial mutations in previous experiments with *E. coli* B in glucose minimal medium [60–65]. Similarly, mutations in *cspC*, a stress protein, confer a fitness advantage for *E. coli* populations evolving at 37°C and higher [62,66]. Finally, mutations in the RNA polymerase gene *rpoC* and the cytoskeletal gene *mreB* have also been commonly found in laboratory evolution [62,67,68].

Surprisingly, multiple MR^{XL} replicates showed the same nucleotide change in 12 of the 20 putatively beneficial genes. As previously discussed, some mutations arose in the ancestor MR^{XL} population before we split it into its replicate populations (S7 Fig). Specifically, mutations in 10 of the 12 genes with the same nucleotide change across the MR^{XL} replicates were also found in the ancestral population (at a frequency between 3% and 22%), before we had split this population into our replicates (S9B Fig). Only two genes (*yfeZ* and *rrlH*) showed no evidence for such identical, pre-existing mutations, although such mutations may have existed below the detection limit of our sequencing coverage. We cannot conclude that these 12 putatively beneficial genes have a beneficial effect in the MR^{XL} populations, because our statistical test relies on the assumption that the mutations occurred and were subject to selection independently. To know with certainty the phenotypic effects of any of these mutations would require additional empirical data from allelic replacement experiments.

Growth and survival in stressful conditions

Thus far, the only phenotype we studied was population growth in one environment—the glucose minimal medium in which we conducted the entire experiment. To expand our analysis to other environments, we used Biolog Phenotype MicroArrays, which help measure the growth and respiration activity of a bacterial strain in multiple environments ([69], but see [70] for caveats). These microarrays determine the ability of our strains to grow in the presence of 96 stressful compounds that include antibiotics and heavy metals. We exposed our evolving replicate populations to these stressors only after completion of laboratory evolution, i.e., the populations could not have adapted to them during the evolution experiment. We selected two populations at random from the MR^S, MR^M, MR^L, and MR^{XL} replicate populations at the end of evolution. Remarkably, the two selected MR^{XL} replicates failed to grow in every single one of the 96 environments, as did the MR^{XL} ancestor. One possible explanation is that the MR^{XL} strain is inherently more sensitive to novel environments, including the medium used in the assay. The remaining populations grew in 42–60 (43.8%–62.5%) of the environments, depending on the population. In order to identify any link between mutation rate and growth in these 96 environments for the MR^S, MR^M, and MR^L replicates, we identified the molecules in which an evolved replicate population grew better (and worse) than its ancestor (S11 Fig). All replicate populations were better able to tolerate stressful conditions than their ancestors in some

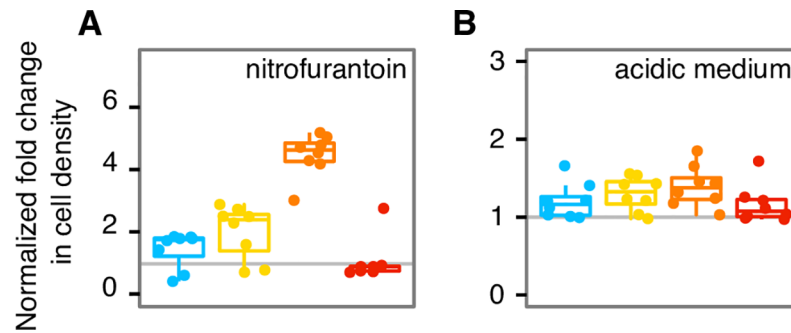


Fig 4. Cell density after 24 hours of growth in stressful conditions increased with increasing mutation rate, except for MR^{XL} replicate populations. We measured the cell density of the MR^S, MR^M, MR^L, and MR^{XL} evolved replicate populations relative to the ancestor (vertical axis) at generation 1000 in (A) medium supplemented with nitrofurantoin (2.2 µg/mL) and in (B) acidic medium (pH 5.25). The MR^S, MR^M, MR^L replicate populations performed better with increasing mutation rate, except for the MR^{XL} replicate populations which performed worst among all populations. See S12 Fig for data on additional timepoints, and nitrofurantoin and pH conditions. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.

<https://doi.org/10.1371/journal.pgen.1007324.g004>

of the tested conditions (between 8% to 30%), which suggests that some (fortuitously) beneficial mutations have occurred. The two MR^L replicate populations we tested tolerated stressful conditions better than the MR^S and MR^M replicate populations. However, the MR^S replicate populations were able to tolerate more stressful conditions than the MR^M populations. In sum, these analyses establish no simple association between ancestral mutation rate and stress tolerance after evolution.

To study the evolutionary dynamics of growth in stressful conditions over the course of the experiment, we periodically tested the growth of the ancestor and all evolving replicate populations in two stressful conditions: the antibiotic nitrofurantoin (a specific, “narrow” stressor) and acidic media (a broader stressor). Nitrofurantoin is a nitrofuran antibiotic with multiple mechanisms of action. Resistance to nitrofurantoin is conferred by mutations in two genes (*nfsA* and *nfsB*), and has a fitness cost in the absence of the antibiotic [71]. Thus, resistance mutations are unlikely to exist at appreciable levels as part of the standing variation in populations not exposed to nitrofurantoin. *E. coli* is known to tolerate acidic conditions due to several acid resistance systems, one of which depends on the alternative sigma factor σ^S , encoded by the *rpoS* gene [72,73]. For both nitrofurantoin-containing and acidic media, we computed the fold change in growth (the cell density after 24 hours) of the evolved populations relative to their ancestors, controlling for changes in carrying capacity (see Methods, Fig 4, S12 Fig). We found that replicate populations with the highest mutation rates grew more slowly in nitrofurantoin than their ancestor. In contrast, populations at low and intermediate mutation rates grew faster than the ancestor (see Methods, $p = 5 \times 10^{-8}$, linear mixed effects analysis, $X^2(4) = 40$; MR^S: 0.85 ± 0.15 ; MR^M: 0.49 ± 0.15 ; MR^L: 0.94 ± 0.15 ; MR^{XL}: -0.29 ± 0.15 , log fold change \pm s.e. m., positive log fold change indicates evolved strains grew better, and negative log fold indicates ancestor strains grew better). The ancestral mutation rate also affected growth in acidic media. In contrast to nitrofurantoin, all strains showed increased growth, but qualitatively similar to nitrofurantoin, the MR^{XL} replicates showed the smallest growth rate increase at low pH ($p = 2 \times 10^{-8}$, linear mixed effects, $X^2(4) = 42$); MR^S: 1.23 ± 0.09 ; MR^M: 1.39 ± 0.09 ; MR^L: 1.56 ± 0.09 ; MR^{XL}: 1.15 ± 0.09 , fold change \pm s.e. m., fold change > 1 indicates evolved strains grew better, and fold change < 1 indicates ancestor strains grew better). In sum, growth in two stressful conditions, nitrofurantoin-containing and acidic media, improved with increasing mutation rates (and thus increasing diversity), except for the MR^{XL} replicates which showed a relative reduction in growth.

Mutation rate decreased for the highest mutation rate replicates

We measured the mutation rate of one randomly selected clone from each evolved replicate population at generation 3000, and of the population’s ancestral strain. To this end, we used fluctuation assays for mutations that cause rifampicin resistance, and estimated the genomic mutation rate U using Drake’s approach [74]. The mutation rates of the evolved MR^{XL} replicates decreased on average by 556%, reaching 18% of the ancestor’s mutation rate (2%-42%, depending on the replicate); Fig 5, S3 Table). At the end of the evolution experiment, the mutation rates in the MR^{XL} populations were no longer statistically distinguishable from those of the MR^L replicates (Wilcoxon rank sum test, $p = 0.38$). The evolved populations’ mutation rates for the MR^M and MR^L strains also tended to decrease (MR^M: 49% of the ancestor’s mutation rate on average, ranging between 11%-90% of the ancestor; MR^L: 87%, range 32%-170%). In contrast, the replicates from the MR^S strain increased their mutation rate somewhat, to 206% of the ancestor’s mutation rate (range 16%-1000%). Having estimated the mutation rates for the ancestor and evolved populations, we also wanted to examine whether prominent theoretical models that predict declines in mean population fitness at high mutation rates apply to our populations (S2 Text). While some of the models we studied (e.g., that of mutational load) predict a small reduction in fitness at the highest mutation rates we employed, none of them could account for the magnitude of the loss of adaptation we found in several of the MR^{XL} replicates (S2 Text; Fig B, Fig C in S2 Text).

Because mutation rates changed between the beginning and the end of the experiment, we wondered whether the final mutation rates were correlated with our measured phenotypes. We found significant correlations between a replicate’s mutation rate and its effective population size, standing genetic diversity, and number of high frequency derived alleles, but no

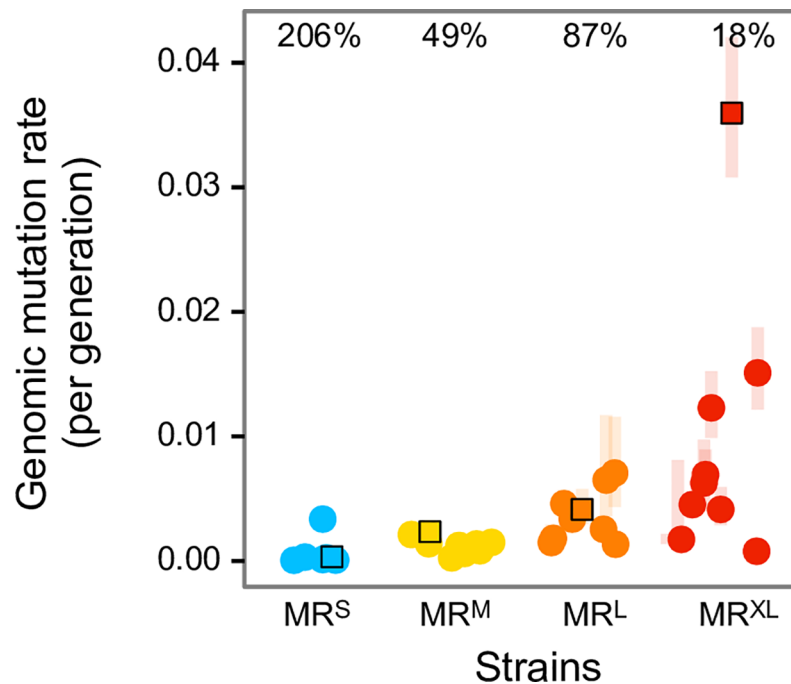


Fig 5. Evolution of the mutation rate. Ancestral genomic mutation rates are shown as squares, evolved mutation rates at generation 3000 as circles, and 95% confidence intervals of mutation rate estimates are shaded. Each evolved strain’s mean change in mutation rate is shown as the percentage of its ancestor’s mutation rate. Data from MR^S strains are shown in blue, from MR^M strains in yellow, from MR^L strains in orange, and from MR^{XL} strains in red.

