

# Evolutionary Dynamics of Asexual Hypermutators Adapting to a Novel Environment

Wei-Chin Ho <sup>1,\*</sup>, Megan G. Behringer<sup>1,2,3</sup>, Samuel F. Miller<sup>1</sup>, Jadon Gonzales<sup>1</sup>, Amber Nguyen<sup>1</sup>, Meriem Allahwerdy<sup>1</sup>, Gwyneth F. Boyer<sup>1</sup>, and Michael Lynch<sup>1,\*</sup>

<sup>1</sup>Center for Mechanisms of Evolution, The Biodesign Institute, Arizona State University, Tempe, Arizona, USA

<sup>2</sup>Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, USA

<sup>3</sup>Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee, USA

\*Corresponding authors: E-mails: weichinh@asu.edu; mlynch11@asu.edu.

Accepted: 14 November 2021

## Abstract

How microbes adapt to a novel environment is a central question in evolutionary biology. Although adaptive evolution must be fueled by beneficial mutations, whether higher mutation rates facilitate the rate of adaptive evolution remains unclear. To address this question, we cultured *Escherichia coli* hypermutating populations, in which a defective methyl-directed mismatch repair pathway causes a 140-fold increase in single-nucleotide mutation rates. In parallel with wild-type *E. coli*, populations were cultured in tubes containing Luria-Bertani broth, a complex medium known to promote the evolution of subpopulation structure. After 900 days of evolution, in three transfer schemes with different population-size bottlenecks, hypermutators always exhibited similar levels of improved fitness as controls. Fluctuation tests revealed that the mutation rates of hypermutator lines converged evolutionarily on those of wild-type populations, which may have contributed to the absence of fitness differences. Further genome-sequence analysis revealed that, although hypermutator populations have higher rates of genomic evolution, this largely reflects strong genetic linkage. Despite these linkage effects, the evolved population exhibits parallelism in fixed mutations, including those potentially related to biofilm formation, transcription regulation, and mutation-rate evolution. Together, these results are generally inconsistent with a hypothesized positive relationship between the mutation rate and the adaptive speed of evolution, and provide insight into how clonal adaptation occurs in novel environments.

**Key words:** adaptation, bottleneck effects, drift barrier, *Escherichia coli*, mutational load, mutation rate.

## Significance

Although mutations are a critical source for the adaptation in a new environment, whether or not elevated mutation rates can lead to elevated adaptation rates remains unclear, especially when the environment is heterogeneous. To address this issue, we evolved *E. coli* populations with different starting mutation rates in a complex medium for 900 days and then examined their fitness and genome profiles. In the populations with higher starting mutation rates, despite faster rates of genome evolution, fitness improvement is not significantly elevated, and most of the accumulated mutations represent passive consequences of linkage effects.

## Introduction

Beneficial mutations are the ultimate source of adaptive evolution. Therefore, it is of interest to study how changes to mutational processes can influence the pace of adaptive processes. In terms of mutation rates, a theoretically complicated

relationship with the rate of adaptation in asexual populations was proposed in the early studies on the evolution of sex (Muller 1932; Crow and Kimura 1965). These studies posited that, when mutation rates are low, such that the waiting time

for a beneficial mutation to arise in a population remains long, increases in the mutation rate can result in linear increases in the adaptation rate. In contrast, when mutation rates are relatively high, such that multiple beneficial mutations frequently arise in different individuals within an asexual population, beneficial mutations may interfere with each other's opportunity to spread through the entire population. Thus, the facilitating effect of mutation on the rate of adaptation becomes diminished; a phenomenon later termed clonal interference (Gerrish and Lenski 1998).

More recent theoretical studies have suggested that the effective number of beneficial mutations per population is critical for the strength of clonal interference, and effective population sizes and effect-size distributions of mutations have also been proposed to be influential (Gerrish and Lenski 1998; Wilke 2004; Kim and Orr 2005; Bollback and Huelsenbeck 2007; Desai and Fisher 2007; Park and Krug 2007; Campos and Wahl 2010; Park et al. 2010; Good et al. 2012). When mutation rates are high, rare beneficial mutations can commonly emerge in genomes already containing some deleterious mutations, and such linkage can further impede adaptation, especially in asexual populations (Bachtrog and Gordo 2004; Good and Desai 2014; Luksza and Lassig 2014; McFarland et al. 2017; Penisson et al. 2017).

The relationship between rates of mutation and adaptation is further complicated because mutation rates can be plastically different in various environments (Williams and Foster 2012; Long et al. 2016; Shewaramani et al. 2017) and evolve over time (Wielgoss et al. 2013; Swings et al. 2017). Given that most mutations are deleterious, high mutation rates create high mutational loads, potentially driving the spread of antimutator alleles and resulting in the evolution of lower mutation rates (Muller 1950; Kimura 1967; Lynch 2008). According to the drift-barrier hypothesis, the reduction of the mutation rate should continue until the addition of new antimutator alleles no longer reduces the mutational load to a significant enough extent to overcome genetic drift (Lynch 2010; Sung et al. 2012; Lynch et al. 2016). Consequently, if two populations with initially different mutation rates adapt to the same constant environment, the resulting difference in fitness-improvement rates can be less than predicted if both populations converge evolutionarily to similar mutation rates. Although ample discussion exists on the relationship between rates of mutation and adaptation, the theory has been mostly focused on constant environments with a simple fitness landscape, neglecting the likely complexity of more natural environments. The limited empirical evidence in asexual populations suggests that adaptation rates are a concave-down function of the mutation rate (Arjan et al. 1999; Desai et al. 2007; Sprouffske et al. 2018). However, the experimental environments in these studies were generally simple and homogeneous (Arjan et al. 1999; Desai et al. 2007; Sprouffske et al. 2018), or the populations of interest were already well-adapted to the experimental environment

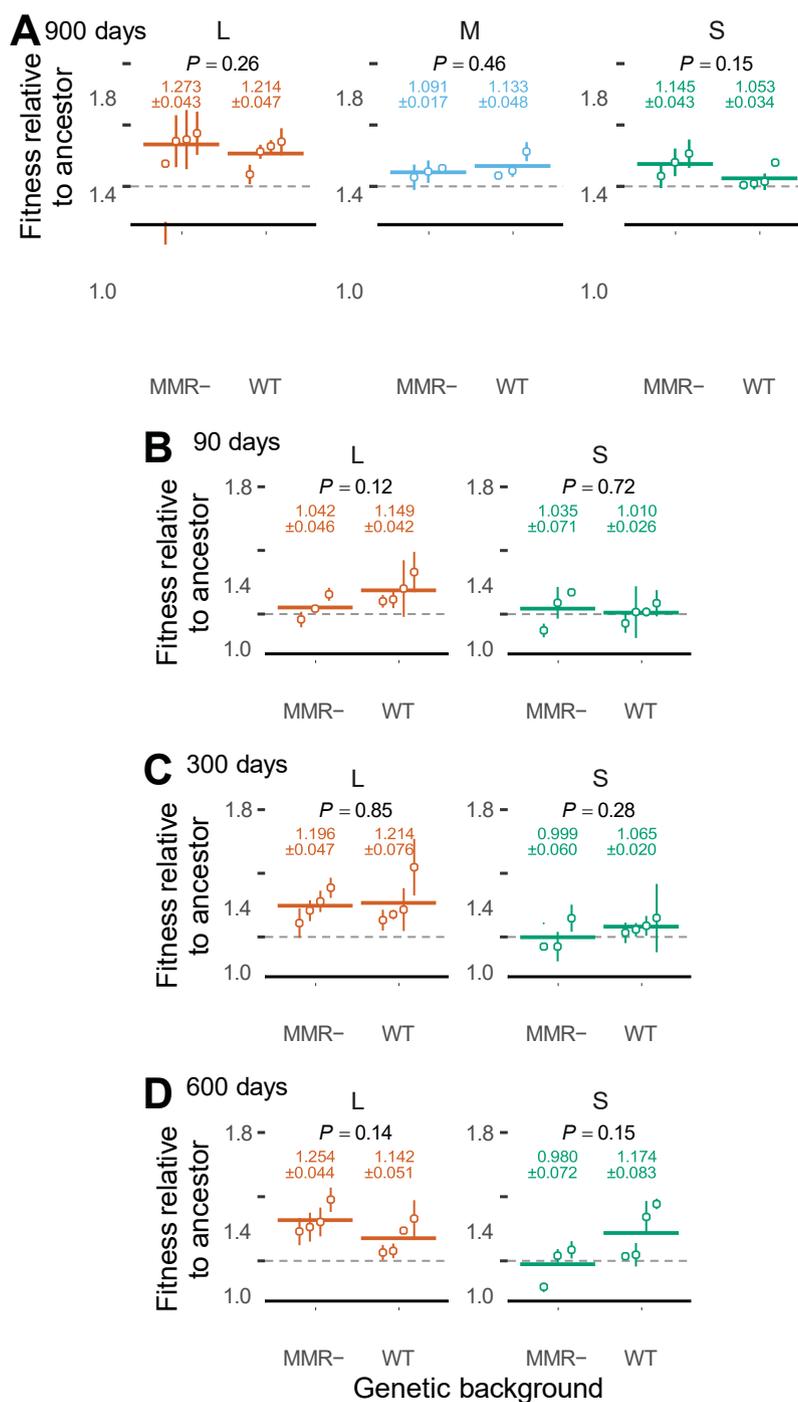
(McDonald et al. 2012). Thus, it remains to be determined whether the relationship between rates of mutation and adaptation in richer and more complex environments follows the same patterns as observed in simpler settings.

To study how the mutation rate affects adaptation in a more complex setting, we investigated the long-term evolutionary changes of *Escherichia coli* grown in culture tubes containing a complex medium, Luria-Bertani (LB) broth, comprised a nutritionally rich mixture of multiple amino-acid based carbon sources (Sezonov et al. 2007). In contrast to evolution in flasks containing glucose-limited media, such environments can facilitate the rapid emergence of stable subpopulations and clonal divergence based on spatial niche differentiation and amino-acid metabolism divergence (Behringer et al. 2018). To vary the mutation rate, we evolved both WT populations (methyl-directed mismatch repair [MMR]<sup>+</sup>) and hypermutator populations with an impaired MMR pathway (MMR<sup>-</sup>, obtained by *mutL* knockout), for which the single-nucleotide mutation rate is 140-fold higher than that for the WT genetic background (Lee et al. 2012). As different population-genetic environments can alter the fixation probability of mutations and the proportion of effectively beneficial or deleterious mutations (Wahl et al. 2002), and different demographic settings have been found to affect the results of experimental evolution (Vogwill et al. 2016; Wein and Dagan 2019), for both WT and MMR<sup>-</sup> populations, we performed three different transfer sizes in daily transfers: 1/10 (large, L), 1/10<sup>4</sup> (medium, M), and 1/10<sup>7</sup> (small, S) to explore the generality of our experimental results. Here, we examine the differences in phenotypic and molecular evolution among these populations over the course of 900 days.

