Title: Shifts in mutation spectra enhance access to beneficial mutations

Authors: Mrudula Sane[1], Gaurav D Diwan[1,2,3], Bhoomika A Bhat[4], Lindi M Wahl[5] and Deepa Agashe[1]*

Affiliations:
[1] National Centre for Biological Sciences, GKVK Campus, Bellary Road, Bengaluru, India
[2] CellNetworks, Bioquant, University of Heidelberg, 69120 Heidelberg, Germany
[3] Biochemie Zentrum Heidelberg (BZH), University of Heidelberg, 69120 Heidelberg, Germany
[4] Undergraduate Programme, Indian Institute of Science, Bengaluru, India
[5] Western University, London, Ontario, Canada

*Correspondence to:
dagashe@ncbs.res.in

One Sentence Summary:

Altered mutational biases allow populations to sample poorly explored mutational space, with wide-ranging evolutionary consequences.

Abstract:

Biased mutation spectra are pervasive, with widely varying direction and magnitude of mutational bias. Why are unbiased spectra rare, and how do such diverse biases evolve? We find that experimentally changing the mutation spectrum increases the beneficial mutation supply, because populations sample mutational classes that were poorly explored by the ancestor. Simulations show that selection does not oppose the evolution of a mutational bias in an unbiased ancestor; but it favours changing the direction of a long-term bias. Indeed, spectrum changes in the bacterial phylogeny are frequent, typically involving reversals of ancestral bias. Thus, shifts in mutation spectra evolve under selection, and may directly alter outcomes of adaptive evolution by facilitating access to beneficial mutations.

Main Text:

The mutation spectrum describes the frequency of various classes of sampled mutations (e.g. transversions vs. transitions (Tv/Ts), or base pair vs. copy number changes), often determined by the action of DNA repair enzymes. Changes in DNA repair function typically alter both mutation rate and spectrum, e.g. in “mutator” genotypes with loss-of-function mutations (*1*), or during stress-induced mutagenesis (*2*, *3*). By determining the pool of genetic variants available for selection, mutation spectra can shape key genome features (e.g. nucleotide (*4*, *5*), codon, and amino acid composition (*6*)); determine the genetic basis of adaptation (*7-9*), driving convergent evolution (*10*, *11*); and shape evolution of resistance to antibiotics (*12-14*) and anti-cancer drugs (*15*). Although mutation spectra are important for adaptive evolution, their role has been underappreciated (*7*).

Most species have a skewed mutation spectrum, with substantial variation in the direction and magnitude of bias (*16*, *17*) (Fig. 1A, Table S1). This diversity implies the occurrence of major evolutionary shifts in mutation spectra (Fig. 1A). However, the frequency, underlying evolutionary processes, and consequences of spectrum shifts are unknown. For instance, why are unbiased spectra rare? Do spectrum shifts – e.g. change from an unbiased to a biased state, or in the direction of bias – evolve under selection? A spectrum shift may persist if the new bias favours mutational classes that are inherently beneficial (*7*, *18*), allowing the evolved bias to hitchhike with beneficial mutations. Similarly, it is speculated that stress-induced changes in mutation spectra may enhance sampling of new beneficial mutations, by altering the distribution of fitness effects (DFE) (*19*, *20*). However, it is difficult to explain how specific mutational classes can be universally beneficial; and harder still to reconcile such a benefit with the observed diversity in bias.