<https://doi.org/10.1371/journal.pgen.1007324.g005>

correlations between a replicate's mutation rate and its final relative fitness, or normalized cell density after 24 hours of growth in acidic medium or medium containing nitrofurantoin (Spearman's rank correlation, S13 Fig). Interpretation of these results requires caution for two reasons. First, for any one population, we do not know exactly when (during the 3000 generations of evolution) the mutation rate changed from its ancestral value. Second, we compared the mutation rate of a single randomly-selected clone from populations which can have considerable genetic diversity, and thus potentially also show diversity in mutation rates. Despite these caveats, we found that the correlations between a representative clone's mutation rate and our other metrics are consistent with our previous analyses and figures (Fig 4, S2 Fig, S12 Fig), which simply considered the effects of ancestral mutation rate (strain identity).

Changes in the mutation rate genome

We call the set of genes potentially involved in modulating the mutation rate the "mutation rate genome".

We wondered whether this part of the whole genome was a preferential target of mutation or selection in our experiments. To find out, we first identified a set of 96 genes potentially involved in modulating the mutation rate (S2 Table) from the literature and EcoCyc [49,75–77]. If mutations or selection did not preferentially affect the mutation rate genome, the amount of genetic change we observe in it would be proportional to its length relative to the rest of the genome. This is indeed the case: We counted the number of synonymous mutations occurring at any frequency in any replicate population at generation 3000, and observed no statistically significant increase in the incidence of such genetic change in the mutation rate genome for any of our evolving strains (S14A Fig). We also found no difference in mean diversity between synonymous sites in the mutation rate genome relative to the rest of the genome (S14B Fig).

Although the mutation rate genome is not a preferential target of genetic change, its genes still accumulated many non-synonymous and nonsense changes, which are the kinds of changes that are especially likely to affect protein function (S15 Fig). To identify mutant alleles putatively associated with the decrease in mutation rates we had observed in MR^{XL} replicate population after 3000 generations (Fig 5), we identified nonsynonymous or nonsense mutations in the mutation rate genome with an allele frequency of at least 50% in any MR^{XL} evolved replicate population. Mutations in ten genes met these criteria (*rpoS*, *umuC*, *dinB*, *dinG*, *dps*, *glyS*, *glyW*, *mutL*, *phr*, and *vsr*), and two were found in multiple replicate populations (*rpoS*: 7 of 8; *umuC*: 2 of 8). The *rpoS* gene encodes the alternative sigma factor, σ^S , which activates the stress response in *E. coli* (reviewed in [78]). Populations with *rpoS* mutations can hold a fitness advantage in nutrient-limiting environments [79], but at a cost to fitness in a variety of stressful environments [28,80]. Because σ^S is a bacterial transcription factor, it can only affect the mutation rate indirectly, by changing the expression of proteins directly involved in DNA copying, repair, and proofreading. For example, σ^S modulates the expression of the error prone DNA Polymerase IV encoded by *dinB* [81]. We found the same *rpoS*N124D mutation in 2.5% of the individuals in the ancestral MR^{XL} population and in all eight evolved replicates. (This mutation reached 40.1% in MR^{XL}₁ and 100% in the rest of the replicates.) Thus, the mutation was likely distributed to the eight replicate populations from the ancestor, and either increased in frequency due to its direct fitness effects, or because it was hitchhiking with a beneficial mutation. The MR^{XL}₄ and MR^{XL}₆ replicate populations each acquired different non-synonymous mutations in *umuC*, which encodes DNA polymerase V. Each of the remaining genes with high frequency mutant alleles in a single replicate population were involved in DNA repair and replication (*dinB*, *dinG*, *glyS*, *glyW*, *mutL*, *phr*, *vsr*) or protection of DNA in

stationary phase (*dps*) in a single replicate population and could have also affected the evolved mutation rate.

Discussion

Here, we studied the effects of mutational pressure on evolutionary adaptation and the evolution of the mutation rate itself. To this end, we engineered four isogenic *E. coli* K12 MG1655 derivative strains with increasing mutation rates (MR^S , MR^M , MR^L , and MR^{XL}), and evolved them for 3000 generations. Our smallest (wild-type) ancestral mutation rate (MR^S : $U = 0.00034$ per genome per generation) was somewhat smaller than rates estimated in *E. coli* strains using a similar experimental approach ($U = 0.0025$) [39,74], but similar to those estimated in a wildtype *E. coli* B strain using a sequencing approach ($U = 0.00041$) [82]. At the opposite extreme was our strain with the highest ancestral mutation rate (MR^{XL}). We originally expected this strain to have a mutation rate approximately 4500-fold higher than wildtype [35], consistent with the large effects that mutations in the *dnaQ* and *mutL* genes have on the mutation rate [75]. However, our measurements of this rate demonstrated that it was lower than expected (MR^{XL} : 139-fold higher than our wildtype; $U = 0.036$ per genome per generation). The discrepancy could in principle be due to the acquisition of an anti-mutator allele during the transfer of the strain between laboratory locations. Alternatively, our mutation rate could be an underestimate for technical reasons discussed in the Methods. The mutation rate for our MR^{XL} strain was also somewhat lower than that of a hypermutable clone which spontaneously evolved from an *E. coli* B strain [50] (*mutT*: $U = 0.061$). The mutation rate of our hypermutable MR^{XL} strain is low enough that we expected its populations to be viable [21]. In sum, we conducted our experiments with strains having a range of viable mutation rates, from wildtype (MR^S), to a 16-, 22-, and 139-fold higher mutation rate (MR^M , MR^L , and MR^{XL}).

We first characterized the general patterns of adaptation in our four strains, and found that their fitness increased significantly by generation 3000 for all replicate populations. Previous experimental evolution studies in constant environments have observed fitness gains that are initially large but decrease over time [17,18,83,84], which is consistent with diminishing returns epistasis, in which the size of the fitness gain in an evolving population depends on its current fitness, such that populations with lower fitness can improve their fitness to a greater extent [85,86]. However, our fitness trajectories differ from those predicted by diminishing returns epistasis in two ways. First, they do not show a decreasing fitness gain over time [18]. Second, the mean fitness of replicate populations with small or modestly high mutation rates (MR^S , MR^M) did not immediately improve, but unexpectedly remained largely unchanged for the first 1000 generations (compared to [87]). While delayed adaptive response is consistent with a lower overall beneficial mutation supply rate, it may not be sufficient to explain our observations. We expected to wait just 44 generations for a new beneficial mutation to establish in our slowest-evolving replicate population (S3 Text). It also raises the possibility that even at moderately high mutation rates, contingent evolution [88], in which the timing and the order of mutational events affects a population's adaptive evolution, may be important in our populations. An instance of such contingent evolution has been documented in *E. coli* [89,90], but the higher mutation rate of some of the strains used in our evolution experiment makes contingent evolution a less likely explanation for delayed adaptation.

We next characterized the effect of mutational pressure on adaptation. We found that strains with higher ancestral mutation rates increased in fitness more than those with lower mutation rates, except for MR^{XL} populations, which we will discuss below. These observations are in agreement with theory [15,91] and previous experimental studies which found that large asexual populations of *E. coli* [17,50] and yeast [15] with high mutation rates outperformed

those with low mutation rates. If we just consider relative fitness after 1000 generations, our data from our four strains are consistent with expectations: MR^S and MR^M populations have lower mean relative fitness than MR^L and MR^{XL}. It is only thereafter that the fitness of MR^{XL} populations stops increasing, such that they have lower mean fitness at generation 3000 than the MR^S, MR^M, and MR^L strains. We do not actually observe the loss of fitness on average across the MR^{XL} replicate populations, but rather a prolonged period in which fitness remains unchanged as a whole. Interestingly, however, the fitness of several MR^{XL} replicate populations decreases from its maximum and arrives at a value that is approximately equal to that of the ancestral population. This is reminiscent of models of extreme mutational pressure developed over the past forty years that predict reduced adaptation and eventual extinction [19,20,22,92–94]. However, these models predict a loss of fitness only at higher mutation rates than we observed, and require unrealistic assumptions (S2 Text), together emphasizing the importance of additional theoretical work. Another possibility is Hill-Robertson interference [7], which can reduce the rate of adaptive evolution by background selection—negative selection against deleterious alleles that removes the most deleterious lineages from a population—and can reduce genetic diversity [8,12]. Empirical evidence supports the action of this mechanism in natural populations of several eukaryotic species (reviewed in [13,14]). However, because background selection removes deleterious mutations from a population, it cannot alone reduce the fitness of a population and it can therefore not explain the loss of fitness we observed in the three MR^{XL} replicates. Overall, our observations support the notion that reduced adaptation can manifest itself at smaller mutation rates than previously thought ($U = 0.036$ in the MR^{XL} strain), even though more than 1000 generations may be needed to manifest its effects.

This observation is all the more striking, because the mutation rate itself had decreased dramatically for all MR^{XL} populations after 3000 generations (and much less so in the MR^S, MR^M, MR^L populations). While a lowering of the mutation rate has been previously observed [46–48,50] and predicted to be favored in some conditions [38,40,42,43,95], its extent and consistency across multiple of our evolving populations is remarkable. The mutation rate decrease probably did not occur very early during evolution, because the MR^{XL} populations show greater genetic diversity than all other populations throughout the experiment (Fig 3). The decreasing mutation rate, together with the observation that the MR^{XL} populations failed to adapt after more than 1000 generations, suggests that the maladaptive effects of hypermutation begin at even lower mutation rates than those in our initial MR^{XL} strain. While we cannot predict whether our hypermutable populations would eventually go extinct, the observation that their mutation rate can decrease adaptively makes this less likely. Indeed, recent mutation accumulation experiments with small bacterial populations suggested that populations with higher mutation rates tend to go extinct more often and have reduced fitness than populations with lower mutation rates [47]. Of the several "mutation rate genome" genes mutated in MR^{XL} strains, only *rpoS* was found in all eight evolved MR^{XL} replicate populations. *rpoS* encodes for the stress response modulator σ^S that can indirectly affect the mutation rate through transcriptional changes. However, we cannot definitively identify the proximal mechanisms driving the drop in mutation rates using bioinformatics alone. Future experimental studies to evaluate the effect of each "mutate rate genome" mutant allele on the mutation rate and fitness would be necessary.

We emphasize that all our experiments use asexual populations, and that the evolutionary dynamics of mutation rates and adaptation may be different in sexual, recombining populations. For example, in our non-recombining populations, any mutator allele remains completely linked to the (mostly deleterious) mutations it helps bring forth, resulting in indirect negative selection on the mutator allele. However, such an allele and its associated mutations can become unlinked in recombining populations, which reduces the strength of indirect

selection on the mutator allele (see [33,39] for reviews). Additionally, beneficial and deleterious alleles can become unlinked in recombining populations, which can lead to increased levels of adaptation and diversity (see [13,14] for reviews).

We also characterized the effect of mutational pressure on the ability of an evolving population to grow better (or worse) than its ancestor in a variety of chemically novel environments, which contain chemical agents that include heavy metal stressors, antibiotics, or acids. Importantly, our populations were never exposed to any of these conditions during the evolution experiment. *A priori*, we reasoned that two outcomes were possible. First, populations with high mutation rates may grow better in novel environments, because they can accumulate more beneficial mutations while evolving in their original environment, and these mutations may also be beneficial in novel environments through pleiotropy. High mutation rate populations can also generate more genotypic diversity, which in turn increases the chances that a population harbors a clone with a latent beneficial mutation that allows it to grow better in a novel environment. Such latent beneficial mutations can indeed occur, as demonstrated by the classic fluctuation test, which relies on such mutations to estimate mutation rates towards resistance to lethal selection [96,97]. Second, populations with high mutation rates may grow worse in novel environments, because they may accumulate more mutations that are either beneficial or neutral in the current environment, but deleterious in a novel environment. Such latent deleterious mutations do indeed exist [36,70,98]. In sum, strains with high mutational pressure may harbor more latent beneficial alleles, but also more latent deleterious alleles, and it is not clear *a priori* which dominates in their effect on fitness.