## Results

### Higher Initial Mutation Rates Do Not Lead to Faster Rates of Fitness Improvement

When batch cultured, *E. coli* commonly adapt to their experimental environments and show fitness improvement compared with their ancestors (Van den Bergh et al. 2018; McDonald 2019). To study fitness improvement in populations originating from genetic backgrounds with different initial mutation rates (MMR<sup>-</sup> and WT), we performed the experimental evolution in three different transfer dilution bottlenecks (L: 1/10, M: 1/10<sup>4</sup>, and S: 1/10<sup>7</sup>) for 900 days (supplementary fig. S1, Supplementary Material online). For each of six genetic background/transfer-size combinations (2 × 3), eight replicates were maintained. After 900 days of experimental evolution, we then performed head-to-head competition assays between evolved populations and their corresponding ancestors using four replicates. Across all 21 populations with data available (three were aborted; see Materials and Methods), mean fitness significantly increased relative to the time-zero ancestor, by a ratio of 1.14 (standard



**FIG. 1.**—Fitness improvement during experimental evolution. For evolved populations under different transfer sizes (orange for L,  $1/10^4$ ; or blue for M,  $1/10^7$ ; or green for S,  $1/10^7$ ) and different genetic backgrounds (MMR- or WT), mean fitnesses relative to the ancestor at (A) Day 900, (B) Day 90, (C) Day 300, or (D) Day 600 are reported. Each open circle represents an estimated mean for an evolved population with at least three independent competition assays. The error bars represent SEs. Gray dashed lines denote the null expectation under no improvement from ancestral fitness. The means across all evolved lines for a combination of transfer size and genetic background are represented by colored horizontal lines; the numeric values of means and SEs are printed on the tops of each set of data. The  $P$  values for nested ANOVA are also shown on the top.

error, or SE  $\frac{1}{4}$  0.019;  $P \frac{1}{4}$   $5.8 \times 10^{-7}$ , two-tailed  $t$ -test), indicating that the evolution of these populations was shaped by adaptive processes.

The amount of fitness improvement of MMR- populations was not significantly different from that for WT populations at any of the transfer sizes (fig. 1A; L:  $P \frac{1}{4}$  0.26; M:  $P \frac{1}{4}$  0.46; S:

$P \approx 0.15$ , nested ANOVA). In addition, no transfer sizes produced a ratio of mean fitness improvement significantly different from 1.0. For example, for the L transfer size, the mean

fitness improvement for MMR<sup>-</sup> and WT backgrounds are respectively 0.27 (SE  $\approx 0.043$ ) and 0.21 (SE  $\approx 0.047$ ), and therefore the ratio (MMR<sup>-</sup> : WT) is 1.28 (SE  $\approx 0.35$ ). Similarly, the ratios are 0.69 (SE  $\approx 0.28$ ) and 2.76 (SE  $\approx 1.98$ ) in the M and S transfer sizes, respectively (fig. 1A). Thus, starting evolution as a hypermutator does not necessarily translate into a faster fitness-improvement rate.

Previous studies of *E. coli* in simpler, more homogeneous environments have shown that the most rapid increases in

population fitness typically occur within 2,500 generations, after which the rate of adaptation significantly slows (Barrick et al. 2009). At 900 days, the L populations, which due to their large transfer size (1:10) experienced the least number of cell divisions ( $\log_2 10$  ∷ 3.3 generations per day), had experienced ~3,000 generations, whereas the S populations had experienced ~21,000 generations ( $\log_2 10^7$  ∷ 23.3 generations per day). As such, the absence of significant differences in the cumulative amount of adaptation over this period—despite large differences in initial mutation rate—might be a consequence of both genetic backgrounds having exited an initial period of rapid fitness evolution. Thus, to better survey any temporal heterogeneity in the rate of adaptation, we further assessed fitness after 90, 300, and 600 days of evolution in response to L and S transfer sizes. The results of nested ANOVA again demonstrate a lack of evidence for an increase in initial mutation rates leading to an increase in the amount of fitness improvement (fig. 1B–D). The results are not qualitatively changed even when the natural-logarithmic transformed fitness is used in the analysis (supplementary fig. S2, Supplementary Material online). Thus, high mutation rates did not result in accelerated fitness improvement in these asexual populations even in the early stages of adaptation.

### Mutation Rates Evolve to Be More Similar Throughout Experimental Evolution

The indifference of adaptation rates to initial mutation rates might be explained if the mutation rates of hypermutator and WT populations became more similar during the evolution experiment, either due to a reduction in the mutation rate in initially hypermutating populations or to an increase of the rate in WT populations. To test this possibility, we performed fluctuation tests, which indirectly measure mutation rates at a resistance locus (Foster 2006), on different clones isolated from evolved populations after 900 days. Although the ratio of rifampicin-resistance mutation rates for the two ancestral lines (MMR–: WT) was 250 (supplementary fig. S3, Supplementary Material online), after 900 days of evolution, the mean difference in mutation rates between MMR– and WT backgrounds greatly decreased across all transfer sizes (fig. 2). For example, in the L transfer size, the mean mutation rate is  $5.0 \times 10^{-7}$  ( $SE \frac{1}{4} 2.7 \times 10^{-7}$ ) and  $1.6 \times 10^{-8}$  ( $SE \frac{1}{4} 1.2 \times 10^{-8}$ ) for MMR– and WT evolved populations, respectively. Therefore, the ratio of mean mutation rates (MMR–: WT) was reduced to 32 ( $SE \frac{1}{4} 29$ ). However, this reduction seems mostly a consequence of the occasional emergence of higher mutation rates in the WT background, as only one population (115) among four tested WT/L populations shows a significant increase in the mutation rate for both tested clones. For the M and S transfer sizes, the ratios of mean mutation rates (MMR–: WT) were also reduced to 32 ( $SE \frac{1}{4} 15$ ) and 12 ( $SE \frac{1}{4} 3.9$ ), respectively. However, here the repeated evolution of lower mutation rates in clones isolated

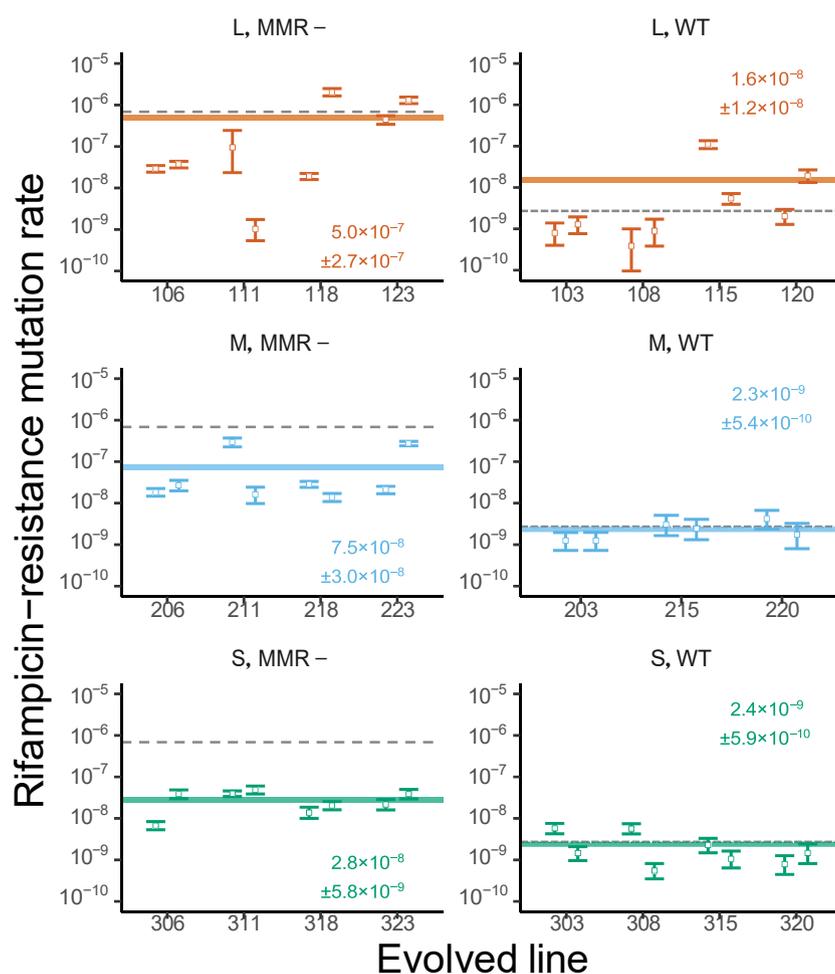
from the MMR– background plays a more important role in the reduction, as all of the tested clones from MMR–/M and MMR–/S populations show a significant mutation rate decrease. Thus, the evolutionary convergence of mutation rates likely contributes to the difference in adaptation rates being less than what might be expected based on initial differences in mutation rates. Interestingly, these observations demonstrate how transfer schemes can affect the evolutionary dynamics of mutation rates in asexual populations.

### Genome Evolution Rates Are Less Different Than Predicted by Initial Mutation Rates

To enhance our understanding of how the tempo and mode of genomic evolution relate to fitness and phenotypic evolution, we performed metapopulation sequencing of each experimental population roughly every 100 days to acquire mutation profiles of derived allele frequencies (DAFs; supplementary fig. S4, Supplementary Material online). For each combination of genetic background and transfer size, we estimated the rate of genomic evolution by regressing the number of mutations per clone (i.e., the sum of DAFs of all observed SNPs) of all populations against the number of generations at each sampled time point. Because a quadratic regression model did not significantly outperform the linear regression ( $P \frac{1}{4} 0.028$  for MMR–/L;  $P > 0.10$  for the others, nested ANOVA), we will focus on the results of the linear regression below.

As with the rate of adaptation, for all three transfer sizes, the ratio of the rate of genomic evolution between the two backgrounds (MMR– versus WT) was much smaller than the initial difference in mutation rates (fig. 3A). For example, under the L transfer size, the rate of genomic evolution is 115 ( $SE \frac{1}{4} 5.3$ ) and 30 ( $SE \frac{1}{4} 4.3$ ) mutations per clone per 1,000 generations for MMR– and WT populations, respectively. Therefore, the ratio of the genomic-evolution rates is only 3.8 ( $SE \frac{1}{4} 0.58$ ). Similarly, in the M and S transfer sizes, the ratios are 20 ( $SE \frac{1}{4} 1.3$ ) and 23 ( $SE \frac{1}{4} 1.6$ ), respectively. This observation remains qualitatively similar even when different kinds of measurements for genomic divergence are used, for example, the number of detected mutations (supplementary fig. S5A, Supplementary Material online), nucleotide diversity (supplementary fig. S5B, Supplementary Material online), or when genomic divergence was estimated only by synonymous SNPs (supplementary fig. S6, Supplementary Material online).

To further survey how genomic evolutionary rates vary across experimental populations and to determine if the mean rates accurately represent the majority of experimental populations, we separately measured the rate of genomic evolution for each experimental population (supplementary fig. S7, Supplementary Material online). This revealed that the distribution of rates in the WT populations under the L transfer size is wider than the distributions of rates in all other



**FIG. 2.**—Evolution of mutation rates after 900 days of experimental evolution. Each panel shows mutation rates of evolved populations in a combination of transfer size (orange for L, 1/10; blue for M, 1/10<sup>4</sup>; or green for S, 1/10<sup>7</sup>) and genetic background (MMR- or WT). In each combination, three or four evolved populations were tested. Two clones per evolved population were isolated and measured. An open circle and an error bar represent the mean and the 95% confidence interval for a clone. The gray dashed line represents the mutation-rate estimate of the corresponding ancestor. Each colored horizontal line represents a mean mutation-rate measurement of a combination; the value of mean and the SE are also printed.