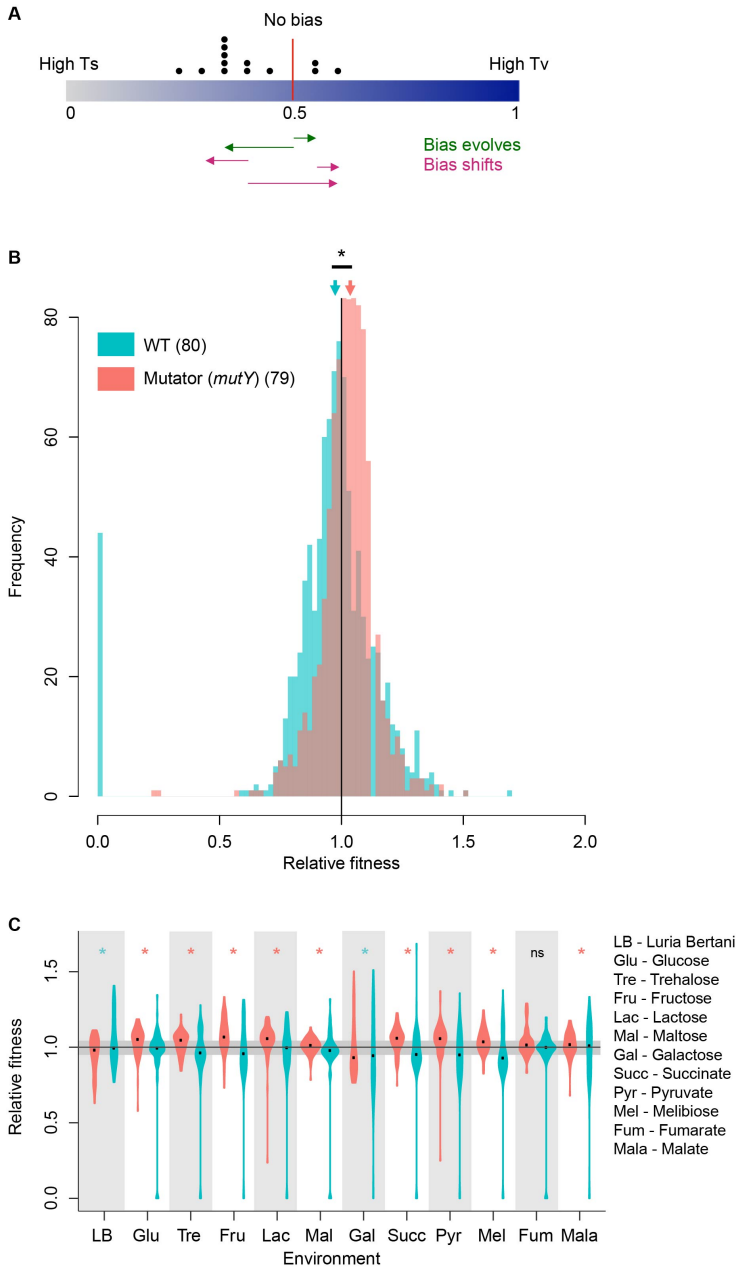
We addressed these gaps using a combination of experiments, simulations, and phylogenetic analyses. To measure the immediate evolutionary impact of the mutation spectrum, we estimated the genome-wide DFE of new mutations in *E. coli* (Fig. S1). We manipulated the mutation spectrum by deleting a DNA repair gene (Δ*mutY* "mutator") from our wild type ("WT") strain, altering both the mutation spectrum and rate (*1*). We evolved independent lineages of WT and mutator under mutation accumulation (MA), and sequenced whole genomes of evolved isolates to identify strains carrying a single new mutation each. Thus, we minimized the impact of selection on the mutation spectrum and the DFE, and decoupled the effects of the mutator's spectrum from its high mutation rate. Measuring the effect of each mutation on growth rate, we determined the effects of mutation spectrum on the beneficial mutation supply and genetic load. Next, to test the generality and longer-term evolutionary consequences of shifts in mutation spectra, we simulated adaptive walks across NK fitness landscapes (*21*). Finally, we inferred evolutionary transitions in DNA repair enzymes across the bacterial phylogeny. Our results show that shifts in mutation spectra can evolve under selection and fuel adaptation via previously unsampled evolutionary paths.

The mutator has a distinct DFE with more beneficial mutations, reduced load, and altered pleiotropic effects

From our MA experiments, we obtained 80 evolved WT (*22*) and 79 mutator strains, each carrying a single distinct mutation with respect to its ancestor (Tables S2, S3). We measured the fitness effect (relative maximal growth rate) of each mutation in 12 environments with different carbon sources, constructing global and environment-specific empirical DFEs (Fig. 1B–C, Fig. S2). The mutator and WT had distinct global DFEs (Fig. 1B; KS test: $D=0.23$, $p=2.2\times10^{-16}$); but the environment alone did not significantly impact fitness effects (ANOVA: relative fitness ~ strain x environment, $p_{strain}=1.25\times10^{-18}$, $p_{env}=0.69$, $p_{strainxenv}=1.8\times10^{-5}$). New mutations had a median fitness advantage of ~3% in the mutator, and an equivalent fitness disadvantage in the WT (Fig. 1B; a different method of calculating fitness yielded median selection coefficients of -0.03 for WT and

0.049 for mutator). This pattern was consistent for environment-specific DFEs in 9 of 12 environments (Fig. 1C, Table S4), indicating that the mutator has a consistently beneficial-shifted DFE compared to the WT.

Figure 1. Mutator DFE is shifted towards more beneficial mutations. (A) Illustration of diversity and potential evolutionary shifts in mutation spectra. Each point represents the Tv bias of a distinct microbial species; Tv = transversion, Ts = transition (Table S1). (B) Global DFE (distribution of fitness effects) of single mutations in wild type (WT) and mutator (Δ*mutY*). Number of mutations per strain is in parentheses in the key (n; each DFE was constructed with n x 12 data points, pooled across 12 environments). Black vertical line: relative fitness of 1 (neutral); arrows: median fitness; asterisk: significant difference. (C) Environment-specific DFEs (except LB, M9 minimal medium + 5mM indicated carbon source). Black dots: medians; horizontal grey shaded area: 5% measurement error margin (0.95 < relative fitness < 1.05); horizontal black line: relative fitness = 1; asterisks: significant differences across strains, colored by strain with higher fitness effects (Table S4); ns: not significant. Relative fitness of each strain was calculated with respect to its ancestor (WT or mutator).