We conducted two tests on how mutational pressure can affect growth in novel conditions. In the first, we measured the growth of eight evolved replicate populations (two each from MR^S , MR^M , MR^L , and MR^{XL}) in 96 chemically novel environments. This test did not yield a clear association between mutation rate and growth for our MR^S , MR^M , and MR^L populations. However, it yielded a very clear pattern for our MR^{XL} populations: They were not able to grow in any one of these environments, which illustrates that at the highest mutation rates we consider, latent deleterious mutations outweigh beneficial ones in both the ancestor and evolved populations. One possible explanation is that the MR^{XL} strain is inherently more sensitive to novel environments, including the assay environment. Because the MR^{XL} ancestor population could not grow at all, we were unable to further quantify the effect of the highest mutation rate in these 96 novel environments. In the second test, we periodically measured growth of all 32 replicate populations (relative to their ancestors) in two stressful conditions: the antibiotic nitrofurantoin (a specific narrow stressor) and an acidic medium (a broader stressor). For both, we found that strains with higher ancestral mutation rates could grow better than those with lower mutation rates, except for MR^{XL} replicate populations, which grew worst of all populations. This experiment shows that latent beneficial alleles may predominate at low and intermediate mutational pressure, but no longer at high mutational pressure. Our observations are consistent with a previous study showing that multidrug resistance in *E. coli* is favored by intermediate mutation rates [99].

In sum, a modest increase in mutation rates can provide an evolutionary advantage in both the constant environment of our long-term laboratory evolution experiment and in novel environments. This advantage disappears at the highest mutation rates ($U = 0.036$) we considered, where populations show signs of decaying adaptation and poor performance in novel environments. These mutation rates are below those commonly considered to limit adaptation, and highlight the need for additional theoretical work. Our observations show that biological systems may be more sensitive to mutational pressure than simple theoretical models suggest, at least when the effects of mutations are allowed to accumulate over many generations. This observation may improve the prospects of using elevated mutagenesis to drive

Table 1. Strains and plasmids used in this work.

Strain	Description	Details	Source
MR ^S	<i>mutL</i> , <i>dnaQ</i> wildtype	<i>F</i> -, <i>fhuA2</i> , <i>lacY1</i> , <i>tsx-1</i> or <i>tsx-70</i> , <i>glnV44</i> (AS), <i>gal-6</i> , λ -, <i>xyl-7</i> , <i>mtlA2</i> , <i>cycA30</i> :: <i>Tn10</i>	This experiment
MR ^M	Impaired mismatch repair (<i>mutL</i>)	<i>F</i> -, <i>fhuA2</i> , <i>lacY1</i> , <i>tsx-1</i> or <i>tsx-70</i> , <i>glnV44</i> (AS), <i>gal-6</i> , λ -, <i>xyl-7</i> , <i>mtlA2</i> , <i>mutL13</i> (Yale's Coli Genetic Stock # 4816)	[35]
MR ^L	<i>mutL</i> , <i>dnaQ</i> wildtype	<i>F</i> -, <i>fhuA2</i> , <i>lacY1</i> , <i>tsx-1</i> or <i>tsx-70</i> , <i>glnV44</i> (AS), <i>gal-6</i> , λ -, <i>xyl-7</i> , <i>mtlA2</i>	This experiment
MR ^{XL}	Impaired mismatch repair (<i>mutL</i>), impaired proofreading (<i>dnaQ</i>)	<i>F</i> -, <i>fhuA2</i> , <i>lacY1</i> , <i>tsx-1</i> or <i>tsx-70</i> , <i>glnV44</i> (AS), <i>gal-6</i> , λ -, <i>xyl-7</i> , <i>mtlA2</i> , <i>mutL13</i> , <i>yafC502</i> :: <i>Tn10</i> , <i>dnaQ905</i>	[35]
ES4	Source for wildtype mismatch repair gene (<i>mutL</i>) in MR ^L	<i>F</i> -, <i>fhuA2</i> , <i>lacY1</i> , <i>tsx-1</i> or <i>tsx-70</i> , <i>glnV44</i> (AS), <i>gal-6</i> , λ -, <i>xyl-7</i> , <i>mtlA2</i> , <i>purA45</i>	[109]
<i>E. coli</i> K12 MG1655	Previously sequenced as the <i>E. coli</i> K12 reference genome	<i>F</i> -, <i>LAM</i> -, <i>rph-1</i> (Yale's Coli Genetic Stock # 7740)	[110]
CAG12073	Wildtype mismatch repair (<i>mutL</i>) linked to tetracycline resistance	λ -, <i>rph-1</i> , <i>cycA30</i> :: <i>Tn10</i>	[107]
pKD4	Plasmid containing <i>bla</i> , <i>FRT</i> , and <i>kan</i>		[105]
pKD46	Plasmid containing <i>bla</i> and <i>araBp-gam-bet-exo</i>		[105]
pCP20	Plasmid containing <i>FLP</i> and <i>cat</i>		[106]

<https://doi.org/10.1371/journal.pgen.1007324.t001>

pathogen or tumor populations to extinction [20,100–104], if high mutation rates can be sustained for a sufficiently long amount of time.

Methods

Bacterial strains

We utilized four isogenic *E. coli* strains derived from K12 MG1655 that have increasing mutation rates. We refer to these strains as the MR^S, MR^M, MR^L, and MR^{XL} strains, corresponding to small (S), medium (M), large (L), and extra-large (XL) mutation rates. Strain genotypes are summarized in Table 1. We obtained MR^M and MR^{XL} strains from our previous experiments (therein called the single- and double-mutator, respectively) [35]. MR^M has a non-synonymous (A120T) mutation in the *mutL* gene relative to the *E. coli* wild type. This gene is involved in the methyl-directed mismatch repair system. We previously constructed the MR^{XL} strain by P1 transduction of the *dnaQ* gene from the *E. coli* CSH116 strain, which has a non-synonymous mutation, T15I, in the *dnaQ* gene, into the MR^M strain [35]. The *dnaQ* gene encodes the epsilon subunit of DNA Polymerase III; mutations in this gene can disrupt proofreading. We constructed the mutator strain MR^L by replacing the *mutL* region in MR^M with the *mutL* region from ES4 with a kanamycin resistance gene inserted upstream of the region, using the method of Datsenko and Wanner [105]. We then excised the kanamycin resistance gene using pCP20 [106], which left a small scar immediately upstream of the *mutL* gene. We constructed the low mutation rate strain MR^S by using P1 transduction to replace the error-prone *mutL* region in MR^M with the wildtype allele from CAG12073 [107]. We confirmed the mutation rates of these ancestral strains using fluctuation tests [108] (see "Mutation rate measurements and calculations" for details), and found that the MR^M, MR^L, and MR^{XL} strains had 16-, 22-, and 139-fold higher mutation rates to rifampicin resistance than MR^S. We also confirmed our manipulations of the MR^S, MR^M, MR^L, and MR^{XL} strains by sequencing their genomes (S1 Table). In addition to these strains, we used the strain *E. coli* K12 MG1655, which we obtained from Yale's Coli Genetic Stock Center.

Evolution experiment

See Fig 1 for an overview. We evolved eight independent replicates from populations starting from a single clone for each of the MR^S, MR^M, MR^L, and MR^{XL} strains for 175 days (2907 generations) in 48-well plates (Fluka 15758-500G-F) in 2 mL of Davis Minimal broth [111] supplemented with 1000 mg/L glucose ('DM1000') at 37°C with shaking at 400 rpm in a microtiter plate shaker (Stuart Microtiter 51505).

Each plate held 24 populations arranged in a checkerboard pattern, such that each well was surrounded only by wells with blank medium, and the populations were assigned to the 24 wells at random by a custom R script. We diluted each culture 100,000-fold every 24 hours into fresh DM1000 medium, which allows almost 17 generations of growth per day. Every 7 days, we archived each evolving population by adding 400 µL of 50% glycerol to 800 µL of stationary phase culture and freezing at -76°C, estimated the cell density by plating subsamples onto LB (Difco 244620) plates with 1.5% agar (Sigma A1296-1KG), and froze cell pellets from 800 µL of stationary phase culture for eventual genome sequencing. We delayed the start of the MR^S replicates by 63 days for technical reasons.

We controlled for contamination in several ways. First, if we observed growth in an empty well, we repeated the transfer from the previous day's 48 well plate stored at 4°C. Second, we periodically checked each evolving culture for contamination by confirming its resistance profile and approximate mutation rate using spot tests. In short, we spotted 25 µL from each evolving culture onto tetracycline (10 µg / mL) and rifampicin (100 µg / mL) plates, and incubated overnight at 37°C. MR^S and MR^{XL} replicates can grow on tetracycline, and the replicates with higher mutation rates display more colonies on rifampicin. Cross-contamination occurred once, which prompted us to restart the experiment from the most recent set of uncontaminated glycerol stocks (day 98 for MR^M, MR^L, and MR^{XL} replicates; day 35 for MR^S replicates). Third, we examined the genome sequence data for cross-contamination, but detected no evidence for cross-contamination in it.

Effective population size

For populations that do not have a constant number of cells, the effective population size is given by the harmonic mean of population sizes over the course of the dilution and growth cycles of the experiment. Previous studies have estimated the effective population size only from the size of the bottleneck measured during one dilution [112,113]. In contrast, because we recorded the census size of the population at carrying capacity (N_{max}) every 7 days, we were able to estimate the effective population size as the harmonic mean of the population sizes both at the beginning and at the end of a cycle of growth and dilution. To obtain $N_{max,d}$ at any one day d , we counted the number of cells in each evolving replicate population in stationary phase just before transferring the population into fresh media. We did so by plating serial dilutions in duplicate on LB agar plates and incubating overnight at 37°C. We discarded plates with fewer than 20 or more than 700 colonies for the purpose of this analysis. Because at the end of each growth cycle we diluted our cultures 100,000-fold into fresh medium, a total of $G = \log_{2z}(10^5) = 16.61$ cell generations (floor($\log_2(10^5)$) \approx 16 complete cell generations) elapsed during each growth cycle, and the minimum population size was $N_{min,d} = N_{max,d} \times 10^{-5}$. Thus, during each generation g of each growth cycle, a population assumed population sizes

$$N_{d,g} = \begin{cases} 2^g N_{min,d}, & \text{if } 0 \leq g \leq 16 \\ N_{max,d}, & \text{if } g = 17 \end{cases}.$$

(Because the precise number of generations in each dilution cycle is $\log_2(10^5) = 16.61$, we included the final number of cells $N_{d,17} = N_{max,d}$ in this calculation). We then determined the nominal effective population size (N_e) of a replicate population during its entire lab evolution as

$$N_e = 25 \times 18 \left/ \sum_{d=1}^{25} \sum_{g=0}^{18} \frac{1}{N_{g,d}} \right.$$

which is the harmonic mean of all the population sizes. We calculated it for all 25 days on which we collected population size data. The number 18 corresponds to the total number of generations g for which we computed population sizes during any one of these 25 days. We also estimated the effect of linkage on reducing the effective population size due to background selection or interference selection [53–55,114]. To this end, we used the R functions (GordoNe and GoodNe, respectively, available from [53]), where we take the size of the deleterious selective effect as $s = 0.03$ [115] and use U as an upper bound on U_d to obtain rough estimates of linkage's effect on effective population size.