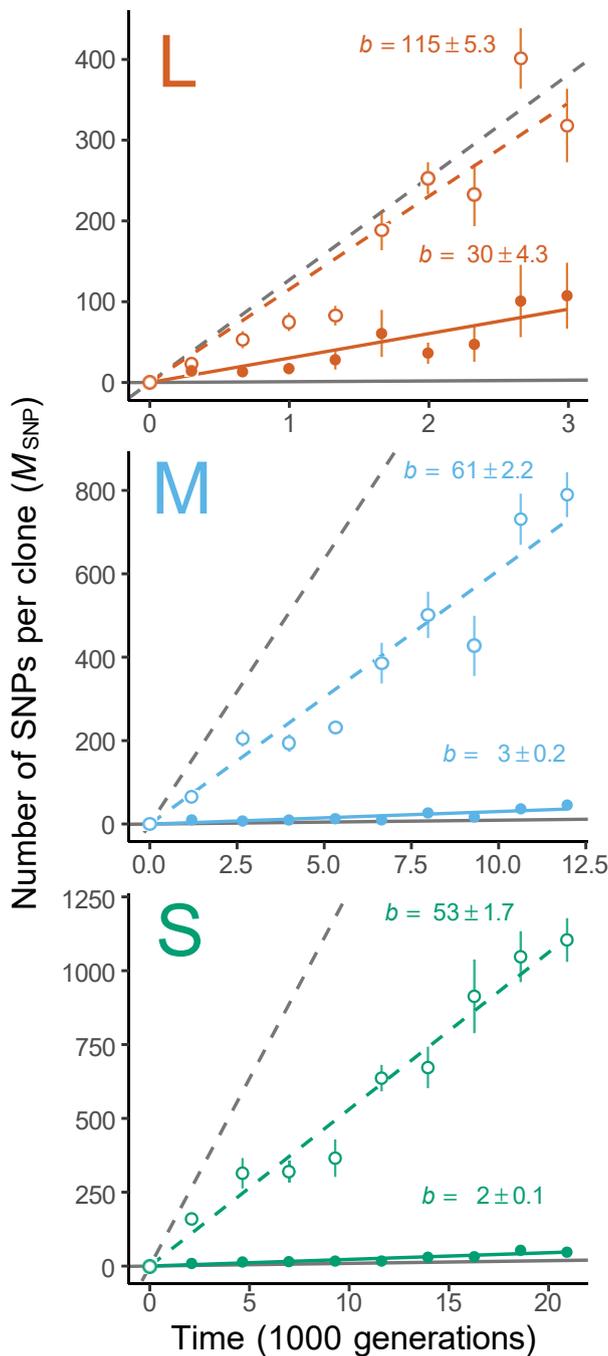
combinations of genetic background and transfer sizes. Specifically, in the L transfer size, although five of the eight WT populations exhibit a genomic evolutionary rate of ~10 mutations per clone per 1,000 generations, one WT population (115) has a rate about ~2x higher, and two WT populations (101 and 113) have a rate close to 100, ~10x higher and similar to MMR- populations. Consistent with the fluctuation-test results noted above, these results suggest that some, but not all, WT populations under the L transfer size evolved a higher mutation rate (see Discussion).

### Mutations Arising in Hypermutators Are More Likely to Be Fixed

Using longitudinal metagenomics-sequencing data allows one to observe evolutionary dynamics at the level of single

mutations and thus better understand the entire adaptive process. Here, we focus on fixed mutations because they are more likely to contribute to adaptation than polymorphic mutations or other mutations that are transient in a population. Because our experimental-environment setting facilitates the development of subpopulation structure, we applied clade-aware hidden Markov chain (caHMM) analysis. Assuming coexistence of two clades (major and minor) in the population, caHMM considers each mutation's DAFs found in different sequencing time-points and then infers which clade each mutation belongs to, whether each mutation reaches within-clade fixation, and when a fixed mutation reaches fixation (Good et al. 2017).

One characteristic parameter in an adaptive process is the fraction of mutations that reached fixation in a population. Using the observed numbers of both detected and fixed



**Fig. 3.**—Rate of genomic evolution in evolved populations. Each panel shows the results of evolved populations in two genetic backgrounds (MMR<sup>-</sup> and WT) in a transfer size (L, M, or S). Each dot shows a mean number of SNPs per clone for MMR<sup>-</sup> (open circles) or WT (closed circles) populations at a sequencing time point. The error bars represent the associated standard errors (some are invisible because smaller than the height of points). The colored dashed and solid lines are linear regressions against the time for MMR<sup>-</sup> and WT populations, respectively. The estimated slope ( $b$ ) and associated standard error are also printed for each regression line. The gray dashed and solid lines represent how evolved populations accumulate mutations under neutral expectation with the initial mutation rates of MMR<sup>-</sup> and WT ancestors.

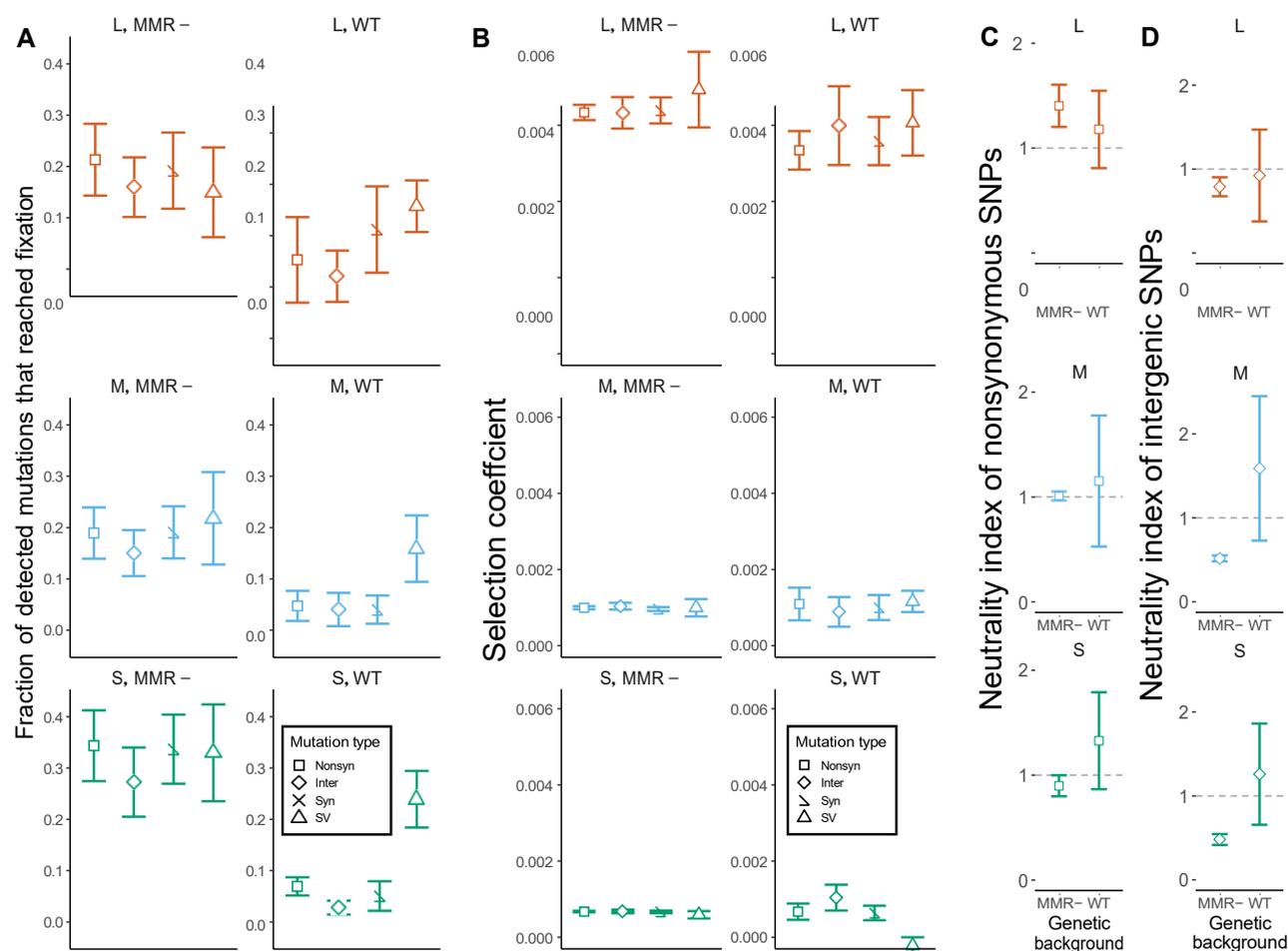
mutations, we first tested whether hypermutator populations have a different fraction of mutation that reached fixation. Different mutation types have a different potential to impact populational fitness. For example, compared with synonymous SNPs, nonsynonymous SNPs have a greater potential to change protein functions; intergenic SNPs have a greater potential to change protein expression; and structural variations (SVs; including indels and mobile-element insertions) have a greater potential to disrupt a protein. Therefore, we performed separate tests on the conditional fractions of these four functional categories of detected mutations that reached fixation.

Although not always significant, the fixation fraction in the MMR<sup>-</sup> populations is generally higher than in WT populations across different transfer sizes and different categories of mutations (fig. 4A). The fixation fraction is similar across different categories of mutations, regardless of their perceived potential to affect fitness, suggesting that the fixation of most mutations is a consequence of genetic hitchhiking as opposed to intrinsic beneficial effects.

Another critical factor determining the temporal dynamics of a mutation is the underlying fitness effect. Using the temporal data of allele frequencies for a mutation, we can quantify the net selection coefficient, which reflects an allele's own fitness effects but can be potentially influenced by the effects of linked mutations. We did not find a significant difference between the selection coefficients of fixed mutations in the two genetic backgrounds (fig. 4B). More importantly, we also found that different categories of mutations show similar mean selection coefficient estimates, which again suggests a pivotal contribution of hitchhiking effects to the mutational dynamics of genome evolution in asexual populations. Given this observation, we assumed that the fixation of mutations that occurred in the same clade at the same time belong to one selective sweep. After calculating the mean selection coefficients per selective sweep, we again found no significant differences of selection coefficients between two genetic backgrounds (supplementary fig. S8, Supplementary Material online).

### Neutrality Tests Reveal Partial Evidence for Positive Selection

In theory, comparing the number of fixed mutations in functional categories of sites with different potential effects on fitness can summarize general patterns in the mode of genome evolution (McDonald and Kreitman 1991; Rand and Kann 1996). If a population has experienced strong positive selection on protein function or expression, it is expected that there will be more fixed mutations with a greater potential to change protein function or expression than mutations with smaller such potential. If a population has experienced purifying selection on protein function or expression, it is expected that there will be fewer fixed mutations with a potential to



**FIG. 4.**—Analysis of strength of natural selection associated with fixed mutations in different categories in different treatments of experimental evolution. (A) Each symbol shows the population mean fraction of detected nonsynonymous mutations (squares), intergenic mutations (diamonds), synonymous mutations (crosses), or SV mutations (triangles) that reached within-clade fixation in each combination of transfer size (L, M, or S) and genetic background (MMR- or WT). The error bars show the 95% confidence intervals. (B) Each symbol shows the mean selection coefficient of nonsynonymous mutations (squares), intergenic mutations (diamonds), synonymous mutations (crosses), or SV (triangles) that are fixed in any population belonging in a combination of transfer size and genetic background. The error bars show the 95% confidence intervals. (C) Each square shows the population mean neutrality index of nonsynonymous mutations for a combination of transfer size and genetic background. The error bars show the 95% confidence intervals. The horizontal gray lines denote the point of neutrality (1.0). (D) Each square shows the population mean neutrality index of intergenic mutations for a combination of transfer size and genetic background. The error bars show the 95% confidence intervals. The horizontal gray lines denote the point of neutrality (1.0).

change protein function/expression than mutations with smaller such potential. Therefore, the ratio of the number of fixed nonsynonymous mutations ( $F_N$ ) to the number of fixed synonymous mutations ( $F_S$ ) or the ratio of the number of fixed intergenic SNPs ( $F_I$ ) to  $F_S$  is predicted to be elevated by strong positive selection. Similarly,  $F_N/F_S$  or  $F_I/F_S$  are predicted to be small with a strong purifying selection. However, these ratios are not necessarily directly comparable in MMR- and WT backgrounds because the two genetic backgrounds have varied mutational spectra and different relative rates of nonsynonymous and synonymous mutations (Lee et al. 2012).