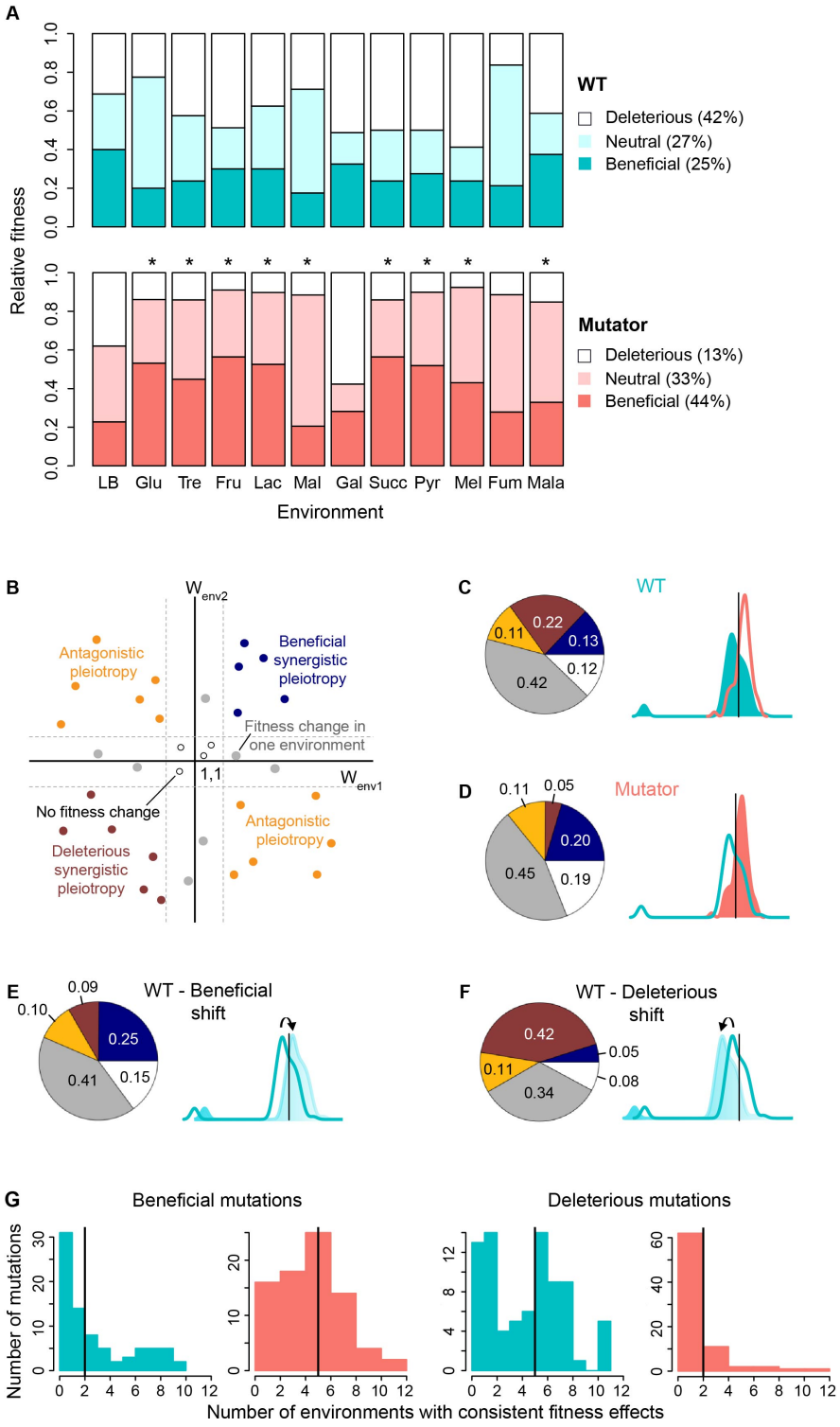


These effects are not explained by differential magnitudes of mutational effects; in fact, new mutations had stronger absolute effects in WT than in the mutator (Mann-Whitney U-test, W = 377300, $p = 3.25 \times 10^{-10}$). The probability of finding distinct mutational effects by chance was 0.0006 (Fig. S3A), and our sample size of ~80 mutations per strain adequately captured the presumed "true" global DFE (Fig. S3B, Table S5). The DFE differences were also robust to removing four conditionally lethal mutations found in evolved WT (Fig. S4, Table S4). Our single-mutation genome-wide DFEs have more beneficial mutations than previously reported DFEs (*23*), which were typically estimated using populations evolved under selection, carrying many or unknown numbers of mutations (i.e. confounded by epistasis and selection). Since there is no precedence for single-mutation genome-wide DFEs, we re-analysed a single-mutation single-gene DFE for antibiotic resistance (*24*) and found ~31% beneficial mutations. Thus, our estimated DFEs likely represent the genome-wide DFE of single mutations, suggesting that beneficial mutations may not be as rare as is usually believed.

To quantify the impact of the mutator's distinct DFE, we coarsely classified each mutation as neutral (0.95 > relative fitness > 1.05, conservatively accounting for ~5% fitness measurement error), beneficial (relative fitness >1.05), or deleterious (relative fitness <0.95). On average, the mutator had a greater fraction of beneficial mutations than the WT, but fewer deleterious mutations (Fig 2A, Table S6). Accounting for its order-of-magnitude higher mutation rate (Table S7), we estimated (following *25*) that the mutator should have an ~14-fold greater genome-wide supply of beneficial mutations, and an ~0.45-fold lower genetic load than the WT (Tables S8, S9). In contrast, if we ignored the mutator's distinct DFE, the supply of beneficial mutations would only be ~1.5 fold greater than WT, and the genetic load would be ~4 fold higher (Tables S8, S9). Thus, depending on their mutation rate and DFE, mutators may have a substantially higher supply of beneficial mutations, and effectively no genetic load.