Fitness measurements

We periodically obtained a proxy for the fitness of the evolving strains by measuring growth curves of the archived populations. For each time point, we restarted all evolving populations as well as three replicates from each ancestral population and three replicates of wild type *E. coli* K12 MG1655 from glycerol stocks in 2 mL of DM1000 and incubated them overnight at 37°C with shaking. We then diluted the overnight cultures 50-fold into 200 μ L final volume of DM1000 in 96-well plates (TPP 92096), and incubated them in a plate reader (Tecan Infinite Pro F200) for 18 hours at 37°C with shaking. During this time, we read the absorbance at 600 nm every 10 minutes. We fit the classic logistic equation describing population growth to the data [116], using the Growthcurver R package [117], and defined the relative fitness of each population as $r_{evo} - r_{anc}$. Here, r_{evo} is the growth rate of the evolved population and r_{anc} is the mean growth rate of the three replicates of the ancestor grown in the same plate reported in units of cell divisions per hour. We measured each growth curve three times. We used the R package lme4 v1.1–12 [118] to perform a linear mixed effects analysis of the relationship between the evolved fitness relative to the ancestor and the mutation rate class. In this analysis, we chose the mutation rate classes as fixed effects, and the replicate population as well as the 96-well plate as random effects. For all linear mixed effects analyses conducted in this paper, we observed no deviations from homoscedasticity according to Levene's test for homogeneity of variance [119] implemented in the R package car v2.1–2 [120]. Also, all residuals were normally distributed unless otherwise specified. We obtained significance values using a likelihood ratio test of the full model against a null model that did not contain the fixed effects. Using the data from the above growth curve experiments, we also compared the fitness of the ancestor populations against each other by obtaining the relative fitness of the ancestors as $r_{anc} - r_{K12}$, where r_{anc} is the growth rate of the ancestor population and r_{K12} is the mean growth rate of the three replicates of *E. coli* MG1655 K12 grown in the same plate. We performed a linear mixed effects analysis of the relationship between the ancestral fitness relative to *E. coli* K12 and the mutation rate class using the lme4 package, as just described. In this analysis, we chose the mutation rate classes as fixed effects, and the identity of the original glycerol stock and of the 96-well plate as random effects. We used the R package multcomp v1.4–6 [121] to test whether fitness values had changed from the ancestral or reference state of 0.

Sequencing

We sequenced samples from the four ancestral populations (day 0, generation 7) and from each of the 32 evolving replicate populations at days 63, 119, and 175 (generations 1046, 1977, and 2907). In total, we thus sequenced 100 populations ($4 + 3 \times 32$). For simplicity, we hereafter designate these time points as generations 0, 1000, 2000, and 3000. For each, we isolated the DNA directly from cell pellets obtained from the evolving populations using Qiagen's DNeasy Blood and Tissue kit (cat. No 69582), with modifications as previously described [122]. We used the TruSeq DNA PCR-Free kit (Illumina FC-121-3002) to prepare and barcode the libraries for paired-end sequencing, as previously described [122]. Importantly, we used no PCR steps in preparing the libraries. We employed qPCR with Roche's FastStart Essential DNA Green Master kit (Cat no. 06402712001) to quantify the libraries, which were then mixed in equimolar amounts for sequencing. We sequenced the populations (paired-end, 125 bp) on a single lane of Illumina's HiSeq 2500 v2. We used *breseq* v0.26.1 [56] to align the reads, and call and annotate the variants relative to the *E. coli* K12 MG1655 reference genome NC_000913.3 [110], downloaded from the NCBI (<https://www.ncbi.nlm.nih.gov/nucleotide/556503834>) on January 6, 2015. We developed scripts in R to identify the alterations that occurred in the evolved populations, but were not fixed in their ancestors. We determined mutational spectra by identifying all mutations that occurred at any detectable frequency in each population at every sequenced time point, and classified them into the following categories: A→C, A→G, A→T, C→A, C→G, or C→T. (Because DNA is double-stranded, the remaining possible point mutations are covered by their reverse complements, e.g., T→G corresponds to A→C.) We computed the relative frequencies of each mutational class for each replicate population, and used these to perform a principal component analysis (PCA) in R with *prcomp*, which uses singular value decomposition for the PCA. All data are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.mh206>.

Diversity and mutational cloud metrics

To quantify the movement and spread of a population as a “cloud” of sequences in sequence space, we first defined the center of this mutational cloud at any given site n in the genome as the majority allele $e_n \in \{A, C, G, T\}$, i.e., the allele whose frequency p exceeded 0.5 (all sites we analyzed had one such allele). We defined the center of the mutational cloud of genomes as the location in genotype space defined by the majority allele at each site. It can also be viewed as the location of the population's consensus sequence. If we denote the fraction of a population with the majority allele at site n as p , then the distance of the population to a given majority allele at site n can be thought of as the fraction of the population not having the majority allele, which is given as $C_n = (1-p)$. We define the population spread metric C as the average of C_n over all sites in the genome. A related quantity is the approximate sequence distance D that an evolving population has moved from its ancestral genotype, i.e.,

$$D = \sum_n \begin{cases} 0, & \text{if } a_n = e_n \\ 1, & \text{otherwise} \end{cases}$$

where $a_n \in \{A, C, G, T\}$ is the ancestral allele of the population at generation 0. In other words, D corresponds to the total number of sites at which the majority allele is different from the ancestral allele. We also computed each population's average genome-scale nucleotide site diversity [123,124] using the pairwise alignment position nucleotide counting approach [125,126]. We

estimated the proportion of pairwise nucleotide differences at each site n as

$$p_n = \frac{m_p(1 - m_p)}{m(m - 1)/2},$$

where m_p is the number of reads corresponding to the majority allele and m is the total number of reads at site n . We estimated the average nucleotide diversity for the L positions in our genome having non-zero coverage as

$$\pi = \sum_{n=1}^L \frac{p_n}{L}.$$

We used the R package lme4 v1.1–12 [118] to perform a linear mixed effects analysis of the relationship between the cube root of C or D (taken to ensure homoscedasticity) and the mutation rate class. In this analysis, we chose the mutation rate classes as fixed effects, and the time points and each of the 32 evolving replicates as random effects. We obtained significance values using a likelihood ratio test of the full model against a null model that did not contain the fixed effects.

Identification of putatively beneficial mutations

We identified putatively beneficial mutations as mutations that occurred in a genomic region more often than one would expect by chance alone. To identify such mutations, we used a numerical approach that focuses on a given gene g among a larger set of genes or genomic regions G (e.g., a gene among the set of all genes), and asked whether more replicate populations experienced a high-frequency genetic change than expected by chance. To this end, we first counted the number n_g of replicate populations with a mutation in gene g that had reached a frequency greater than 50% at generation 3000.

If all sites in the genomes of all samples were equally likely to experience a mutation, and if different genes were likely to experience mutations only in proportion to their length, then the probability p_g that any one gene g receives such a mutation in any given replicate would depend only on the length of the gene l_g ,

$$p_g = \frac{l_g}{\sum_{\gamma \in G} l_\gamma}.$$

The n_g total mutations found in gene g could be distributed in $\binom{32}{n_g}$ ways across the 32 replicate populations $R = \{MR^S_1, \dots, MR^S_{32}, MR^M_1, \dots, MR^M_{32}, MR^L_1, \dots, MR^L_{32}, MR^{XL}_1, \dots, MR^{XL}_{32}\}$. For example, for $n_g = 2$, the mutations could be distributed across the 32 replicate populations in $\binom{32}{2} = 496$ ways, i.e., one could occur in MR^S_1 and the other in MR^S_2 , one could occur in MR^S_1 and the other in MR^S_3 , etc. We computed the probability of observing the n_g mutations in any given set of replicates as the probability that gene g was mutated in each member of the set of replicates times the probability that it was not mutated in any of the other replicates. For $n_g = 2$ and replicate populations $r_i, r_j \in R$, this quantity is given by the binomial distribution adjusted to account for the number of observed mutant genes in the replicate populations, n_{r_i} and n_{r_j}

$$p_{r_i, r_j} = n_{r_i} n_{r_j} p_g^{n_g} (1 - p_g)^{(n_{tot} - n_{r_i} - n_{r_j})},$$

where

$$n_{tot} = \sum_{r \in R} n_r.$$

The probability of observing exactly $n_g = 2$ mutations in gene g in any pair of replicate populations is the sum of the probabilities that n_g mutations occurred in each of the 496 pairs, and is given by

$$P_g = \sum_{r_i, r_j \in R, r_i \neq r_j} p_{r_i, r_j}.$$

This quantity P_g is our null expectation that two replicates acquire mutations in gene g , if each replicate population's mutations were randomly distributed across its genome. We were interested in genes containing mutations in improbably many replicate populations, which we identified as those genes having less than a 0.005 percent chance of finding n_g replicate populations with a high frequency derived allele in the gene. We performed analogous analyses for $n_g > 2$. For example, for genes in which we observed three replicates with mutations ($n_g = 3$), we computed the probability that three replicates r_i , r_j , and r_k each contained a mutation in gene g as

$$p_{r_i, r_j, r_k} = n_{r_i} n_{r_j} n_{r_k} p_g^{n_g} (1 - p_g)^{(n_{tot} - n_{r_i} - n_{r_j} - n_{r_k})}.$$

Similarly, the probability of observing a mutation in exactly 3 replicate populations is given by

$$P_g = \sum_{r_i, r_j, r_k \in R, r_i \neq r_j \neq r_k} p_{r_i, r_j, r_k}.$$

Mutation rate measurements and calculations

We estimated the mutation rate of a single clone isolated from each ancestor and from each evolved replicate population through fluctuation assays that screened for mutants resistant to rifampicin [127], which can be caused by mutations in the *rpoB* gene. Specifically, we performed the following procedure for each replicate population. We isolated a single random clone and incubated it overnight in 2 mL DM1000 in 48 well plates at 37°C with shaking. We diluted the resulting overnight culture 100,000-fold to yield a culture with approximately 1000 cells in 100 μL. We then transferred 100 μL of the diluted culture into 5–7 sterile 50 mL tubes (Sarstedt 62.547.254) containing 30 mL of DM1000, and incubated for 48 hours at 37°C with shaking. We estimated the number of cells in each tube by plating dilutions on LB agar plates, and estimated the number of resistant cells in each tube by plating dilutions on LB agar plates supplemented with 50 mg/mL rifampicin (Sigma R3501-5G). We calculated the mutation rate to rifampicin μ_{rif} using the method and program provided by Philip Gerrish [108]. We obtained the genomic mutation rate U using Drake's approach [74] by first determining the "correction factor" C , which counts the number of single nucleotide mutations in *rpoB* that show rifampicin resistance. By counting all possible nucleotide changes underlying the amino acid changes in *rpoB* previously shown to confer rifampicin resistance [57], we determined that $C = 71$. Finally, we estimated the genomic mutation rate as $U = L\mu_{rif}/C$, and the mutation rate per base pair as $\mu = \mu_{rif}/C$, where $L = 4641652$ is the number of nucleotides in the *E. coli* K12 genome. This mutation rate may be an underestimate because we neglected other types of mutations (e.g., indels) and mutations in other genes that may lead to rifampicin resistance.