To address this issue, we normalized the observed  $F_N/F_S$  by the ratio of nonsynonymous and synonymous mutations ( $U_N/$

$U_S$ ) previously observed in a mutation accumulation experiment that utilized our exact ancestral genotypes (Lee et al.

2012). The ratio of these ratios ( $F_N/F_S)/(U_N/U_S)$ , will be referred to as the neutrality index of nonsynonymous SNPs, because ratios significantly  $>1.0$  or  $<1.0$  imply, respectively, a predominance of positive selection in driving mutation fixation or a predominance of purifying selection driving mutation extinction. This index is analogous to Tachida's index for neutrality (Tachida 2000), but with a different normalizing approach—via results from mutation-accumulation experiment instead of polymorphism, and is a preferred approach as it eliminates any potential problems with selection on silent sites. This definition of the neutrality index can be applied to any kind of mutation. For example, we also define the neutrality index of intergenic SNPs as  $(F_i/F_s)/(U_i/U_s)$ , where  $U_i$  is the number of intergenic mutations observed in mutation accumulation (Lee et al. 2012).

The neutrality index of nonsynonymous SNPs in MMR–populations under the L transfer size is significantly larger than one, consistent with strong positive selection (fig. 4C). Moreover, the neutrality index of intergenic SNPs in MMR–populations under all transfer sizes is significantly smaller than one, suggesting overall strong purifying selection on intergenic SNPs (fig. 4D). On the contrary, the neutrality index in WT populations never significantly differs from 1.0. Although this may indicate a weaker strength of selection in WT populations, it may also be the result of insufficient statistical power due to an overall smaller number of fixed SNPs in WT populations, evident by their large confidence intervals associated with the point estimation.

### Parallel Evolution of Fixed Mutations at the Genic and Nucleotide Level

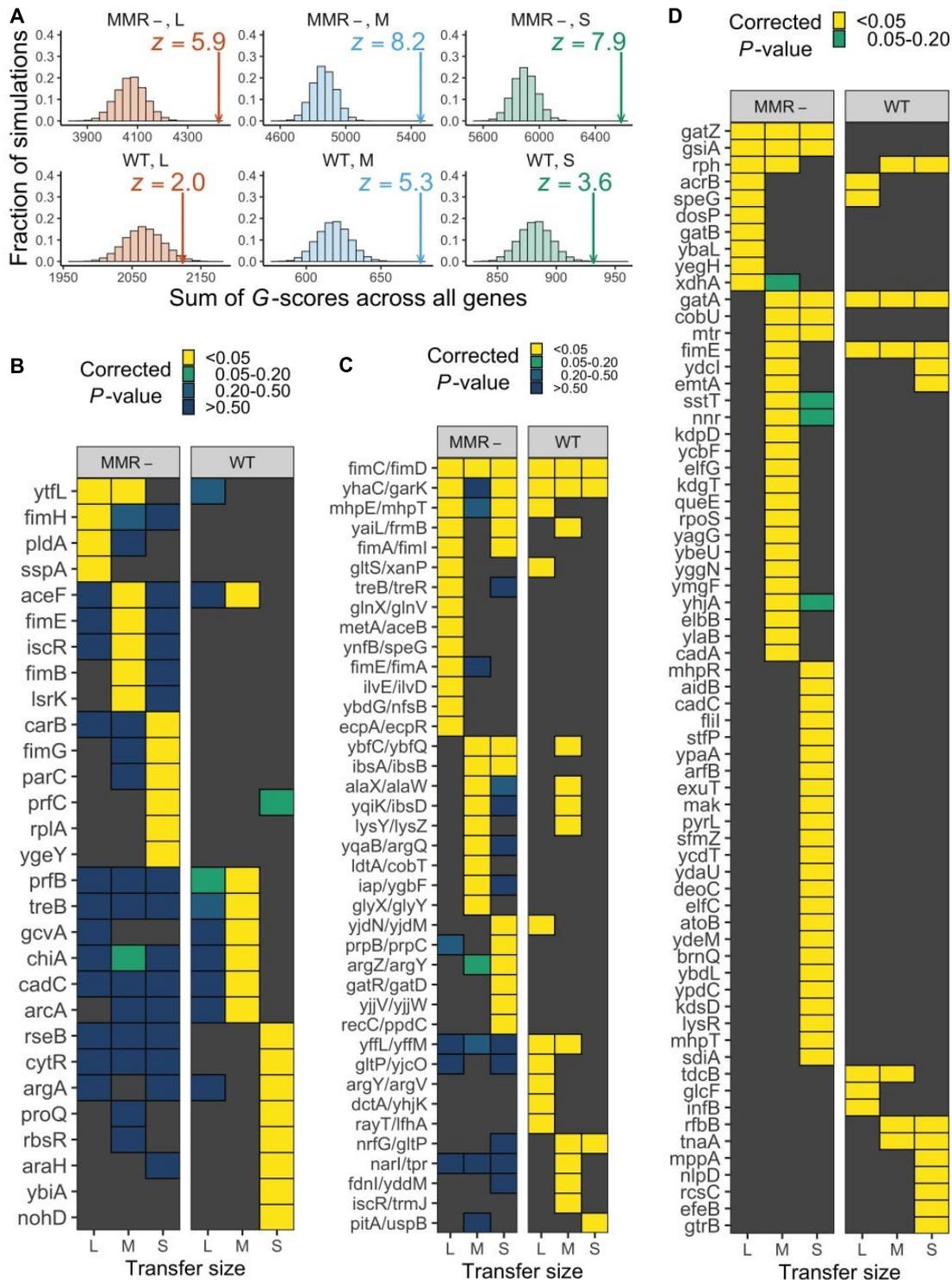
As the previous analysis suggests that the statistical power of the neutrality index may be insufficient, we also considered parallelism of fixed mutations to examine the action of positive selection using two different metrics. First, we utilized a sum of *G*-scores to measure the excess parallelism across fixed nonsynonymous mutations relative to the expectation based on gene lengths (Tenailon et al. 2016). For each gene, a higher *G*-score implies that the gene is more enriched for fixed nonsynonymous mutations than expected by gene length. Summing the *G*-scores of all genes for each combination of genetic background and transfer size revealed statistical significance in all cases (*z*-test, fig. 5A). Second, we calculated the mean Bray–Curtis similarity of the number of fixed nonsynonymous mutations across all genes (Turner et al. 2018) for all pairwise comparisons of evolved populations in each combination of genetic background and transfer size, again revealing statistically significant similarity in all cases (*z*-test, supplementary fig. S9, Supplementary Material online). Therefore, both findings suggest that positive selection has shaped the genomic evolution of populations in all genetic-background/transfer-size combinations.

The analysis of parallelism also helps reveal which mutations are most likely to be drivers of adaptation. Using the same *G*-score analysis, in each genetic-background/transfer-size combination, we identified genes that were overrepresented for fixed nonsynonymous mutations ( $P < 0.05$ , Bonferroni correction; fig. 5B). Gene Ontology (GO) analysis on these gene subsets revealed that significantly enriched GO terms were often related to transcription regulation and biofilm formation (supplementary table S1, Supplementary Material online). Using similar methods, we also identified subsets of genes enriched for fixed mutations in intergenic regions (fig. 5C) and for structural mutations, including indels and IS-element insertions (fig. 5D). Together, the genes in these lists serve as good candidates for revealing the various mechanisms important to adaptation in complex environments (see Discussion).

In addition to genic-level parallelism, we also identified nucleotide-level parallelism. In particular, 197 cases of parallel fixed nonsynonymous mutations were identified in at least two experimental populations within the same combination of transfer size and genetic background (supplementary table S2, Supplementary Material online). Because the probability that fixed mutations would occur by chance at a given nonsynonymous site in at least two experimental populations is very low, the observed parallelism at the nucleotide-level again suggests positive selection and identifies important candidates for studying the molecular mechanisms associated with adaptation in complex media. For example, three instances of parallel mutation within *fimH* were located in its mannose-binding domain and therefore may be good candidates for future functional studies of cell adhesion (Schembri et al. 2001). Furthermore, three instances of parallel-fixed nonsynonymous mutations were found in genes with GO terms associated DNA repair or DNA replication, including *dnaE* (DNA pol III subunit *α*), *yajL* (protein/nucleic acid deglycase 3), and *nrdA* (ribonucleoside-diphosphate reductase 1, *α* subunit dimer). As all three of these cases arose in M and S transfer sizes with MMR– backgrounds, they may be good candidates for studying the molecular mechanisms associated with lowering mutation rates.

### Discussion

Here, we describe the effect of the initial mutation rate on the rate of fitness improvement and genomic evolution by experimentally evolving *E. coli* with two distinct mutation-rate backgrounds. Although the initial difference in mutation rates of these two genetic backgrounds is >100x, after 900 days of evolution, the differences in the net rates of fitness improvement (0.7–2.7 fold), in the mutation rates (12- to 32-fold), and in the level of genome evolution (4- to 23-fold) are much smaller. These results suggest that elevating the mutation rates in asexual populations does not proportionally increase the rate of fitness improvement, possibly due to clonal interference, linkage to deleterious mutations, and/or evolution of lower mutation rates. In addition, we found that mutations arising in MMR– populations exhibit a higher fraction of fixation than mutations in WT populations, whereas mutations in categories with different fitness effects show similar fixation probabilities within each genetic-background treatment, suggesting the strong influence of hitchhiking effects in genome evolution in these populations. Despite the strong linkage, a high degree of parallelism in mutations arising in particular genes and nucleotides still suggests positive selection shaping genome evolution in all transfer-size/genetic-background combinations. The observed mutations with high parallelism serve as excellent candidates for understanding the mechanisms of adaptation to the complex conditions provided by the experimental environment.



**Fig. 5.**—Evolutionary parallelism as evidence for positive selection. (A) The vertical arrow shows the observed sum of G-scores, representing the extent of parallel mutation for a particular combination of transfer size and genetic background. The histogram shows the distribution of 20,000 simulated sums of G-scores, representing the null distribution of evolutionary parallelism. The significance of the observed sums can be evaluated by z-scores ( $z > 1.65$  for one-tailed  $P < 0.05$ ). (B) List of genes with fixed nonsynonymous mutations. Significance levels (simulated P values with Bonferroni correction) are shown by the different nonblack colors of tiles. Genes with no such hits in a particular combination are shown by black tiles. (C) List of genes with fixed intergenic mutations. Significance levels (simulated P values with Bonferroni correction) are shown by the different nonblack colors of tiles. Intergenic regions with no such hits in a particular combination are shown by black tiles. (D) List of genes significantly overrepresented for structural mutations that are likely under positive selection. Yellow tiles highlight genes observed in at least two populations and yielding simulated P values  $< 0.05$  after Bonferroni correction.