The pleiotropic effect of beneficial mutations (Fig. 2B) is important in new environments, because it can facilitate adaptation (via synergistic pleiotropic benefits) and shape fitness tradeoffs (via antagonistic pleiotropy). The mutator had a distinct distribution of pleiotropic effects in 11 of 12 environments (Fig. S5, Table S10). These differences were driven by a higher incidence of beneficial synergistic pleiotropy and lower deleterious synergistic pleiotropy in the mutator; but equally low antagonistic pleiotropy in both strains (Fig. 2C–D, Fig. S5). Beneficial mutations in the mutator were also beneficial across many more environments (KS test, D=0.36, $p=6.7 \times 10^{-5}$); whereas deleterious mutations were deleterious in fewer environments than WT (KS test, D=0.51, $p=1.9 \times 10^{-9}$) (Fig. 2G, Fig. S6).

Figure 2. Evolutionary consequences of a DFE shifted towards beneficial mutations. (A) Proportion of beneficial, deleterious and neutral mutations in WT and mutator. Median values across environments are indicated in the key; asterisks: significant differences between WT and mutator (Table S6). (B) Schematic showing possible pleiotropic effects of mutations in two environments (W: relative fitness). (C–F) Median proportions of pleiotropic fitness effects of mutations in WT and mutator, and WT DFE artificially shifted towards either beneficial or deleterious mutations (see Methods). Schematics indicate global DFEs (WT=cyan, mutator=pink, WT-shifted=light blue); black lines: relative fitness = 1. (G) Pleiotropic effects of mutations across multiple environments. Frequency distributions show the proportion of mutations with consistent fitness effects (beneficial or deleterious) in a given number of environments. Black lines indicate medians.