Phenotype screening

Our phenotype screening revolved around the density of cells after growth in various chemicals. Specifically, we determined the cell density after 24 hours of growth using Biolog

Phenotype MicroArrays PM11C, PM12B, PM13B, and PM14A MicroPlates (Biolog, Inc., Hayward CA, USA), which assay the sensitivity of bacteria to diverse chemicals that range from antibiotics to heavy metals. We screened the ancestors at generation 0, and two randomly selected replicates of the evolved populations (MR^S_1 , MR^S_6 , MR^M_3 , MR^M_4 , MR^L_2 , MR^L_7 , MR^{XL}_3 , MR^{XL}_4) at the final time point of laboratory evolution. To do so, we streaked population samples from glycerol stocks onto LB agar plates, incubated them at 37°C for 24 hours, restreaked the resulting colonies onto fresh LB agar plates (37°C, 24 hours), and repeated this streaking and incubation procedure once more. We resuspended the colonies from the final (third round) plates in IF-0 solution (Biolog, Inc., Hayward CA, USA) to a final absorbance reading at 600 nm of approximately 0.18 (200 μ L suspension in a 96-well TPP plate, ref 92096). We diluted this suspension 6-fold using IF-0+dye (Biolog, Inc., Hayward CA, USA), and diluted the resulting suspension 201-fold using IF-10+dye (Biolog, Inc., Hayward CA, USA). We added 100 μ L of the final solution to each well of a Phenotype MicroArray and incubated the array in the dark at 37°C for 24 hours, taking absorbance readings at 600 nm after 10 minutes and 24 hours. We then computed the Biolog phenotype $B_{S,C} = A_{S,C,24h} - A_{S,C,10m}$ where $A_{S,C,24h}$ and $A_{S,C,10m}$ are the absorbance readings for each sample S and compound C at 600 nm after 24 hours and 10 minutes.

To determine the minimum threshold for detection of growth in a given compound C, we computed the absolute difference between the readings in a given well across all pairs of samples (i,j) after 10 minutes (before cells had started to grow and divide), i.e., $A_{i-j,C,10m} = |A_{i,C,10m} - A_{j,C,10m}|$. The values of $A_{i-j,C,10m}$ quantify the expected experimental noise of wells with no growth. We found that $A_{i-j,C,10m} < 0.097$ for more than 99% of sample pairs. Based on this observation, we considered differences between readings smaller than the threshold value $A_{thresh} = 0.097$ as due to experimental noise.

Each compound in the Biolog Phenotype MicroArrays we used occurs in four wells at increasing concentrations. For further analysis, we used data only from the concentration (the well) that showed the highest variation in the difference between matched evolved and ancestor strains across all samples. We considered a sample to have evolved tolerance to a compound C if it improved its phenotype after 3000 generations of evolution more than expected based on experimental noise, i.e., $B_{SEvo,C} - B_{SAnc,C} > A_{thresh}$. Likewise, we considered that a sample had lost tolerance if its phenotype had degenerated after 3000 generations of evolution, i.e., if $B_{SAnc,C} - B_{SEvo,C} > A_{thresh}$. We note that both cellular growth and respiration contribute to the Biolog phenotype $B_{S,C}$, because respiration can occur independently of cellular growth [70,128].

We were also interested in observing the evolutionary dynamics of phenotypes over time. The phenotypes we selected for this analysis are the cell density after 24 hours of growth of the evolved populations relative to their ancestors in two conditions: a narrow antibiotic (nitrofurantoin) stress, and a broader environmental (low pH) stress. Specifically, we chose DM1000 medium with 1.5–2.4 μ g/mL nitrofurantoin for the narrow antibiotic stress, and acidic DM1000 (pH 4–5.25) for the broad stress. Nitrofurantoin is one of the phenotypes where evolved populations gained tolerance in the Biolog analyses, and acid stress has been well-studied in *E. coli* [72]. To control for changes in cell density at stationary phase, we also performed a control measurement in the standard medium, DM1000. Specifically, we measured the growth of evolved replicate populations at days 28, 63, 91, 119, 147, and 175 (generations 465, 1046, 1511, 1977, 2442, and 2907, hereafter designated as generations 500, 1000, 1500, 2000, 2500, and 3000). To do so, we inoculated each population in triplicate from glycerol stock in DM1000 and grew it at 37°C for at least 18 hours. We then diluted the resulting culture 50-fold into DM1000 medium immediately before adding 10 μ L from the diluted culture to 190 μ L media with nitrofurantoin, low pH, or just DM1000 in a 96-well plate. We incubated

the resulting 96-well plates for 24 hours, and then measured the absorbance at 600 nm. We computed the normalized fold change in cell density in each condition at six time points by obtaining the average value of $G = (A_{X,E} / A_{X,A}) / (A_{DM1000,E} / A_{DM1000,A})$ for the three replicate cultures, where $A_{X,Y}$ is the absorbance reading in condition X (e.g., 2.2 $\mu\text{g}/\text{mL}$ nitrofurantoin or acidic DM1000 medium) of a given replicate before evolution ($Y = A$ for ancestral) or after evolution ($Y = E$ for evolved). The denominator, $(A_{DM1000,E} / A_{DM1000,A})$, removes the effect of changes in the evolved carrying capacity, which otherwise could confound cell density changes observed in stressful media with evolved cell density changes in the medium without stressor. We considered that evolution had increased cell density relative to the ancestor when the numerator of G was greater than its denominator, $(A_{X,E} / A_{X,A}) > (A_{DM1000,E} / A_{DM1000,A})$. In order to quantify the relationship between the normalized fold-change in cell density G and ancestral mutation rate, we performed a linear mixed effects analysis using the R package lme4 v1.1–12 [118] to obtain the relationship between G and the mutation rate class. In this analysis, we chose the mutation rate classes (MR^S , MR^M , MR^L , and MR^{XL}) as fixed effects, and the measurement time points, the experimental condition, and each of the 32 evolving replicates as random effects. For the nitrofurantoin and pH stressors, we used data from the experimental condition with the most variability between replicates (2.3 $\mu\text{g}/\text{mL}$ nitrofurantoin and pH 5.25) in this analysis. We tested for homoscedasticity using the R package car v 2.1–2 [120], and found that the untransformed pH data and the log-transformed nitrofurantoin data were homoscedastic (Levene's test, pH 5.25: $F_{3,188} = 1.71$, $p = 0.17$; nitrofurantoin 2.3 $\mu\text{g}/\text{mL}$: $F_{3,188} = 1.25$, $p = 0.29$).

Supporting information

S1 Fig. Cell density (vertical axes) after 24 hours of growth as a function of generation time (horizontal axes). We counted the number of cells in stationary phase just before our daily transfer at regular intervals. Each point is the average cell density of an evolving replicate population at a given generation. One standard deviation above and below the mean is depicted with a shaded line. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains. (PDF)

S2 Fig. Effective population size (N_e) of replicate populations over the course of the evolution experiment. We counted the number of cells at regular intervals, and used these counts to estimate (A) the nominal effective population size N_e for each replicate population. Because our populations are asexual, the effects of selection on polymorphisms linked to neutral sites will make drift at neutral sites appear much stronger than indicated by these estimates. To account for such effects, we also made rough estimates of the effect of linkage on the effective population size using two published methods (further described in Methods), which compute the "Gordo" N_e (B), and the "Good" N_e (C). Together, panels B and C suggest that the effective population size may be much smaller than the nominal population size. Each circle shows the N_e estimate of a replicate population, the center line of the box plot is the median value, and the top and bottom edges of the box correspond to the first and third quartiles. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains. (PDF)

S3 Fig. (A) Fitness differences between ancestral replicate populations and *E. coli* K12 MG1655. Each circle shows the growth rate of a replicate population for a given strain (horizontal axis) minus the growth rate of *E. coli* K12 MG1655 from the same experimental batch. Overall, 54 experimental estimates were made for each strain. (B) Fitness differences between

each evolving replicate population and a common reference strain *E. coli* K12 MG1655 over time are depicted in separate panels for each strain and replicate. Shaded areas indicate one s.e.m. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.

(PDF)

S4 Fig. (A) (C) The fitness difference between each evolving replicate population and its ancestor and its change over time is depicted in separate panels for each strain and replicate. Panels corresponding to the replicates randomly chosen for further characterization in Biolog plates are outlined with a heavy black border. Shaded areas indicate one s.e.m. (B) Variance in relative fitness for the replicate populations of each strain. Strains with higher ancestral mutation rates have more variability in the relative fitness of their evolving populations than those with lower mutation rates. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.

(PDF)

S5 Fig. Percentage of the genome with no sequencing coverage for all 100 sequenced populations. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.

(PDF)

S6 Fig. Frequency and type of SNP in each evolving population over time. Each line in a given panel shows the frequency of one SNP in one replicate population (vertical axis) at generations 0, 1000, 2000, and 3000 (horizontal axis). The color of the line indicates the type of SNP. Types of SNPs with likely functional consequences are emphasized in brown (nonsense mutations) and green (nonsynonymous mutations). Data from all eight independently-evolving replicates (rows of panels) are plotted for each strain (MR^S, MR^M, MR^L, and MR^{XL}; columns of panels).

(PDF)

S7 Fig. The frequency of newly-arising SNPs after one day of growth in the ancestral populations. Several of the observed SNPs, particularly those occurring at higher frequencies, may have been transferred to the eight replicates. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.

(PDF)

S8 Fig. The mutational spectra at four-fold degenerate sites for each evolving replicate strain. (A) Nucleotide changes are depicted along the horizontal axis. For each type of mutation, we computed how often it occurred at any time point during the evolution experiment relative to all other types (Methods). (B) The mutational spectra from replicate populations evolved from ancestors with different mutation rates do not clearly separate when projected onto the first two principal components (PC1 and PC2) in a principal component analysis (Methods). (C) The scree plot shows that PC1 and PC2 account for 43% and 32% of the variability, respectively. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.

(PDF)

S9 Fig. The number of replicates for which at least half of the population harbors any mutation in a putatively beneficial gene. Each dot indicates that a single replicate acquired a mutation within the gene and that this mutation rose to a frequency of at least 50% in a given generation (horizontal axes). In this analysis, we included all genes (rows, labeled in left-most subpanel of each row) that rose to a frequency of >50% in significantly more replicate

populations at generation 3000 than expected by chance alone (see [Methods](#)). Because we evolved eight replicate populations for each strain, each vertical stack of dots can harbor at most eight dots. For many genes, all MR^{XL} replicates share the same nucleotide change, which likely already occurred in the shared ancestor. (A) Genes with different mutations in the same gene in different replicates, and (B) genes where all the MR^{XL} replicates share the same nucleotide change (the nucleotide changes found in the MR^L replicate populations for *betI* and *torA* are not the same as found in the MR^{XL} populations). Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains. (PDF)

S10 Fig. The evolutionary dynamics of mutations in the eight putatively beneficial genes.