Downloaded from <https://academic.oup.com/gbe/article/13/12/evab257/6432043> by Arizona State University user on 08 June 2022

### Effects of High Mutation Rates on Evolution

Our experiments reveal that, even over 900 days of evolution in a complex medium, hypermutating *E. coli* does not necessarily exhibit a faster rate of adaptation than wild-type *E. coli*. This observation is consistent with previous experimental-evolution results over a shorter period and in simpler media, such as glucose medium DM25 for 1,000 generations (Arjan et al. 1999) and glucose medium DM1,000 for 3,000 generations (Sprouffske et al. 2018). Our experiments further demonstrate that populations with initial hypermutator backgrounds can rapidly evolve lower mutation rates. Together with other empirical work on prokaryotic (Sprouffske et al. 2018) and eukaryotic hypermutators (McDonald et al. 2012), these results suggest that strong genetic load due to deleterious mutations remains a pivotal factor in the evolution of mutation rates, consistent with the drift-barrier hypothesis (Lynch 2010; Sung et al. 2012). Beneficial mutations may also play a role in the evolution of hypermutators, suggested by our finding that the evolution of higher mutation rates tends to be found in populations with the L transfer size. Even when some new hypermutator alleles can spread in a population by linkage with other beneficial mutations during the adaptive process and thus briefly improve the rate of adaptation (Sniegowski et al. 1997; Tenailon et al. 1999, 2001), such events are usually transitory (Giraud et al. 2001; Desai and Fisher 2007; Wielgoss et al. 2013). Moreover, the fact that lowering the mutation rate could not have involved a reversion of deleted MMR in our experiments implies that there is excess capacity for improving replication fidelity through other parts of *E. coli* genome. As mutation-rate evolution occurred over a relatively short period in our study, the results bear on several critical questions for future studies, including the rapidity of the dynamics of the evolution of mutation rates and the consequences for the mutational spectrum. Another interesting question on the relationship between genetic backgrounds and evolutionary outcomes is whether different genetic backgrounds alter the set of beneficial mutations in the fitness landscape. As there is significant parallelism between WT and MMR- populations in each of the three transfer sizes (supplementary fig. S10, Supplementary Material online), the differences in their fitness landscapes are likely limited.

### Effect of Genetic Linkage on Evolution

In asexual populations with reduced recombination, the fate of a mutation is largely affected by its association with other mutations due to strong genetic linkage (Gillespie 2000; Neher 2013; Couce et al. 2017). Extensive hitchhiking is a feature of our evolving *E. coli* populations, as we observe similar fractions of fixed mutations and net associated fitness effects for fixed mutations across different functional categories. To better illustrate this effect of genetic linkage, we showed that the temporal DAF changes of any two SNPs in

the same genome are highly correlated (i.e., a rightly skewed distribution of correlation coefficients) and result in more largely positive correlation coefficients than for a random expectation of nonlinked mutations (supplementary fig. S11, Supplementary Material online). Therefore, even though MMR- populations show a range of genome-evolution rates 4–23x higher than WT populations, the excess of fixed mutations does not directly contribute to adaptation rates. Rather, with no recombination in the experimental populations, the adaptation rates are heavily affected by clonal interference and linkage to deleterious mutations (Schiffels et al. 2011). In the future, it will be interesting to measure the fitness of single clones within a population to further reveal the effect of clonal interference. By sequencing single clones and measuring fitness effects of single mutations, it will be possible to separate the driver mutations from hitchhiking mutations and to dissect the evolutionary dynamics under linked selection.

### Effects of Transfer Sizes on Evolution

The three different transfer sizes (L, M, and S) implemented in our evolution experiment allow us to compare adaptive processes in different population-genetic environments. In theory, when the transfer size is large, there should be more effectively beneficial mutations. Consistent with this theoretical prediction, the populations cultured with the L-transfer size show the highest rates of fitness gain (fig. 1). The fitness gain may be due to the improvement of protein functions, as the observed neutrality index of nonsynonymous mutations in MMR- populations under L transfer size is  $>1.0$  and the highest (fig. 4C). On the contrary, the populations cultured with the S-transfer size may have accumulated more deleterious mutations and thus show less fitness improvement.

We also found that the populations under the L-transfer size tend to evolve higher mutation rates (fig. 2), as hypermutator alleles may hitchhike with more selective sweeps of beneficial mutations. Previous research has also demonstrated that the spread of hypermutator alleles tends to be found in populations with larger sizes or weaker bottleneck effects (Raynes et al. 2014). As a result, even though the mutation rates of populations in two genetic backgrounds (WT and MMR-) evolved to be closer to each other during the experiments with all three transfer sizes, the driving mechanisms are different between populations under the L transfer size and those under M or S transfer size (fig. 2).

We note that an L-size transfer requires fewer cell divisions to reach the stationary phase than an M-size or S-size transfer. Therefore, populations under the L transfer size may experience a longer stationary phase during experimental evolution. Considering different staying times in the stationary phase potentially is critical for interpreting many results across transfer-size treatments. For example, when calculating the rate of genomic evolution (fig. 3), we estimated generation numbers by the required number of cell divisions in the

exponential growth phase, assuming no cell death during the stationary phase. Such an assumption can overestimate the rate of genomic evolution, and the overestimation in L populations can be the strongest due to the longest stationary phase.

Given that the physiological features of *E. coli* can be different in exponential and stationary phases (Pletnev et al. 2015), the evolutionary pressure under different transfer sizes may also vary as their relative durations within the two phases differ. A population can adapt by improving its maximum growth rate in the exponential phase, increasing its carrying capacity, and decreasing its death rate during the stationary phase, or a combination of both. The fitness effects combined among different growth phases were focused in our fitness assays. In our case, however, the adaptive mechanisms of the populations in different transfer sizes are probably similar, as we still found significant parallelism in the fixed enriched mutations across different transfer sizes (supplementary fig. S12, Supplementary Material online).

This issue of different durations in different growth phases should not directly affect our main conclusion regarding evolved fitness under different genetic backgrounds in the same transfer size. However, in our competition assays, the transferred ratio at inoculation (1/100) was not the same as the daily transferred ratios applied to the experimentally evolved populations, especially for S populations (1/10<sup>7</sup>). Because the relative durations in different growth phases in competition assay can be different from the ones in experimental evolution, the measured fitness may not reflect the genuine fitness improvement during the experimental evolution. To explore whether this problem has limited our abilities to identify the fitness differences in different genetic backgrounds, we performed another set of competition assays for S/MMR– and S/WT populations using 1/10<sup>7</sup> as the transferred ratio at inoculation. The results show no significant differences between the evolved populations in two genetic backgrounds (supplementary fig. S13, Supplementary Material online), similar to the previous results (fig. 1). Therefore, our conclusion, that there is no strong relationship between hypermutators and the overall rate of fitness improvement, remains unchanged, although how much *E. coli* evolved specifically in different growth phases is worth studying in the future.

### The Possible Role of Biofilm Formation in Adaptation

With respect to the specific genes and biological processes that appear to be targeted for adaptation, our analysis of candidate genes suggests that biofilm formation is an important characteristic in adapting to the complex setting imposed by the experimental environment. In particular, the formation of type I fimbriae is critical for biofilm formation in *E. coli* (Pratt and Kolter 1998). Consistently, several genes enriched for fixed nonsynonymous mutations are related to formation of

type I fimbriae (fig. 5B), including *fimH* and *fimG*, which account for components the type I fimbriae (Waksman and Hultgren 2009; Le Trong et al. 2010), *fimB*, and *fimE*, which regulate the expression of *fimAICDFGH* operon (Olsen et al. 1998), and *proQ* (RNA chaperone) and *IsrK* (autoinducer-2 kinase), which also facilitate biofilm formation (Li et al. 2007; Sheidy and Zielke 2013). We additionally observed an enrichment of mutations in the intergenic region of *fimE/fimA* (fig. 5C) which contains a phase-variable promoter for regulating the expression of the *fimAICDFGH* operon (Abraham et al. 1985; Spears et al. 1986). Lastly, the list of genes enriched with structural mutations (fig. 5D) also includes *fimE*, which primarily turns off the expression of *fimAICDFGH* operon, and several genes related to *gatYZABCD* operon, including *gatZ*, *gatB*, and *gatA*, whose deletions can increase biofilm formation (Domka et al. 2007). Many of these candidate genes related to type I fimbriae contributed to the adaptation in an earlier experimental environment of a similar nature (Behringer et al. 2018). Moreover, structural mutations in *gatZ* and *gatA* genes have been found to contribute to the initial adaptation of *E. coli* in a mouse gut, another example of a complex environment (Barroso-Batista et al. 2014). Studying the genetic variants promoting the evolution of biofilm in complex environments is of particular interest in the field of public health, as the evolution of biofilm has been considered to be related to the evolution of microbial social behaviors (Tarnita 2017), the evolution of pathogenicity (Kaper et al. 2004; Naves et al. 2008; Rossi et al. 2018), and the evolution of antibiotic resistance (Avalos Vizcarra et al. 2016; Sharma et al. 2016). Thus, understanding the evolution of biofilm formation may be key to increasing the efficiency of treatments in patients to combat the fast emergence of antibiotic resistance and pathogenicity.

### Candidate Transcriptional Regulators Involved in Adaptation

Because we have provided fresh media every day during experimental evolution, a fast switch from stationary phase to exponential growth phase may bring benefits to the evolving populations (Monod 1949; Navarro Llorens et al. 2010). Interestingly, several genes enriched in fixed nonsynonymous mutations in our study are transcription regulators, including *arcA*, *cadC*, *cytR*, *rbsR*, *rseB*, and *sspA* (fig. 5B). The genes enriched in structured variations also include transcription regulators, such as *arcB*, *cadC*, *rpoS*, and *nlpD* (fig. 5C). The mutations in these genes may contribute to the transcriptomic reprogramming for the fast switch from stationary phase to exponential phase. For example, *rseB* is a negative regulator of the stationary-phase effector, sigma factor E (Missiakas et al. 1997); and both *cytR* and *rbsR* are repressors to the carbon limitation effector, cAMP-CRP (Bell et al. 1986; Mauzy and Hermodson 1992; Kristensen et al. 1996). Therefore, gain-of-function mutations in these three genes

can theoretically reduce the chance that cells stay in stationary phase. In addition, *arcA*, *arcB*, *cadC*, *rpoS*, and *sspA* are known as stress-responding activators (Iuchi et al. 1989; Lange and Hengge-Aronis 1991; Watson et al. 1992; Williams et al. 1994; Rolfe et al. 2011), suggesting that loss-of-function mutations in these genes are beneficial in the experimental environment involving a frequent supply of fresh media imposing low stress. Further investigation is needed to determine whether any of these mutations do indeed bring such benefits to the experimentally evolved populations.

### Concluding Remarks

To sum up, our results reveal that high mutation rates in *E. coli* have only a very limited influence on the rate of adaptation. Our findings may provide useful insight for understanding clinically relevant processes involving asexual populations, such as the evolution of improved growth rates in pathogens and the emergence of antibiotic resistance in natural or host environments. Compared with the previous lab experiments in simpler environmental settings, our experimental evolution results in complex media are likely to be more representative of natural or host environments, which are usually highly heterogeneous. For example, a combination of the occasional emergence of hypermutators under weaker bottlenecks and the consistent evolution of antimutators under stronger bottlenecks may explain why a low to intermediate frequency of hypermutators is usually found in pathogen populations (Couce et al. 2016; Veschetti et al. 2020). Whether the elevated mutation rates affect adaptation rates and the pattern of genome evolution in these populations under natural or host environments merits further research. For example, hypermutators may be critical for a founding population in a new environment, especially with epistasis in the fitness landscape, but such an effect of hypermutation can diminish after in subsequent generations (Mehta et al. 2019).

## Materials and Methods

### Strains

The ancestral strains used in experimental evolution are descendants of PMF2 and PMF5, provided by the Foster Lab (Lee et al. 2012). PMF2 is a prototrophic derivative of *E. coli* K-12 str. MG1655, and its genetic background is called WT in the article. PMF5 is derived from PMF2 with *mutL* deletion, providing the MMR- genetic background. For both kinds of strains, a 3513 bp deletion to the *araBAD* operon is further introduced by lambda red recombineering as a neutral marker. Plates with TA agar (1% arabinose, 1% tryptone, 0.5% NaCl, 0.1% yeast extract, 0.005% TTC [Sigma T8877]) are used for examining the deletion of *araBAD* operon. The colonies with an *araBAD* deletion (*ara-*) appear to be pink; otherwise, the colonies (*araB*) appear to be purple.