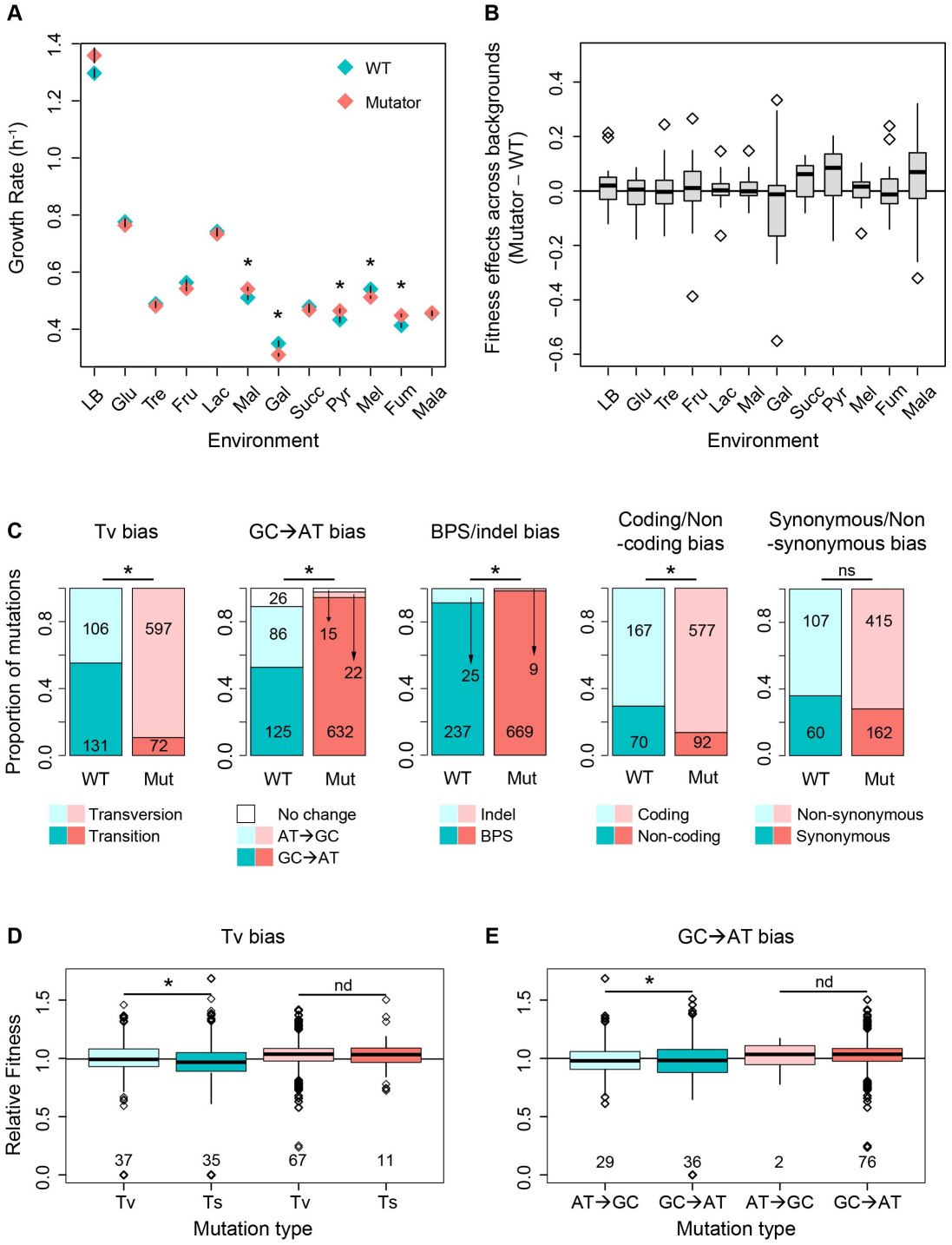
Thus, new beneficial mutations in the mutator are more likely to facilitate adaptation across many environments, but are no more likely to generate tradeoffs. We speculated that these differences may arise because the high proportion of beneficial mutations in the mutator might increase the likelihood of beneficial synergistic pleiotropy. Indeed, simulating an increase in the median fitness effect of the WT DFE without changing its shape ("WT-beneficial shift") mimicked pleiotropic effects observed in the mutator in all but one environment (compare Fig. 2C-D, Fig. S5, Table S10). Conversely, simply reducing the median fitness effect in the WT DFE ("WT-deleterious shift") lowered beneficial synergistic pleiotropy in 7 of 12 environments (compare Fig. 2C and 2F, Fig. S5, Table S10). Ignoring the mutation spectrum can thus cause overestimation of a mutator's genetic load, and underestimation of both the beneficial supply of mutations as well as their pleiotropic effects during adaptation.

The mutator's distinctive DFE arises from its distinct mutation spectrum

We had hypothesized that the biased spectrum of the mutator could generate a distinct DFE. However, other global effects of the initial *mutY* gene deletion could also lead to a DFE shift. For instance, if deleting *mutY* reduced the mutator's fitness, a larger fraction of new mutations might be beneficial (*26*). However, the mutator ancestor had lower fitness than WT in only two environments (Fig 3A, Table S11). Alternatively, epistatic interactions with *mutY* could increase the deleterious effect of new mutations in the WT. However, paired strains carrying the same mutation in either the WT or mutator background had similar fitness in all environments (Fig. 3B, Table S12). Thus, the observed shift in the mutator DFE cannot be explained by global effects of the original *mutY* deletion.

To directly test the effect of the spectrum, we asked whether the mutator over-samples specific classes of mutations that happen to be more beneficial (or less deleterious). The mutator strongly favours transversions, GC→AT mutations, base pair substitutions (BPS) and coding mutations (Fig. 3C, Table S7). The resulting skew in sample size of mutation classes limited our ability to determine the association between mutation bias and fitness effects across strains. However, in the WT, transversions and GC→AT mutations were indeed less deleterious than transitions and AT→GC mutations (Fig. 3D–E; fitness effects were pooled across environments, Table S13). We could not test the effect of BPS/indel bias, and the coding/non-coding bias had no effect (Fig. S7, Table S13). Other properties of the mutations (e.g. their location in the genome or specific genes, or the resulting amino acid changes) did not differ across strains (Fig. S8, S9; Tables S14, S15). Thus, the mutation spectrum – specifically the strong Tv and GC→AT bias – shapes the distinct DFE of the mutator.