Each circle corresponds to one evolving replicate population. The size of a circle is proportional to the frequency at which a mutation is found in a population, and can change over time (horizontal axes). All replicates for all strains (circles inside each panel) are depicted for each gene (labeled in the top, left of each panel). Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains. (PDF)

S11 Fig. Cell density of two randomly selected evolved and ancestral strains in 96 different environments on Biolog plates.

Importantly, all tested ancestor and evolved MR^{XL} strains failed to grow in every one of the 96 environments. Each circle represents the ancestor's density (horizontal axes) and the evolved replicate population's density (vertical axes) in a particular environment. Points above the diagonal line correspond to conditions in which an evolved replicate population outperformed its ancestor; points below the line correspond to conditions in which an evolved replicate population underperformed its ancestor. We consider a population to have evolved tolerance to a condition when its density is larger than the ancestral density in the same condition, excluding differences attributable to experimental noise. Conversely, we consider a population as having experienced decay if its density after evolution is smaller than that of its ancestor (see [Methods](#)). Both gains and decays are indicated by solid circles. Open circles indicate that no gain or decay was detected for that condition, or that the difference between the evolved and ancestral cell density could be due to experimental noise. Different colors distinguish data from the MR^S (blue), MR^M (yellow), and MR^L (orange) strains. (PDF)

S12 Fig. The fold-change in cell density after 24 hours of growth of evolved replicate populations (relative to their ancestor) in media with nitrofurantoin and low pH media.

We measured the cell density of the MR^S, MR^M, MR^L, and MR^{XL} evolved replicate populations relative to their ancestors over the course of the experiment (horizontal axis) in (A) the antibiotic nitrofurantoin (1.5 µg/mL–2.4 µg/mL), and (B) acidic media (pH 4.0–5.25). Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains. (PDF)

S13 Fig. Comparisons of multiple population properties (matrix diagonal) for the experimental data from each evolved replicate population.

We plotted each property in a pairwise fashion to identify correlations between properties. Each property is listed on the diagonal ("log(U)" is the logarithm of the genomic mutation rate, "relative fitness" is the evolved growth rate relative to the ancestor, "N_e" is the effective population size, "cell density at 3000" is the absorbance reading at 600 nm at generation 3000 after 24 hours of growth in minimal medium, "log(derived alleles)" is the logarithm of the number of high frequency derived alleles at generation 3000, "log(cloud size)" is the logarithm of the population's average distance to the center of the cloud at generation 3000, "log(pH cell density)" is the logarithm of the

normalized fold change in cell density after 24 hours of growth in acidic media at pH 5.25, and "log(nitro cell density)" is the logarithm of the normalized fold change in cell density after 24 hours of growth in media containing 2.2 $\mu\text{g}/\text{mL}$ nitrofurantoin). Pairwise comparisons are plotted below the diagonal; each circle corresponds to a different replicate population. The Spearman correlation coefficient of each panel below the diagonal is reported in the corresponding panel above the diagonal. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.
(PDF)

S14 Fig. No evidence that the mutation rate genome is preferentially subject to genetic change. (A) We calculated the percentage of synonymous nucleotide changes (at any frequency) that occurred within genes belonging to the mutation rate genome (vertical axes) during the evolution experiment (horizontal axes) at any frequency in each evolving replicate population (circles). Horizontal gray lines indicate the percentage of coding regions in the *E. coli* genome that belong to the mutation rate genome (2.8%). There are no more mutations in the mutation rate genome than expected by chance alone at generation 3000 (one-sided binomial test, MR^S:n = 55, p = 1.0; MR^M:n = 48, p = 1.0; MR^L:n = 433, p = 0.79; MR^{XL}:n = 1050, p = 0.99). (B) We calculated the mean synonymous nucleotide site diversity and its standard error (Methods). The mean synonymous nucleotide site diversity for the mutation rate genome is depicted in the right panel, and for all other genes in the left panel. Note that no or very few sites may contribute to average diversity at low mutation rates. Shaded areas indicate one standard error of the mean. Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.
(PDF)

S15 Fig. The evolutionary dynamics of possibly function-altering mutations in the mutation rate genome. Each circle corresponds to a putatively function-altering mutation (nonsynonymous or nonsense mutations in protein-coding genes, or any mutation in tRNA-encoding genes) in one evolving replicate population. The size of a circle is proportional to the frequency at which a mutation is found in a population, and can change over time (horizontal axes). All replicates for all strains (circles inside each panel) are depicted for each gene (labeled on the top left of each panel). Different colors distinguish data from the MR^S (blue), MR^M (yellow), MR^L (orange), and MR^{XL} (red) strains.
(PDF)

S1 Table. All single nucleotide differences between the ancestor strains MR^S, MR^M, MR^L, and MR^{XL}. The columns are as follows: "Position" is the genomic coordinate of the SNP; "Ref" is the consensus ancestral sequence; "MR^S", "MR^M", "MR^L", and "MR^{XL}" display the SNPs in the each of the respective ancestral strains (a period means that the nucleotide is the same as in the consensus ancestral sequence); "Gene" is the name of the gene in or near which the SNP occurs; "SNP type" defines what effect the polymorphism may have (e.g., nonsynonymous, intergenic, noncoding); and "AA effect" is a SNP's effect on the protein sequence for SNPs occurring within protein-coding genes. SNPs occurring in intergenic regions are annotated with the nearest 5' and 3' genes.
(XLSX)

S2 Table. Genes putatively involved in modulating the mutation rate.
(XLSX)

S3 Table. Genomic mutation rates. The columns are as follows: "Strain" is the identity of the ancestral strain (e.g., MR^S, MR^M, MR^L, MR^{XL}); "Replicate" identifies the replicate population;

" μ " is the mutation rate in units of number of mutations per bp per cell generation; " μ 95% confidence intervals" "Upper" and "Lower" give the confidence in the mutation rate estimate; " U " is the genomic mutation rate in number of mutations per genome and per cell generation; " U 95% confidence intervals" "Upper" and "Lower" give the confidence in the genomic mutation rate estimate; " $N_e U$ " is the estimated number of mutations to occur in an evolving population each generation. Mutation rates were measured at generations 0 ("Ancestor" replicates) and 3000 (all other replicates).

(XLSX)

S1 Text. Area Under the Curve (AUC) is a complementary fitness metric that also demonstrates reduced adaptation at very high mutation rates.

(PDF)

S2 Text. Applicability of several theoretical models predicting loss or reduction of adaptation at high mutation rates.

(PDF)

S3 Text. The waiting time for the establishment of a new beneficial allele.

(PDF)

Author Contributions

Conceptualization: Kathleen Sprouffske, Paul Sniegowski, Andreas Wagner.

Data curation: Kathleen Sprouffske.

Formal analysis: Kathleen Sprouffske.

Funding acquisition: Kathleen Sprouffske, Andreas Wagner.

Investigation: Kathleen Sprouffske, José Aguilar-Rodríguez, Andreas Wagner.

Methodology: Kathleen Sprouffske.

Project administration: Kathleen Sprouffske, Andreas Wagner.

Resources: Kathleen Sprouffske, Paul Sniegowski, Andreas Wagner.

Software: Kathleen Sprouffske.

Supervision: Kathleen Sprouffske, Paul Sniegowski, Andreas Wagner.

Validation: Kathleen Sprouffske, José Aguilar-Rodríguez.

Visualization: Kathleen Sprouffske.

Writing – original draft: Kathleen Sprouffske.

Writing – review & editing: Kathleen Sprouffske, José Aguilar-Rodríguez, Paul Sniegowski, Andreas Wagner.

References

1. Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations. *Nat Rev Genet.* 2007; 8: 610–618. <https://doi.org/10.1038/nrg2146> PMID: 17637733
2. Kunkel TA, Bebenek K. DNA replication fidelity. *Annu Rev Biochem.* 2000; 69: 497–529. <https://doi.org/10.1146/annurev.biochem.69.1.497> PMID: 10966467
3. Kunkel TA, Erie DA. DNA mismatch repair. *Annu Rev Biochem.* 2005; 74: 681–710. <https://doi.org/10.1146/annurev.biochem.74.082803.133243> PMID: 15952900
4. Fisher RA. The genetical theory of natural selection. Oxf Univ Press. 1930; 318.

5. Muller H. Some genetic aspects of sex. *Am Nat.* 1932; 66: 118–138.
6. Felsenstein J. The evolutionary advantage of recombination. *Genetics.* 1974; 78: 737–756. PMID: [4448362](#)
7. Hill WG, Robertson A. The effect of linkage on limits to artificial selection. *Genet Res.* 1966; 8: 269–294. PMID: [5980116](#)
8. Charlesworth B, Morgan MT, Charlesworth D. The effect of deleterious mutations on neutral molecular variation. *Genetics.* 1993; 134: 1289–1303. PMID: [8375663](#)
9. Johnson T, Barton NH. The effect of deleterious alleles on adaptation in asexual populations. *Genetics.* 2002; 162: 395. PMID: [12242249](#)
10. Desai MM, Fisher DS. Beneficial mutation selection balance and the effect of linkage on positive selection. *Genetics.* 2007; 176: 1759–1798. <https://doi.org/10.1534/genetics.106.067678> PMID: [17483432](#)
11. Barton NH. Genetic linkage and natural selection. *Philos Trans R Soc B Biol Sci.* 2010; 365: 2559–2569. <https://doi.org/10.1098/rstb.2010.0106> PMID: [20643746](#)
12. Charlesworth B. The effects of deleterious mutations on evolution at linked sites. *Genetics.* 2012; 190: 5–22. <https://doi.org/10.1534/genetics.111.134288> PMID: [22219506](#)
13. Cutter AD, Payseur BA. Genomic signatures of selection at linked sites: unifying the disparity among species. *Nat Rev Genet.* 2013; 14: 262–274. <https://doi.org/10.1038/nrg3425> PMID: [23478346](#)
14. Charlesworth B, Campos JL. The relations between recombination rate and patterns of molecular variation and evolution in *Drosophila*. *Annu Rev Genet.* 2014; 48: 383–403. <https://doi.org/10.1146/annurev-genet-120213-092525> PMID: [25251853](#)
15. Desai MM, Fisher DS, Murray AW. The speed of evolution and maintenance of variation in asexual populations. *Curr Biol.* 2007; 17: 385–394. <https://doi.org/10.1016/j.cub.2007.01.072> PMID: [17331728](#)
16. Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. Role of mutator alleles in adaptive evolution. *Nature.* 1997; 387: 700–702. <https://doi.org/10.1038/42696> PMID: [9192893](#)
17. de Visser JAGM, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE. Diminishing returns from mutation supply rate in asexual populations. *Science.* 1999; 283: 404–406. PMID: [9888858](#)
18. Wisner MJ, Ribick N, Lenski RE. Long-term dynamics of adaptation in asexual populations. *Science.* 2013; 342: 1364–1367. <https://doi.org/10.1126/science.1243357> PMID: [24231808](#)
19. Eigen M, Schuster P. The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften.* 1977; 64: 541–565. PMID: [593400](#)
20. Bull JJ, Wilke CO. Lethal mutagenesis of bacteria. *Genetics.* 2008; 180: 1061–1070. <https://doi.org/10.1534/genetics.108.091413> PMID: [18780744](#)
21. Gerrish PJ, Colato A, Sniegowski PD. Genomic mutation rates that neutralize adaptive evolution and natural selection. *J R Soc Interface.* 2013; 10: 20130329–20130329. <https://doi.org/10.1098/rsif.2013.0329> PMID: [23720539](#)
22. Lynch M, Bürger R, Butcher D, Gabriel W. The mutational meltdown in asexual populations. *J Hered.* 1993; 84: 339–344. PMID: [8409355](#)
23. Swetina J, Schuster P. Self-replication with errors. A model for polynucleotide replication. *Biophys Chem.* 1982; 16: 329–345. PMID: [7159681](#)
24. Smith JM. Models of evolution. *Proc R Soc B-Biol Sci.* 1983; 219: 315–325.
25. Mao EF, Lane L, Lee J, Miller JH. Proliferation of mutators in a cell population. *J Bacteriol.* 1997; 179: 417–422. PMID: [8990293](#)
26. Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature.* 1997; 387: 703–705. <https://doi.org/10.1038/42701> PMID: [9192894](#)
27. Giraud A, Matic I, Tenailon O, Clara A, Radman M, Fons M, et al. Costs and benefits of high mutation rates: Adaptive evolution of bacteria in the mouse gut. *Science.* 2001; 291: 2606–2608. <https://doi.org/10.1126/science.1056421> PMID: [11283373](#)
28. Notley-McRobb L, Seeto S, Ferenci T. Enrichment and elimination of *mutY* mutators in *Escherichia coli* populations. *Genetics.* 2002; 162: 1055–1062. PMID: [12454055](#)
29. Notley-McRobb L, King T, Ferenci T. *rpoS* mutations and loss of general stress resistance in *Escherichia coli* populations as a consequence of conflict between competing stress responses. *J Bacteriol.* 2002; 184: 806–811. <https://doi.org/10.1128/JB.184.3.806-811.2002> PMID: [11790751](#)
30. Pal C, Maciá MD, Oliver A, Schachar I, Buckling A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature.* 2007; 450: 1079–1081. <https://doi.org/10.1038/nature06350> PMID: [18059461](#)
31. Swings T, Van den Bergh B, Wuyts S, Oeyen E, Voordeckers K, Verstrepen KJ, et al. Adaptive tuning of mutation rates allows fast response to lethal stress in *Escherichia coli*. *eLife.* 2017; 6: e22939. <https://doi.org/10.7554/eLife.22939> PMID: [28460660](#)