### Experimental Evolution

When we established the experimental populations, the ancestral strains were first cultivated overnight at 37 °C on LB agar plates, and then their single-isolated progenitor colonies were inoculated in a 16- x 100-mm glass tube with 10 ml of LB-Miller broth (BD Difco). The tubes were cultured at 175 rpm shaking at 37 °C. Every day, cultures are thoroughly vortexed and transferred into a new tube with 10 ml of fresh LB broth. Three different transfer sizes are used: 1 ml (large), 1 ml (medium), and 1 nl (small), which correspond to different dilution factors:  $10^{-1}$ ,  $10^{-4}$ , and  $10^{-7}$ . Specifically, for small-size transfers, we first diluted 1 ml vortexed bacterial liquid with 10 ml fresh LB broth and then mixed 10 ml of this  $1:10^4$  diluted bacterial liquid with another 10 ml fresh LB broth. Initially, for each combination of genetic background and transfer size, we set up eight replicate tubes. For preventing cross-contamination, four replicates are *ara-* and four replicates are *araB*. During the experimental evolution, experimental populations at Day 90, 200, 300, 400, 500, 600, 700, 800, 900 were frozen in -80 °C freezers for analysis.

### Competition Assay

When evaluating the fitness of an evolved population (*ara-*), we performed head-to-head competition assays between evolved populations and the corresponding ancestor (*araB*). We used the WT ancestor and the MMR- ancestor for evolved populations with WT or MMR- genetic backgrounds, respectively, as there may be some slight fitness effects for the *mutL* deletion (fitness  $\frac{1}{4}$  0.97 with SE  $\frac{1}{4}$  0.04 by the competition assay described here,  $n \frac{1}{4}$  5). At the beginning of the competition assay, we inoculated the frozen samples of the evolved population and the corresponding ancestor in two different tubes each of which contains 10 ml of fresh LB broth at 37 °C shaking at 175 rpm overnight. For the competition, we then put both 50 ml aliquot from the evolved population and 50 ml aliquot from the ancestor into a new tube with 10 ml of LB broth at 37 °C shaking at 175 rpm for 24 h. The shaking speed and the temperature used in the competition are the same as what was used in experimental evolution. Immediately after inoculation (Day 0) and after 24-h competition (Day 1), we used plate-counting to determine the colony-forming units (CFU) of both the ancestor and the evolved population. Specifically, we serially diluted 100 ml aliquot in 900 ml phosphate buffered saline. To distinguish the evolved population from the ancestor, we plated serially diluted aliquots on TA agar because *ara-* colonies will appear to be pink, whereas the ancestor (*araB*) will appear to be purple. The TA agar plates were incubated at 37 °C overnight. We then identified the plate with 30–300 total colonies and counted the number of colonies for the evolved population and the ancestor. Based on the colony numbers and the dilution factor during the serial dilution, we then calculated the

CFU of the ancestor before the competition ( $A_0$ ), the CFU of the ancestor after the competition ( $A_1$ ), the CFU of the evolved population before the competition ( $E_0$ ), and the CFU of the evolved population after the competition ( $E_1$ ). The fitness of the evolved population ( $w$ ) relative to the ancestor was then calculated by the following formula:

$$w = \frac{1}{4} \ln \frac{E_1}{E_0} = \ln \frac{A_1}{A_0} = \ln \frac{b}{b_0}$$

which is the ratio of two Malthusian parameters (Lenski et al. 1991).

At Day 0 of the competition assay, we also serially diluted and plated 100  $\mu$ l aliquot of evolved population as a control. For an evolved population, if both purple and pink colonies were found in the countable control plates (30–300 colonies per plate), the data will be discarded from the analysis because we are not sure about the source of the pink colonies in the experimental plates.

### Fluctuation Test and Mutation Rate Estimation

We quantified mutation rates of evolved populations (at Day 900) and ancestors by fluctuation tests (Foster 2006). Briefly, fluctuation tests measure the rate of resistance to the antimicrobial rifampicin which is conferred by mutations to *rpoB*. For each combination of genetic background and transfer size, the four *ara*<sup>-</sup> populations were assayed. For each population, two biological replicates with different starting clones were assayed. For each of the WT or MMR<sup>-</sup> ancestor, we also run replicate experiments in different starting clones. For each clone, 40 replicate experiments were performed. A number of mutants as determined by CFU/ml were converted to an estimated mutation rate and a corresponding 95% confidence interval by the function “newton.LD” function in the R package “rSalvador” (Zheng 2017).

### DNA Isolation and High-Throughput Sequencing

We conducted high resolution population tracking by collecting 1 ml of culture at Day 90, 200, 300, 400, 500, 600, 700, 800, and 900 of experimental evolution. We used the DNeasy UltraClean Microbial Kit (Qiagen 12224; formerly MO BIO UltraClean Microbial DNA Kit) to extract DNA. For library preparation and sequencing, we submitted DNA to either the Hubbard Center for Genomic Analysis at the University of New Hampshire, the Center for Genomics and Bioinformatics at Indiana University, or the CLAS Genomics Facility at Arizona State University for library preparation and sequencing. Library preparation was done by the Nextera DNA Library Preparation Kit (Illumina, FC-121-1030) following an augmented protocol for optimization of reagent use (Baym et al. 2015) before being pooled and sequenced as paired-end reads on an Illumina HiSeq 2500 (UNH) or an Illumina NextSeq 500 (Indiana; ASU). The target depth is 100x.

### Sequencing Analysis

We performed Sequencing analysis on the Mason and Carbonate high-performance computing clusters at Indiana University. The quality control of sequencing reads was performed by Cutadapt version 1.9.1 (Martin 2011), which removes residual adapters and trims low-quality sequences. The qualified sequencing reads were then mapped to the *E. coli* K-12 substr. MG1655 reference genome (NC\_000913.3). All mutations and their DAFs were identified using Breseq version 0.30.2 with the predict-polymorphisms parameter setting (Deatherage and Barrick 2014). Furthermore, several criteria of further quality checks were applied to the samples which we will only include in the following analysis: 1) mean sequencing depths >10; 2) any WT sample identified to contain the 1,830 bp deletion in *mutL* from the PMF5 progenitor strain was discarded; 3) regions lacking sequencing coverage (i.e., depth  $\frac{1}{4}$  0) must be smaller than 5% of the genome; and 4) the sequencing result should reflect the correct genetic background in terms of *ara* markers, including a nonsynonymous SNP at position 66528, an intergenic SNP at position 70289, and a multiple base substitution mutation (SUB) at position 66533. For an *ara*<sup>-</sup> population, we required either of two SNPs showing DAF < 0.2. For an *ara*<sup>+</sup> population, we required either of two SNPs showing DAF > 0.8 or the SUB is detected.

In the end, 396 genomic profiles passed QC and were included in the following analysis (supplementary table S3, Supplementary Material online). In the case of M transfer size under WT background, only six out of eight replicates of evolved populations are left for the following analysis because the other two were potentially contaminated by MMR<sup>-</sup> strains. In the other five combinations of transfer size and genetic background, there are still eight replicates of evolved populations for the following analysis. For every evolved population subject to the following analysis, its sequencing profiles are available in at least seven different time points.

To make sure that we do not use mutations that originated from the starting clone before experimental evolution in the analysis, we discarded any mutations with a DAF  $\frac{1}{4}$  1.0 at one time point for at least 11 experimental populations with the same genetic background from the analysis. Furthermore, the highly repetitive sequences in *rsx* genes are known to cause errors in SNP calling (McCloskey et al. 2018), so they were also discarded from the analysis.

### Number of Guaranteed Generations

Given the observation that carrying capacity of experimental populations can be recovered within a transfer period (i.e., 1 day), the guaranteed generation numbers in 900 days for dilution factors  $\frac{1}{4}$   $10^{-1}$ ,  $10^{-4}$ , and  $10^{-7}$  were, respectively, estimated as 900 times of  $\log_2(10)$ ,  $\log_2(10^4)$ , and  $\log_2(10^7)$ , which are equal to  $\sim$  3.0k, 12k, and 21k.

### Rate of Genomic Evolution

The level of genomic divergence for each experimental population at each time point is defined by summing all DAFs of detected mutations. We then calculated the mean genomic divergence across all eligible experimental populations in each combination of transfer size and genetic background. We further performed the linear regression by the function “lm” in R with formula “mean genomic divergence ~ guaranteed generations  $\beta$  0,” which enforces the y-intercept as 0. The slope of the regression is the estimated rate of genomic evolution. We also performed a nonlinear regression using the formula “mean genomic divergence ~ guaranteed generations  $\beta$  square root of guaranteed generations  $\beta$  0,” which was previously proposed to catch the trend of diminishing returns (Tenailon et al. 2016).

### Identification of Fixed Mutations by Hidden Markov Chain

For each population, caHMM was performed using a modified version (Behringer et al. 2020) of previously released code (Good et al. 2017). For the populations in which caHMM did not finish, we instead performed well-mixed hidden Markov chain (wmHMM) using a modified version (Behringer et al. 2020) of previously released code (Good et al. 2017). The single clade in wmHMM is defined as the basal clade. Fixed mutations are then defined as mutations that are inferred to be fixed in basal, major, or minor clade in the results of either analysis.

### Estimation of Selection Coefficients

For each fixed mutation, we estimated its selection coefficients by its temporal data of DAFs. If a mutation is in the basal clade, no correction is needed. If caHMM infers a mutation belongs to the major or minor clade, its corrected allele frequency will be the DAF divided by the proportion of the population belonging to the major or minor clade (also inferred by caHMM). For each of two available consecutive time points  $i$  and  $j$ , if the corrected allele frequencies at both time points ( $p_i$  and  $p_j$ ) are smaller than 0.95 and larger than 0.05, the selection coefficient is calculated as

$$\ln \frac{p_j \left( \frac{1 - p_i p_j}{1 - p_j} \right)}{p_i} = s(t - t_i);$$

where  $t_i$  and  $t_j$  are the numbers of guaranteed generations at time point  $i$  and  $j$ , respectively. The negative values were discarded. The largest positive value across all pairs of time point was used as the final measurement.

### Calculation of Neutrality Index

As discussed in the main text, we defined  $(F_N/F_S)/(U_N/U_S)$  as the neutrality index of nonsynonymous SNPs, where  $F_N$  is the

number of nonsynonymous SNPs fixed within a clade or fixed in an entire population,  $F_S$  is the number of synonymous SNPs fixed within a clade or fixed in an entire population,  $U_N$  is the number of nonsynonymous SNPs in the mutation-accumulation experiment, and  $U_S$  is the number of synonymous SNPs in the mutation-accumulation experiment. We also similarly defined  $(F_i/F_S)/(U_i/U_S)$  as neutrality index of intergenic SNPs, where  $F_i$  is the number of intergenic SNPs fixed within a clade or fixed in an entire population, and  $U_i$  is the number of intergenic SNPs in the mutation-accumulation experiment. The values of  $U_N$ ,  $U_S$ , and  $U_i$  are from a previously published mutation-accumulation experiment of our ancestral lines (Lee et al. 2012). We calculated population-specific indexes and then acquired population-wise mean and SE. The populations with  $F_N \approx 0$  and  $F_i \approx 0$  were discarded in the calculation of neutrality index of nonsynonymous and intergenic SNPs, respectively.

### Calculation of G-Scores

For each combination of genetic background and transfer size, we quantified the parallelism of the fixed nonsynonymous mutations using the sum of  $G$ -scores across genes (Tenailon et al. 2016). A larger  $G$ -score for a gene suggests that the fixed nonsynonymous mutations are more overrepresented in that gene. Specifically, to calculate a genic  $G$ -score ( $G_i$ ), we first counted the observed number of fixed nonsynonymous mutations in gene  $i$  ( $O_i$ ) per a combination of genetic background and transfer size. We then calculated the expected number for gene  $i$  ( $E_i$ ) by  $O_{\text{tot}}(L_i/L_{\text{tot}})$ , where  $O_{\text{tot}} \approx R_i O_i$ ,  $L_i$  is the number of nonsynonymous sites for gene  $i$ , and  $L_{\text{tot}} \approx R_i L_i$ . In the end,  $G_i$  is defined by  $2O_i \ln(O_i/E_i)$  or defined as zero when  $O_i \approx 0$  or when  $2O_i \ln(O_i/E_i) < 0$ .