Figure 3. Differences in WT and mutator DFEs are best explained by differences in mutation spectra. (A) Growth rate of WT and mutator ancestors (n=12-24/strain/environment). Asterisks: significant differences between strains (Table S11). (B) Pairwise differences in the fitness effects of 19 mutations across genetic backgrounds (Mutator – WT); none were different from zero (black line; Table S12). (C) Observed mutation biases in WT and mutator MA lines (including those with multiple mutations); the number of mutations of each type is noted. Asterisks: significant differences in proportions across strains (Table S7). (D–E) Fitness effects of different mutation classes, for MA-evolved isolates carrying single mutations. Boxplots (colored by strain and mutation class) show median, quartiles and range of relative fitness effects of mutations (open diamonds: outliers). Each boxplot was constructed with n x 12 data points (n = number of mutations, shown below each plot; pooled effects from 12 environments). Black horizontal lines: relative fitness = 1; asterisks: significant differences across mutation classes (Table S13); ns: not significant; nd: not determined.



Simulations demonstrate a general benefit of reversing mutation spectra

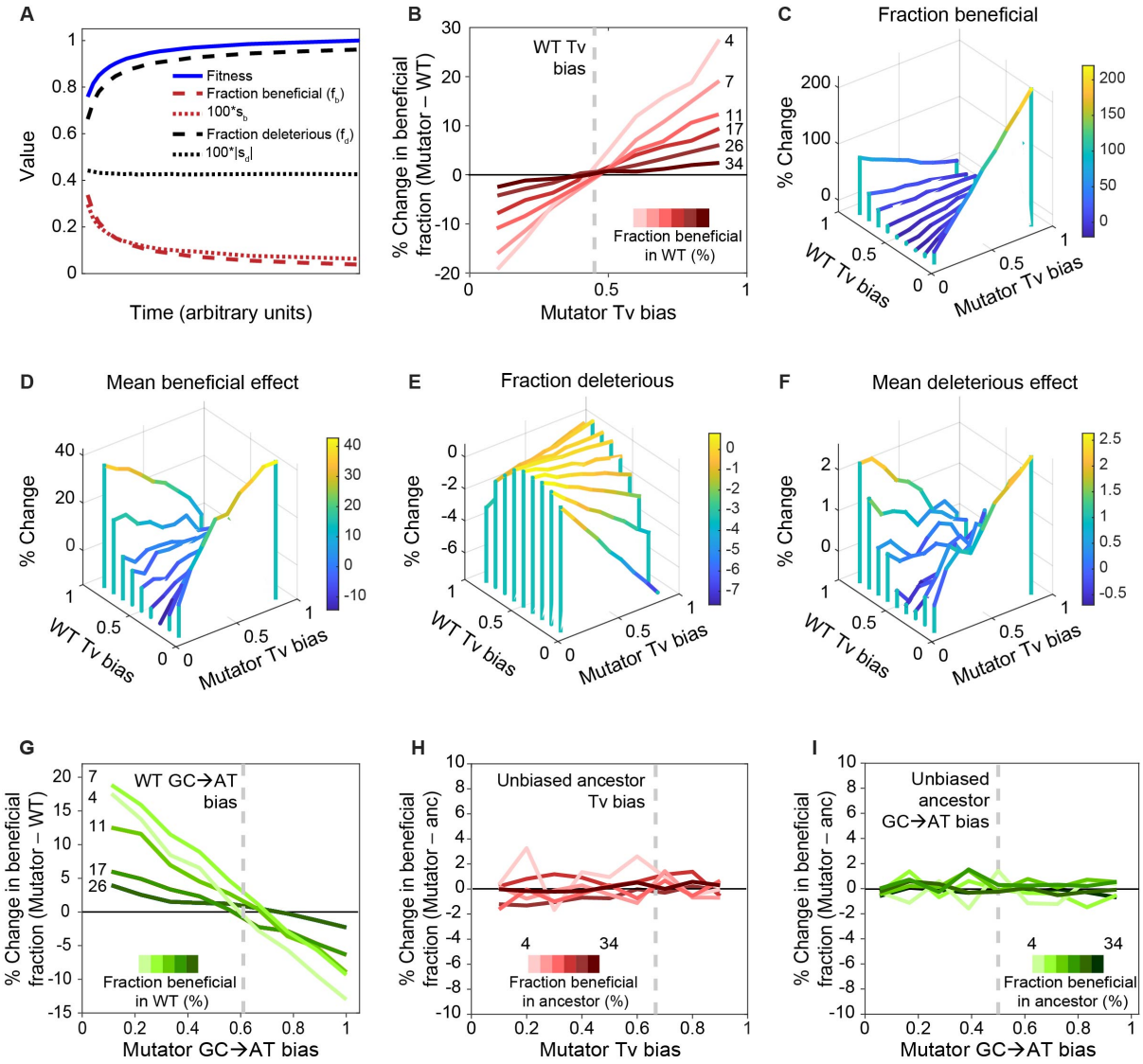
To test the generality and impact of our empirical results, we simulated adaptive walks (following (*27*)), modelling sequences of length N such that mutations could be classified as Ts vs. Tv or GC→AT vs. AT→GC. Starting from a genetically uniform population (i.e. a strong-selection, weak-mutation regime) with a randomly chosen ancestor sequence and mutation spectrum (absolute Tv bias = fraction of transversions), we allowed populations to explore successive mutational steps at randomly chosen loci (bases). At various points during the walk, we generated a DFE by simulating 300 possible substitutions, and computing their fitness given the underlying fitness landscape (affected by K other randomly chosen loci, to incorporate epistasis). From the DFE, we calculated the fraction of beneficial (f_b) and deleterious (f_d) mutations, and their effect size. We initially set N=200 and K=1 (i.e. mild epistasis, with each locus epistatically coupled to 1/199=0.5% loci), and present the average outcomes for 500 adaptive walks on 500 randomly-generated fitness landscapes (Fig. S10 shows variation across walks).