32. Desai MM, Fisher DS. The balance between mutators and nonmutators in asexual populations. *Genetics*. 2011; 188: 997–1014. <https://doi.org/10.1534/genetics.111.128116> PMID: 21652523
33. Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. The evolution of mutation rates: separating causes from consequences. *BioEssays*. 2000; 22: 1057–1066. [https://doi.org/10.1002/1521-1878\(200012\)22:12<1057::AID-BIES3>3.0.CO;2-W](https://doi.org/10.1002/1521-1878(200012)22:12<1057::AID-BIES3>3.0.CO;2-W) PMID: 11084621
34. Chao L, Cox EC. Competition between high and low mutating strains of *Escherichia coli*. *Evolution*. 1983; 37: 125–134. <https://doi.org/10.1111/j.1558-5646.1983.tb05521.x> PMID: 28568016
35. Gentile CF, Yu S-C, Serrano SA, Gerrish PJ, Sniegowski PD. Competition between high- and higher-mutating strains of *Escherichia coli*. *Biol Lett*. 2011; 7: 422–424. <https://doi.org/10.1098/rsbl.2010.1036> PMID: 21227974
36. Giraud A, Radman M, Matic I, Taddei F. The rise and fall of mutator bacteria. *Curr Opin Microbiol*. 2001; 4: 582–585. PMID: 11587936
37. Cox EC, Gibson TC. Selection for high mutation rates in chemostats. *Genetics*. 1974; 77: 169–184. PMID: 4603159
38. Sturtevant AH. Essays on evolution. I. On the effects of selection on mutation rate. *Q Rev Biol*. 1937; 12: 464–467.
39. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. *Genetics*. 1998; 148: 1667–1686. PMID: 9560386
40. Kimura M. On the evolutionary adjustment of spontaneous mutation rates. *Genet Res*. 1967; 9: 23–34.
41. Kondrashov AS. Deleterious mutations and the evolution of sexual reproduction. *Nature*. 1988; 336: 435–440. <https://doi.org/10.1038/336435a0> PMID: 3057385
42. Kondrashov AS. Modifiers of mutation-selection balance: general approach and the evolution of mutation rates. *Genet Res*. 1995; 66: 53–69.
43. Lynch M. The lower bound to the evolution of mutation rates. *Genome Biol Evol*. 2011; 3: 1107–1118. <https://doi.org/10.1093/gbe/evr066> PMID: 21821597
44. Lynch M, Ackerman MS, Gout J-F, Long H, Sung W, Thomas WK, et al. Genetic drift, selection and the evolution of the mutation rate. *Nat Rev Genet*. 2016; 17: 704–714. <https://doi.org/10.1038/nrg.2016.104> PMID: 27739533
45. Loh E, Salk JJ, Loeb LA. Optimization of DNA polymerase mutation rates during bacterial evolution. *Proc Natl Acad Sci*. 2010; 107: 1154–1159. <https://doi.org/10.1073/pnas.0912451107> PMID: 20080608
46. McDonald MJ, Hsieh Y-Y, Yu Y-H, Chang S-L, Leu J-Y. The evolution of low mutation rates in experimental mutator populations of *Saccharomyces cerevisiae*. *Curr Biol*. 2012; 22: 1235–1240. <https://doi.org/10.1016/j.cub.2012.04.056> PMID: 22727704
47. Singh T, Hyun M, Sniegowski P. Evolution of mutation rates in hypermutable populations of *Escherichia coli* propagated at very small effective population size. *Biol Lett*. 2017; 13: 20160849. <https://doi.org/10.1098/rsbl.2016.0849> PMID: 28250208
48. Tröbner W, Piechocki R. Selection against hypermutability in *Escherichia coli* during long term evolution. *Mol Genet Genomics*. 1984; 198: 177–178.
49. Turrientes M-C, Baquero F, Levin BR, Martínez J-L, Ripoll A, González-Alba J-M, et al. Normal mutation rate variants arise in a mutator (Mut S) *Escherichia coli* population. *Marinus MG*, editor. *PLOS ONE*. 2013; 8: e72963. <https://doi.org/10.1371/journal.pone.0072963> PMID: 24069167
50. Wielgoss S, Barrick JE, Tenaillon O, Wisner MJ, Dittmar WJ, Cruveiller S, et al. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc Natl Acad Sci*. 2013; 110: 222–227. <https://doi.org/10.1073/pnas.1219574110> PMID: 23248287
51. Raynes Y, Sniegowski PD. Experimental evolution and the dynamics of genomic mutation rate modifiers. *Heredity*. 2014; 113: 375–380. <https://doi.org/10.1038/hdy.2014.49> PMID: 24849169
52. Robertson A. Inbreeding in artificial selection programmes. *Genet Res*. 1961; 2: 189. <https://doi.org/10.1017/S0016672300000690>
53. Price MN, Arkin AP. Weakly deleterious mutations and low rates of recombination limit the impact of natural selection on bacterial genomes. *mBio*. 2015; 6: e01302–15. <https://doi.org/10.1128/mBio.01302-15> PMID: 26670382
54. Gordo I, Navarro A, Charlesworth B. Muller's ratchet and the pattern of variation at a neutral locus. *Genetics*. 2002; 161: 835–848. PMID: 12072478
55. Good BH, Walczak AM, Neher RA, Desai MM. Genetic diversity in the interference selection limit. *PLOS Genet*. 2014; 10: e1004222. <https://doi.org/10.1371/journal.pgen.1004222> PMID: 24675740

56. Deatherage DE, Barrick JE. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. In: Sun L, Shou W, editors. Engineering and analyzing multicellular systems: Methods and protocols. New York, NY: Springer Science+Business Media; 2014. pp. 165–188.
57. Garibyan L, Huang T, Kim M, Wolff E, Nguyen A, Nguyen T, et al. Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. DNA Repair. 2003; 2: 593–608. [https://doi.org/10.1016/S1568-7864\(03\)00024-7](https://doi.org/10.1016/S1568-7864(03)00024-7) PMID: 12713816
58. Horst JP, Wu TH, Marinus MG. *Escherichia coli* mutator genes. Trends Microbiol. 1999; 7: 29–36. PMID: 10068995
59. Wu TH, Clarke CH, Marinus MG. Specificity of *Escherichia coli* *mutD* and *mutL* mutator strains. Gene. 1990; 87: 1–5. PMID: 2185133
60. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. Negative epistasis between beneficial mutations in an evolving bacterial population. Science. 2011; 332: 1193–1196. <https://doi.org/10.1126/science.1203801> PMID: 21636772
61. Schneider D, Duperchy E, Coursange E, Lenski RE, Blot M. Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. Genetics. 2000; 156: 477–488. PMID: 11014799
62. Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, et al. The molecular diversity of adaptive convergence. Science. 2012; 335: 457–461. <https://doi.org/10.1126/science.1212986> PMID: 22282810
63. Wang Y, Diaz Arenas C, Stoebel DM, Cooper TF. Genetic background affects epistatic interactions between two beneficial mutations. Biol Lett. 2012; 9: 20120328. <https://doi.org/10.1098/rsbl.2012.0328> PMID: 22896270
64. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. Proc Natl Acad Sci. 2006; 103: 9107–9112. <https://doi.org/10.1073/pnas.0602917103> PMID: 16751270
65. Woods RJ, Barrick JE, Cooper TF, Shrestha U, Kauth MR, Lenski RE. Second-order selection for evolvability in a large *Escherichia coli* population. Science. 2011; 331: 1433–1436. <https://doi.org/10.1126/science.1198914> PMID: 21415350
66. Rath D, Jawali N. Loss of expression of *cspC*, a cold shock family gene, confers a gain of fitness in *Escherichia coli* K-12 strains. J Bacteriol. 2006; 188: 6780–6785. <https://doi.org/10.1128/JB.00471-06> PMID: 16980479
67. Barrick JE, Lenski RE. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. Cold Spring Harb Symp Quant Biol. 2009; 74: 119–129. <https://doi.org/10.1101/sqb.2009.74.018> PMID: 19776167
68. Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, et al. Tempo and mode of genome evolution in a 50,000-generation experiment. Nature. 2016; 536: 165–170.
69. Bochner BR, Gadzinski P, Panomitros E. Phenotype MicroArrays for high-throughput phenotypic testing and assay of gene function. Genome Res. 2001; 11: 1246–1255. <https://doi.org/10.1101/gr.186501> PMID: 11435407
70. Leiby N, Marx CJ. Metabolic erosion primarily through mutation accumulation, and not tradeoffs, drives limited evolution of substrate specificity in *Escherichia coli*. PLOS Biol. 2014; 12: e1001789. <https://doi.org/10.1371/journal.pbio.1001789> PMID: 24558347
71. Sandegren L, Lindqvist A, Kahlmeter G, Andersson DI. Nitrofurantoin resistance mechanism and fitness cost in *Escherichia coli*. J Antimicrob Chemother. 2008; 62: 495–503. <https://doi.org/10.1093/jac/dkn222> PMID: 18544599
72. Foster JW. *Escherichia coli* acid resistance: Tales of an amateur acidophile. Nat Rev Microbiol. 2004; 2: 898–907. <https://doi.org/10.1038/nrmicro1021> PMID: 15494746
73. Kanjee U, Houry WA. Mechanisms of acid resistance in *Escherichia coli*. Annu Rev Microbiol. 2013; 67: 65–81. <https://doi.org/10.1146/annurev-micro-092412-155708> PMID: 23701194
74. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci U S A. 1991; 88: 7160–7164. <https://doi.org/10.1073/pnas.88.16.7160> PMID: 1831267
75. Miller JH. Spontaneous mutators in bacteria: Insights into pathways of mutagenesis and repair. Annu Rev Microbiol. 1995; 50: 625–643. <https://doi.org/10.1146/annurev.micro.50.1.625> PMID: 8905093
76. Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martinez C, et al. EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Res. 2013; 41: D605–D612. <https://doi.org/10.1093/nar/gks1027> PMID: 23143106
77. MacLean RC, Torres-Barceló C, Moxon R. Evaluating evolutionary models of stress-induced mutagenesis in bacteria. Nat Rev Genet. 2013; 14: 221–227. <https://doi.org/10.1038/nrg3415> PMID: 23400102