It was noted that the null expectation of  $G$ -scores varies with total number of fixed nonsynonymous mutations (Behringer et al. 2020). Therefore, for each combination of genetic background and transfer size, we performed 20,000 simulations in each of which  $O_{\text{tot}}$  hits are randomly distributed among all  $L_{\text{tot}}$  sites across all genes in the reference genome. Then the significance of the sum of  $G$ -scores was evaluated by the  $z$  score defined by (the observed sum—mean of simulated sums)/(standard deviation of simulated sums).

### Calculation of Mean Bray-Curtis Similarity

For each combination of genetic background and transfer size, we also quantified the parallelism of the fixed nonsynonymous mutations using the mean Bray-Curtis similarity across all pairs of experimental populations (Turner et al. 2018; Behringer et al. 2020). Specifically, for a pair of populations  $j$  and  $k$ , their Bray-Curtis similarity is defined by

$$P_{i, o_{ij} - o_{ik}} = 1 - P\left(\frac{o_{ij} - o_{ik}}{\sigma_j - \sigma_k}\right);$$

where  $o_{ij}$  and  $o_{ik}$  are the observed number of fixed nonsynonymous mutations in gene  $i$  for population  $j$  and  $k$ , respectively.

For each combination of genetic background and transfer size, we also performed 1,000 simulations to acquire the null distribution. In each simulation, we randomly sample the nonsynonymous sites up to the number of observed fixed nonsynonymous mutations for each population and calculated mean Bray-Curtis similarity as described above. After acquiring the null distribution, we evaluated the significance of the observed mean Bray-Curtis similarity by calculating the  $z$  score defined by (the observed value—mean of simulated values)/(standard deviation of simulated values).

### Overrepresentation of the Genes Affected by Nonsynonymous Mutations

To evaluate the significance of  $G$ -score for gene  $i$ , we directly compared the  $G_i$  to the distribution of 20,000 simulated  $G_i$ , and the  $P$  value was defined as the proportion of simulated  $G_i$  larger or equal to the observed  $G_i$ . For multiple test correction, we multiplied each gene's  $P$  value by the number of genes with at least one hit by the set of fixed nonsynonymous mutations (Bonferroni correction). The genes are called significant only if the genes show Bonferroni corrected  $P$  value  $< 0.05$ .

### Enrichment Test of GO Terms and KEGG Pathways

Using the set of significant genes, we performed the enrichment test of GO terms using the function "enrichGO" in R package "DOSE" (Yu et al. 2015) with  $q$ -value cut-off  $\frac{1}{4} 0.05$  and the organismal database as org. ECK12.eg.db. We also performed the enrichment test of KEGG pathways using the function "enrichKEGG" in the same package but found no terms with  $q$ -value  $< 0.05$ .

### Overrepresentation of the Intergenic Regions Affected by Fixed Mutations

The identification of the intergenic regions affected by mutations was also performed by the way similar to identify the genes affected by nonsynonymous fixed mutations in genic  $G$ -score approach (Tenaillon et al. 2016). Instead of focusing on genic regions, genome-wide intergenic regions are focused. For each combination of genetic background and transfer size, we first counted the observed number of intergenic mutations in intergenic region  $i$  ( $O_i$ ), and the expected number for intergenic region  $i$  ( $E_i$ ) was calculated by  $O_{\text{tot}}(L_i/L_{\text{tot}})$ , where  $O_{\text{tot}} \frac{1}{4} R_i$ ,  $O_i$ ,  $L_i$  is the length for intergenic region  $i$ , and  $L_{\text{tot}} \frac{1}{4} R_i L_i$ . The  $G$ -score for intergenic region  $i$  ( $G_i$ ) was then calculated by  $2O_i \ln(O_i/E_i)$ , following the methods described in the above section. We also performed 20,000 sim-

each gene  $i$  following the methods described in the above section.

ulations and determined the Bonferroni corrected  $P$  value for

The identification of the genes affected by structural fixed mutations was performed by the way similar to identify the genes affected by nonsynonymous fixed mutations in genic G-score approach (Tenailon et al. 2016). Structural mutations include indels and IS-element insertions. For each combination of genetic background and transfer size, we first counted the observed number of populations with any structural mutations in gene  $i$  ( $O_i$ ), and the expected number for gene  $i$  ( $E_i$ ) was calculated by  $O_{\text{tot}} \cdot \frac{L_i}{L_{\text{tot}}}$ , where  $O_{\text{tot}}$  is the total number of structural mutations,  $L_i$  is the gene length for gene  $i$ , and  $L_{\text{tot}}$  is the total length of the genome. The G-score for gene  $i$  ( $G_i$ ) was then calculated by  $2 \cdot \ln(O_i/E_i)$ , following the methods described in the above section.

We also performed 20,000 simulations and determined the Bonferroni corrected  $P$  value for each gene  $i$  following the methods described in the above section. As a result, we found all the genes with  $O_i \geq 2$  show Bonferroni corrected  $P$  value  $< 0.05$ .

### Correlations Between Pairs of SNPs

For each evolved population, we focused on the nonsynonymous, synonymous, and intergenic SNPs in which at least two nonzero DAFs were found. For each pair of two such SNPs, we calculated the change of DAFs. Then we calculated Pearson's correlation coefficients across only all the odd-numbered changes of DAFs to avoid nonindependence (Lynch and Ho 2020). That is to say, if a population has sequencing profiles available for analysis at every sampling time points (Days 90, 200, 300, 400, 500, 600, 700, 800, 900), we calculate Pearson's correlation coefficients using the five changes of DAFs: the one between Day 0 and 90, between Day 200 and 300, between Day 400 and 500, between Day 600 and 700, and between Day 800 and 900. For another example of population, if its sequencing profile at Day 90 is discarded from analysis due to its low quality, we calculate the Pearson's correlation coefficients using only the four changes of DAFs: the one between Day 0 and 200, between Day 300 and 400, between Day 500 and 600, and between Day 700 and 800. Note that at least four changes of DAFs are used for each evolved population because no populations have more than two missing profiles. We then get the distribution of Pearson's correlation coefficients for each evolved population.

To establish the baseline for comparison in each combination of transfer size and genetic background, we also generated the set of Pearson's correlation coefficients using two random mutations from two different random populations with all sequencing profiles available (i.e., they are unlinked for sure). We followed the same procedure above to calculate Pearson's correlation coefficients for each pair of unlinked

mutations. When simulating a distribution of Pearson's correlation coefficients, we used 100 pair of unlinked mutations. We then repetitively performed 100 rounds of simulation to get the mean and SE for the distribution.

## Supplementary Material

**Supplementary data** are available at *Genome Biology and Evolution* online.

## Acknowledgments

We thank D. Allan Drummond, Jay T. Lennon, James B. McKinlay, and Ryan Stikeleather for their helpful comments. High Performance Computing Resources were provided and maintained by the National Center for Genome Analysis Support at Indiana University. This work was supported by the Army Research Office [ARO65308-LS-MU and W911NF-14-1-0411 to M.L.]; and the National Institutes of Health [F32GM123703 to M.G.B., and R35GM122566 to M.L.].

## Author Contributions

Conceptualization, W-C.H., M.G.B., and M.L.; Methodology, W-C.H., M.G.B., and S.F.M.; Investigation, W-C.H., M.G.B., S.F.M., J.G., A.N., M.A., and G.F.B.; Formal Analysis, W-C.H., and M.G.B.; Writing—Original Draft, W-C.H., and M.G.B.; Writing—Review and Editing, W-C.H., M.G.B., W-C.H., and M.L.; Visualization, W-C.H. and M.G.B.; Funding Acquisition, M.G.B. and M.L.

## Data Availability

Sequencing data generated during this study are available at NCBI's Sequencing Read Archive: BioProject PRJNA532905 and PRJNA722381. Codes generated to analyze sequencing data are available at [https://github.com/LynchLab/ECEE\\_Hypermutator](https://github.com/LynchLab/ECEE_Hypermutator) (last accessed November 24, 2021)

## Literature Cited

- Abraham JM, Freitag CS, Clements JR, Eisenstein BI. 1985. An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Proc Natl Acad Sci USA*. 82(17):5724–5727.
- Arjan JA, et al. 1999. Diminishing returns from mutation supply rate in asexual populations. *Science* 283(5400):404–406.
- Avalos Vizcarra I, et al. 2016. How type 1 fimbriae help *Escherichia coli* to evade extracellular antibiotics. *Sci Rep*. 6:18109.
- Bachtrog D, Gordo I. 2004. Adaptive evolution of asexual populations under Muller's ratchet. *Evolution* 58(7):1403–1413.
- Barrick JE, et al. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461(7268):1243–1247.
- Barroso-Batista J, et al. 2014. The first steps of adaptation of *Escherichia coli* to the gut are dominated by soft sweeps. *PLoS Genet*. 10(3):e1004182.
- Baym M, et al. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10(5):e0128036.
- Behringer MG, et al. 2018. *Escherichia coli* cultures maintain stable sub-population structure during long-term evolution. *Proc Natl Acad Sci USA*. 115(20):E4642–E4650.
- Behringer MG, et al. 2020. Evolutionary response to resource deprivation: parallelism and nonmonotonicity. *bioRxiv*: 865584. Available from: <https://doi.org/10.1101/865584>.
- Bell AW, et al. 1986. The nucleotide sequences of the rbsD, rbsA, and rbsC genes of *Escherichia coli* K12. *J Biol Chem*. 261(17):7652–7658.
- Bollback JP, Huelsenbeck JP. 2007. Clonal interference is alleviated by high mutation rates in large populations. *Mol Biol Evol*. 24(6):1397–1406.
- Campos PR, Wahl LM. 2010. The adaptation rate of asexuals: deleterious mutations, clonal interference and population bottlenecks. *Evolution* 64(7):1973–1983.
- Couce A, Alonso-Rodriguez N, Costas C, Oliver A, Blazquez J. 2016. Intrapopulation variability in mutator prevalence among urinary tract infection isolates of *Escherichia coli*. *Clin Microbiol Infect*. 22:566 e561–567.
- Couce A, et al. 2017. Mutator genomes decay, despite sustained fitness gains, in a long-term experiment with bacteria. *Proc Natl Acad Sci USA*. 114(43):E9026–E9035.
- Crow JF, Kimura M. 1965. Evolution in sexual and asexual populations. *Am Nat*. 99(909):439–450.
- Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol Biol*. 1151:165–188.
- Desai MM, Fisher DS. 2007. Beneficial mutation selection balance and the effect of linkage on positive selection. *Genetics* 176(3):1759–1798.
- Desai MM, Fisher DS, Murray AW. 2007. The speed of evolution and maintenance of variation in asexual populations. *Curr Biol*. 17(5):385–394.
- Domka J, Lee J, Bansal T, Wood TK. 2007. Temporal gene-expression in *Escherichia coli* K-12 biofilms. *Environ Microbiol*. 9(2):332–346.
- Foster PL. 2006. Methods for determining spontaneous mutation rates. *Methods Enzymol*. 409:195–213.
- Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* 102–103:127–144.
- Gillespie JH. 2000. Genetic drift in an infinite population. The pseudohitchhiking model. *Genetics* 155(2):909–919.
- Giraud A, et al. 2001. Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* 291(5513):2606–2608.
- Good BH, Rouzine IM, Balick DJ, Hallatschek O, Desai MM. 2012. Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proc Natl Acad Sci USA*. 109(13):4950–4955.
- Good BH, Desai MM. 2014. Deleterious passengers in adapting populations. *Genetics* 198(3):1183–1208.
- Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. 2017. The dynamics of molecular evolution over 60,000 generations. *Nature* 551(7678):45–50.
- Iuchi S, Cameron DC, Lin EC. 1989. A second global regulator gene (arcB) mediating repression of enzymes in aerobic pathways of *Escherichia coli*. *J Bacteriol*. 171(2):868–873.
- Kaper JB, Nataro JP, Mobley HL. 2004. Pathogenic *Escherichia coli*. *Nat Rev Microbiol*. 2(2):123–140.
- Kim Y, Orr HA. 2005. Adaptation in sexuals vs. asexuals: clonal interference and the Fisher-Muller model. *Genetics* 171(3):1377–1386.
- Kimura M. 1967. On the evolutionary adjustment of spontaneous mutation rates. *Genet Res*. 9(1):23–34.
- Kristensen HH, Valentin-Hansen P, Sogaard-Andersen L. 1996. CytR/cAMP-CRP nucleoprotein formation in *E. coli*: the CytR repressor binds its operator as a stable dimer in a ternary complex with cAMP-CRP. *J Mol Biol*. 260(2):113–119.