As expected, the mean population fitness increased during the adaptive walk, with a concomitant reduction in f_b and beneficial effect size, an increase in f_d, and a relatively constant deleterious effect size (Fig. 4A). Setting ancestral Tv bias to 0.45 (mimicking our WT), we compared the DFE generated by the ancestor at various time points, to the DFE that would be created if the bias were changed (mimicking a mutator). The mutator thus started with the same fitness and sequence as the ancestor, differing only in spectrum. As the population evolves (i.e. ancestor f_b decreases), mutators with a stronger bias (i.e. higher Tv) sample proportionally more beneficial mutations than the ancestor (Fig. 4B). More generally, exploring well-adapted populations (f_b = 0.04) but varying the ancestral mutation spectrum, we found that mutators that reverse the ancestral Tv bias have the greatest advantage. If the ancestor is biased toward Ts, mutators with higher Tv show a greater increase in f_b values compared to the ancestor (Fig. 4C), with larger effect sizes (Fig. 4D). Simultaneously, f_d also decreases (Fig. 4E), with a small increase in the effect size (Fig. 4F). These results were consistent for more rugged fitness landscapes with higher epistasis (up to 16% or 43% interacting loci, Fig. S11).

The symmetric nature of these results implies that a specific mutation class or change in spectrum is not always beneficial. In the example presented here, a bias towards more transversions was beneficial only because the ancestor underwent a prior adaptive walk with a spectrum that favoured transitions. Analogous results hold for GC→AT bias: reversing the WT bias yields more beneficial mutations, and the effect is stronger with better-adapted WT populations (Fig. 4G, Fig. S12), suggesting that these results can be generalized for any axis of the mutation spectrum. Note that although our Δ*mutY* mutator reversed the WT Tv bias (explaining Fig. 2D), it reinforced the WT GC→AT bias, rendering the results in Figure 2E puzzling. However, ~40% of all GC→AT

mutations in the mutator are also transversions. Hence, our simulations suggest that the mutator's strong Tv bias (which opposes the WT bias) leads to the observed DFE differences.

Figure 4. Simulations show general benefit of reversing mutation spectra. (A) Change in mean population fitness, fraction of beneficial and deleterious mutations (f_b and f_d), and magnitude of beneficial and deleterious effects (s_b and s_d) over the course of adaptive walks of a WT ancestor. (B, G) Impact of altering the WT mutation bias on f_b at different points along the adaptive walk (Tv bias = Tv/(Ts+Tv) mutations; GC→AT bias = GC→AT/(GC→AT+AT→GC) mutations; grey dashed line: WT bias), as a function of the new mutation bias (mutator Tv or GC→AT bias). Line color: WT f_b at a given time; steeper lines (lighter colors): later evolutionary time, better adapted WT. (C–F) Waterfall plots show the impact of altering Tv bias (% change, Mutator – WT), as a function of WT (ancestor) and mutator (new) bias, for a well-adapted WT population (f_b = 0.04). (H–I) Impact of introducing a mutation bias in an unbiased ancestor ("anc"), as a function of the new bias (mutator Tv or GC→AT bias).