78. Landini P, Egli T, Wolf J, Lacour S. sigmaS, a major player in the response to environmental stresses in *Escherichia coli*: Role, regulation and mechanisms of promoter recognition. *Environ Microbiol Rep*. 2014; 6: 1–13. <https://doi.org/10.1111/1758-2229.12112> PMID: 24596257
79. Zambrano MM, Siegele DA, Almirón M. Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. *Science*. 1993; 259: 1757–1760. PMID: 7681219
80. King T, Seeto S, Ferenci T. Genotype-by-environment interactions influencing the emergence of *rpoS* mutations in *Escherichia coli* populations. *Genetics*. 2005; 172: 2071–2079. <https://doi.org/10.1534/genetics.105.053892> PMID: 16489226
81. Layton JC, Foster PL. Error-prone DNA polymerase IV is controlled by the stress-response sigma factor, RpoS, in *Escherichia coli*. *Mol Microbiol*. 2003; 50: 549–561. <https://doi.org/10.1046/j.1365-2958.2003.03704.x> PMID: 14617178
82. Wielgoss S, Barrick JE, Tenaillon O, Cruveiller S, Chane-Woon-Ming B, Médigue C, et al. Mutation rate inferred from synonymous substitutions in a long-term evolution experiment with *Escherichia coli*. *G3 Genes Genomes Genet*. 2011; 1: 183–186. <https://doi.org/10.1534/g3.111.000406/-DC1>
83. Schoustra SE, Bataillon T, Gifford DR, Kassen R. The properties of adaptive walks in evolving populations of fungus. Barton NH, editor. *PLOS Biol*. 2009; 7: e1000250. <https://doi.org/10.1371/journal.pbio.1000250> PMID: 19956798
84. Lenski RE, Wisser MJ, Ribbeck N, Blount ZD, Nahum JR, Morris JJ, et al. Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with *Escherichia coli*. *Proc R Soc B-Biol Sci*. 2015; 282: 20152292. <https://doi.org/10.1098/rspb.2015.2292> PMID: 26674951
85. 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. <https://doi.org/10.1126/science.1203799> PMID: 21636771
86. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science*. 2014; 344: 1519–1522. <https://doi.org/10.1126/science.1250939> PMID: 24970088
87. Lenski RE, Travisano M. Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. *Proc Natl Acad Sci U S A*. 1994; 91: 6808–6814. PMID: 8041701
88. Gould SJ. *Wonderful life: The burgess shale and the nature of history*. New York: W.W. Norton & Company; 1989.
89. 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. <https://doi.org/10.1073/pnas.0803151105> PMID: 18524956
90. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature*. 2012; 489: 513–+. <https://doi.org/10.1038/nature11514> PMID: 22992527
91. Kimura M, Ohta T. The average number of generations until fixation of a mutant gene in a finite population. *Genetics*. 1969; 61: 763–771. PMID: 17248440
92. Muller HJ. The relation of recombination to mutational advance. *Mutat Res*. 1964; 106: 2–9. PMID: 14195748
93. Haigh J. The accumulation of deleterious genes in a population—Muller’s Ratchet. *Theor Popul Biol*. 1978; 14: 251–267. PMID: 746491
94. Kimura M, Maruyama T. The mutational load with epistatic gene interactions in fitness. *Genetics*. 1966; 54: 1337–1351. PMID: 17248359
95. Good BH, Desai MM. Evolution of mutation rates in rapidly adapting asexual populations. *Genetics*. 2016; 204: 1249–1266. <https://doi.org/10.1534/genetics.116.193565> PMID: 27646140
96. Luria S, Delbrück M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*. 1943; 28: 491–511. PMID: 17247100
97. Rosche WA, Foster PL. Determining mutation rates in bacterial populations. *Methods*. 2000; 20: 4–17. <https://doi.org/10.1006/meth.1999.0901> PMID: 10610800
98. Funchain P, Yeung A, Stewart JL, Lin R, Slupska MM, Miller JH. The consequences of growth of a mutator strain of *Escherichia coli* as measured by loss of function among multiple gene targets and loss of fitness. *Genetics*. 2000; 154: 959–970. PMID: 10757746
99. Denamur E, Tenaillon O, Deschamps C, Skurnik D, Ronco E, Gaillard JL, et al. Intermediate mutation frequencies favor evolution of multidrug resistance in *Escherichia coli*. *Genetics*. 2005; 171: 825–827. <https://doi.org/10.1534/genetics.105.045526> PMID: 15965238
100. Loeb LA, Essigmann JM, Kazazi F, Zhang J, Rose KD, Mullins JI. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc Natl Acad Sci U S A*. 1999; 96: 1492–1497. PMID: 9990051

101. Lynch M, Gabriel W. Mutation load and the survival of small populations. *Evol Int J Org Evol*. 1990; 1725–1737.
102. Solé RV, Deisboeck TS. An error catastrophe in cancer? *J Theor Biol*. 2004; 228: 47–54. <https://doi.org/10.1016/j.jtbi.2003.08.018> PMID: 15064082
103. Fox EJ, Loeb LA. Lethal mutagenesis: targeting the mutator phenotype in cancer. *Semin Cancer Biol*. 2010; 20: 353–359. <https://doi.org/10.1016/j.semcancer.2010.10.005> PMID: 20934515
104. Frank SA, Nowak MA. Problems of somatic mutation and cancer. *BioEssays*. 2004; 26: 291–299. <https://doi.org/10.1002/bies.20000> PMID: 14988930
105. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci*. 2000; 97: 6640–6645. <https://doi.org/10.1073/pnas.120163297> PMID: 10829079
106. Cherepanov PP, Wackernagel W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*. 1995; 158: 9–14. PMID: 7789817
107. Singer M, Baker TA, Schnitzler G, Deischel SM, Goel M, Dove W, et al. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. *Microbiol Rev*. 1989; 53: 1–24. PMID: 2540407
108. Gerrish P. A simple formula for obtaining markedly improved mutation rate estimates. *Genetics*. 2008; 180: 1773–1778. <https://doi.org/10.1534/genetics.108.091777> PMID: 18832356
109. Siegel EC, Kamel F. Reversion of frameshift mutations by mutator genes in *Escherichia coli*. *J Bacteriol*. 1974; 117: 994–1001. PMID: 4360543
110. Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of *Escherichia coli* K-12. *Science*. 1997; 277: 1453–1462. PMID: 9278503
111. Carlton BC, Brown BJ. Gene mutation. In: Gerhardt P, editor. *Manual of methods of general bacteriology*. Washington, DC: American Society for Microbiology Press; 1981. pp. 222–242.
112. Lenski RE, Rose MR, Simpson SC, Tadler SC. Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat*. 1991; 138: 1315.
113. Samani P, Bell G. Adaptation of experimental yeast populations to stressful conditions in relation to population size. *J Evol Biol*. 2010; 23: 791–796. <https://doi.org/10.1111/j.1420-9101.2010.01945.x> PMID: 20149025
114. Gessler DD. The constraints of finite size in asexual populations and the rate of the ratchet. *Genet Res*. 1995; 66: 241–253. PMID: 16553995
115. Trindade S, Perfeito L, Gordo I. Rate and effects of spontaneous mutations that affect fitness in mutator *Escherichia coli*. *Philos Trans R Soc Lond B Biol Sci*. 2010; 365: 1177–1186. <https://doi.org/10.1098/rstb.2009.0287> PMID: 20308092
116. Crow JF, Kimura M. *An introduction to population genetics theory*. (reprinted by Blackburn Press, Caldwell, NJ). New York, NY: Harper and Row, Publishers, Inc.; 1970.
117. Sprouffske K, Wagner A. Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*. 2016; 17: 172. <https://doi.org/10.1186/s12859-016-1016-7> PMID: 27094401
118. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015; 67: 1–48. <https://doi.org/10.18637/jss.v067.i01>
119. Levene H. Robust tests for equality of variances. In: Olkin Ingram, Ghurye Sudhish G, Hoeffding Wasily, Madow William G, Mann Henry B, editors. *Contributions to probability and statistics: Essays in honor of Harold Hotelling*. Stanford, CA: Stanford University Press; 1960. pp. 278–292.
120. Fox J, Weisberg S. *An R companion to applied regression*. 2nd ed. Los Angeles: SAGE; 2011.
121. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. *Biom J*. 2008; 50: 346–363. <https://doi.org/10.1002/bimj.200810425> PMID: 18481363
122. Sprouffske K, Aguilar-Rodríguez J, Wagner A. How archiving by freezing affects the genome-scale diversity of *Escherichia coli* populations. *Genome Biol Evol*. 2016; 8: 1290–1298. <https://doi.org/10.1093/gbe/eww054> PMID: 26988250
123. Nei M, Li W-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U S A*. 1979; 76: 5269–5273. PMID: 291943
124. Hartl DL, Clark AG. *Principles of population genetics*. Sinauer Associates Incorporated; 1997.
125. Shao W, Kearney MF, Boltz VF, Spindler JE, Mellors JW, Maldarelli F, et al. PAPNC, a novel method to calculate nucleotide diversity from large scale next generation sequencing data. *J Virol Methods*. 2014; 203: 73–80. <https://doi.org/10.1016/j.jviromet.2014.03.008> PMID: 24681054

126. Nelson CW, Hughes AL. Within-host nucleotide diversity of virus populations: Insights from next-generation sequencing. *Infect Genet Evol.* 2015; 30: 1–7. <https://doi.org/10.1016/j.meegid.2014.11.026> PMID: [25481279](https://pubmed.ncbi.nlm.nih.gov/25481279/)
127. Crane GJ, Thomas SM, Jones ME. A modified Luria-Delbrück fluctuation assay for estimating and comparing mutation rates. *Mutat Res.* 1996; 354: 171–182. PMID: [8764946](https://pubmed.ncbi.nlm.nih.gov/8764946/)
128. Bochner BR. Global phenotypic characterization of bacteria. *FEMS Microbiol Rev.* 2009; 33: 191–205. <https://doi.org/10.1111/j.1574-6976.2008.00149.x> PMID: [19054113](https://pubmed.ncbi.nlm.nih.gov/19054113/)