- Lange R, Hengge-Aronis R. 1991. Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. *Mol Microbiol*. 5(1):49–59.
- Le Trong I, et al. 2010. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like beta sheet twisting. *Cell* 141(4):645–655.
- Lee H, Popodi E, Tang H, Foster PL. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci USA*. 109(41):E2774–2783.
- Lenski RE, Rose MR, Simpson SC, Tadler SC. 1991. Long-Term experimental evolution in *Escherichia coli*. I. adaptation and divergence during 2,000 generations. *Am Nat*. 138(6):1315–1341.
- Li J, et al. 2007. Quorum sensing in *Escherichia coli* is signaled by Al-2/LsrR: effects on small RNA and biofilm architecture. *J Bacteriol*. 189(16):6011–6020.
- Long H, et al. 2016. Antibiotic treatment enhances the genome-wide mutation rate of target cells. *Proc Natl Acad Sci USA*. 113(18):E2498–E2505.
- Luksza M, Lassig M. 2014. A predictive fitness model for influenza. *Nature* 507(7490):57–61.
- Lynch M. 2008. The cellular, developmental and population-genetic determinants of mutation-rate evolution. *Genetics* 180(2):933–943.
- Lynch M. 2010. Evolution of the mutation rate. *Trends Genet*. 26(8):345–352.
- Lynch M, et al. 2016. Genetic drift, selection and the evolution of the mutation rate. *Nat Rev Genet*. 17(11):704–714.
- Lynch M, Ho WC. 2020. The limits to estimating population-genetic parameters with temporal data. *Genome Biol Evol*. 12(4):443–455.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 17(1):10–12.
- Mauzy CA, Hermodson MA. 1992. Structural homology between rbs repressor and ribose binding protein implies functional similarity. *Protein Sci*. 1(7):843–849.
- McCloskey D, et al. 2018. Evolution of gene knockout strains of *E. coli* reveal regulatory architectures governed by metabolism. *Nat Commun*. 9(1):3796.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351(6328):652–654.
- McDonald MJ, Hsieh YY, Yu YH, Chang SL, Leu JY. 2012. The evolution of low mutation rates in experimental mutator populations of *Saccharomyces cerevisiae*. *Curr Biol*. 22(13):1235–1240.
- McDonald MJ. 2019. Microbial experimental evolution—a proving ground for evolutionary theory and a tool for discovery. *EMBO Rep*. 20(8):e46992.
- McFarland CD, et al. 2017. The damaging effect of passenger mutations on cancer progression. *Cancer Res*. 77(18):4763–4772.
- Mehta HH, et al. 2019. The essential role of hypermutation in rapid adaptation to antibiotic stress. *Antimicrob Agents Chemother*. 63(7):e00744–19.
- Missiakas D, Mayer MP, Lemaire M, Georgopoulos C, Raina S. 1997. Modulation of the *Escherichia coli* sigmaE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. *Mol Microbiol*. 24(2):355–371.
- Monod J. 1949. The growth of bacterial cultures. *Annu Rev Microbiol*. 3(1):371–394.
- Muller HJ. 1932. Some genetic aspects of sex. *Am Nat*. 66(703):118–138.
- Muller HJ. 1950. Our load of mutations. *Am J Hum Genet*. 2(2):111–176.
- Navarro Llorens JM, Tormo A, Martinez-Garcia E. 2010. Stationary phase in gram-negative bacteria. *FEMS Microbiol Rev*. 34(4):476–495.
- Naves P, et al. 2008. Correlation between virulence factors and in vitro biofilm formation by *Escherichia coli* strains. *Microb Pathog*. 45(2):86–91.
- Neher RA. 2013. Genetic draft, selective interference, and population genetics of rapid adaptation. *Annu Rev Ecol Evol Syst*. 44(1):195–215.
- Olsen PB, Schembri MA, Gally DL, Klemm P. 1998. Differential temperature modulation by H-NS of the fimB and fimE recombinase genes which control the orientation of the type 1 fimbrial phase switch. *FEMS Microbiol Lett*. 162(1):17–23.
- Park SC, Krug J. 2007. Clonal interference in large populations. *Proc Natl Acad Sci USA*. 104(46):18135–18140.
- Park SC, Simon D, Krug J. 2010. The speed of evolution in large asexual populations. *J Stat Phys*. 138(1–3):381–410.
- Penisson S, Singh T, Sniegowski P, Gerrish P. 2017. Dynamics and fate of beneficial mutations under lineage contamination by linked deleterious mutations. *Genetics* 205(3):1305–1318.
- Pletnev P, Osterman I, Sergiev P, Bogdanov A, Dontsova O. 2015. Survival guide: *escherichia coli* in the stationary phase. *Acta Nat*. 7(4):22–33.
- Pratt LA, Kolter R. 1998. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol*. 30(2):285–293.
- Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Mol Biol Evol*. 13(6):735–748.
- Raynes Y, Halstead AL, Sniegowski PD. 2014. The effect of population bottlenecks on mutation rate evolution in asexual populations. *J Evol Biol*. 27(1):161–169.
- Rolfe MD, et al. 2011. Transcript profiling and inference of *Escherichia coli* K-12 ArcA activity across the range of physiologically relevant oxygen concentrations. *J Biol Chem*. 286(12):10147–10154.
- Rossi E, et al. 2018. “It’s a gut feeling”—*Escherichia coli* biofilm formation in the gastrointestinal tract environment. *Crit Rev Microbiol*. 44(1):1–30.
- Schembri MA, Kjaergaard K, Sokurenko EV, Klemm P. 2001. Molecular characterization of the *Escherichia coli* FimH adhesin. *J Infect Dis*. 183 Suppl 1:S28–31.
- Schiffels S, Szollosi GJ, Mustonen V, Lassig M. 2011. Emergent neutrality in adaptive asexual evolution. *Genetics* 189(4):1361–1375.
- Sezonov G, Joseleau-Petit D, D’Ari R. 2007. *Escherichia coli* physiology in Luria-Bertani broth. *J Bacteriol*. 189(23):8746–8749.
- Sharma G, et al. 2016. *Escherichia coli* biofilm: development and therapeutic strategies. *J Appl Microbiol*. 121(2):309–319.
- Sheidy DT, Zielke RA. 2013. Analysis and expansion of the role of the *Escherichia coli* protein ProQ. *PLoS One* 8(10):e79656.
- Shewaramani S, et al. 2017. Anaerobically grown *Escherichia coli* has an enhanced mutation rate and distinct mutational spectra. *PLoS Genet*. 13(1):e1006570.
- Sniegowski PD, Gerrish PJ, Lenski RE. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387(6634):703–705.
- Spears PA, Schauer D, Orndorff PE. 1986. Metastable regulation of type 1 piliation in *Escherichia coli* and isolation and characterization of a phenotypically stable mutant. *J Bacteriol*. 168(1):179–185.
- Sprouffske K, Aguilar-Rodriguez J, Sniegowski P, Wagner A. 2018. High mutation rates limit evolutionary adaptation in *Escherichia coli*. *PLoS Genet*. 14(4):e1007324.
- Sung W, Ackerman MS, Miller SF, Doak TG, Lynch M. 2012. Drift-barrier hypothesis and mutation-rate evolution. *Proc Natl Acad Sci USA*. 109(45):18488–18492.
- Swings T, et al. 2017. Adaptive tuning of mutation rates allows fast response to lethal stress in *Escherichia coli*. *Elife* 6:e22939.
- Tachida H. 2000. DNA evolution under weak selection. *Gene* 261(1):3–9.
- Tarnita CE. 2017. The ecology and evolution of social behavior in microbes. *J Exp Biol*. 220(Pt 1):18–24.
- Tenaillon O, Toupance B, Le Nagard H, Taddei F, Godelle B. 1999. Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. *Genetics* 152(2):485–493.

- Tenaillon O, Taddei F, Radmian M, Matic I. 2001. Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Res Microbiol.* 152(1):11–16.
- Tenaillon O, et al. 2016. Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature* 536(7615):165–170.
- Turner CB, Marshall CW, Cooper VS. 2018. Parallel genetic adaptation across environments differing in mode of growth or resource availability. *Evol Lett.* 2(4):355–367.
- Van den Bergh B, Swings T, Fauvart M, Michiels J. 2018. Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiol Mol Biol Rev.* 82(3):e00008-18.
- Veschetti L, Sandri A, Krogh Johansen H, Lleo' MM, Malerba G. 2020. Hypermutation as an evolutionary mechanism for *Achromobacter xylosoxidans* in cystic fibrosis lung Infection. *Pathogens* 9(2):72.
- Vogwill T, Phillips RL, Gifford DR, MacLean RC. 2016. Divergent evolution peaks under intermediate population bottlenecks during bacterial experimental evolution. *Proc Biol Sci.* 283:20160749.
- Wahl LM, Gerrish PJ, Saika-Voivod I. 2002. Evaluating the impact of population bottlenecks in experimental evolution. *Genetics* 162(2):961–971.
- Waksman G, Hultgren SJ. 2009. Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat Rev Microbiol.* 7(11):765–774.
- Watson N, Dunyak DS, Rosey EL, Slonczewski JL, Olson ER. 1992. Identification of elements involved in transcriptional regulation of the *Escherichia coli* cad operon by external pH. *J Bacteriol.* 174(2):530–540.
- Wein T, Dagan T. 2019. The effect of population bottleneck size and selective regime on genetic diversity and evolvability in bacteria. *Genome Biol Evol.* 11(11):3283–3290.
- Wielgoss S, et al. 2013. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc Natl Acad Sci USA.* 110(1):222–227.
- Wilke CO. 2004. The speed of adaptation in large asexual populations. *Genetics* 167(4):2045–2053.
- Williams AB, Foster PL. 2012. Stress-induced mutagenesis. *EcoSal Plus.* 5(1):10.1128/ecosalplus.7.2.3.
- Williams MD, Ouyang TX, Flickinger MC. 1994. Starvation-induced expression of SspA and SspB: the effects of a null mutation in sspA on *Escherichia coli* protein synthesis and survival during growth and prolonged starvation. *Mol Microbiol.* 11(6):1029–1043.
- Yu GC, Wang LG, Yan GR, He QY. 2015. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* 31(4):608–609.
- Zheng Q. 2017. rSalvador: an R package for the fluctuation experiment. *G3 (Bethesda)* 7(12):3849–3856.

Associate editor: Brandon Gaut