Interestingly, for an unbiased ancestor (Tv bias = 0.66 since 2/3 of possible mutations are transversions), introducing any bias is selectively neutral, regardless of how long it has evolved in the landscape (Fig. 4H–I). However, if the ancestor evolved with even a slightly biased spectrum (e.g. our WT), the population would have already sampled most beneficial mutations within the

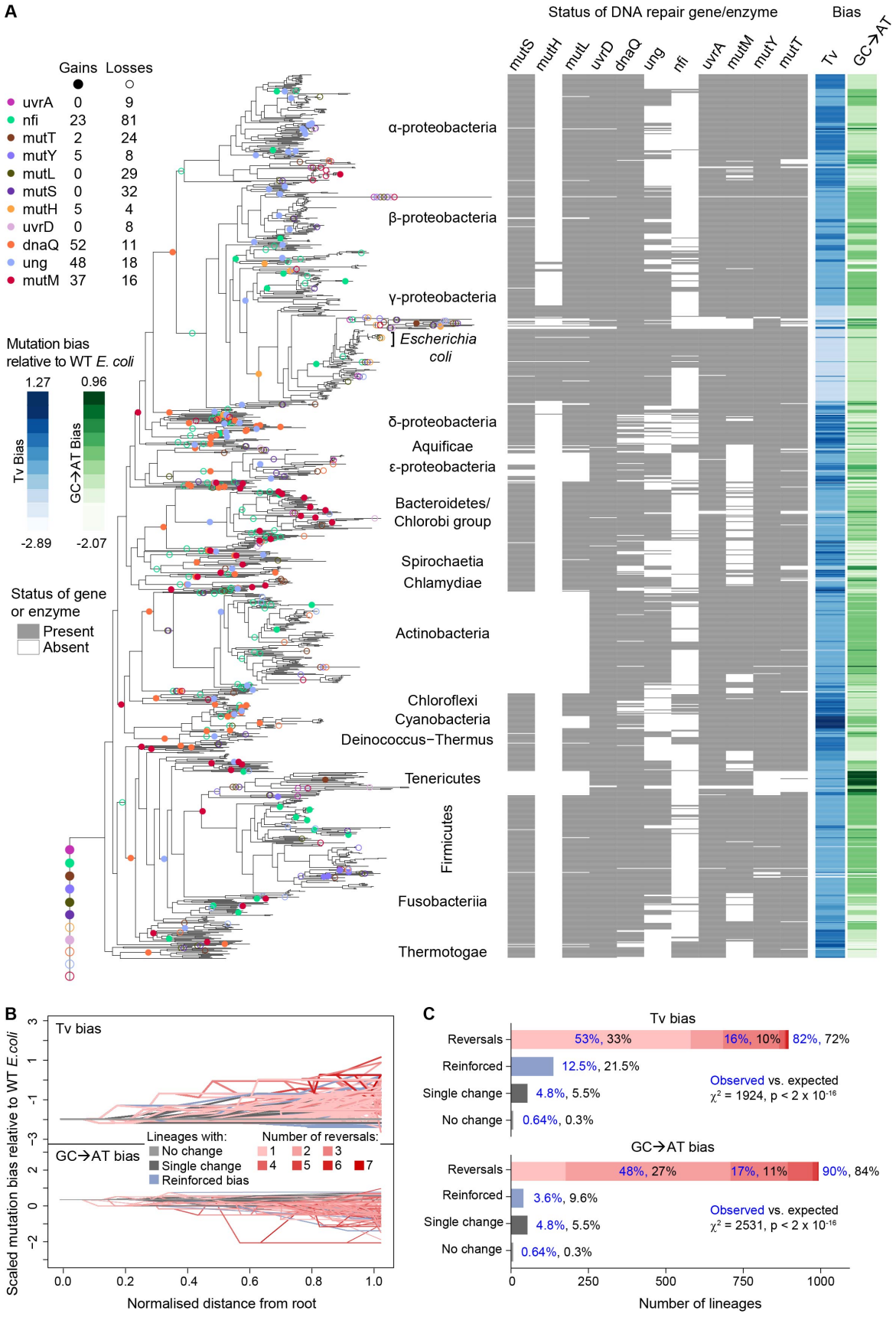
class favoured by the existing bias. Hence, a shift in the opposite direction (e.g. in our mutator) allows the population to explore mutational space that was not previously well-sampled, increasing the probability of finding new beneficial mutations. After a period of evolving with a new spectrum, a change in the fitness landscape (e.g. due to epistasis or environmental change) may again render a spectrum shift advantageous. Together, these results generate two predictions: (a) Since an unbiased state is not selectively favoured, biased spectra can evolve through drift (b) In due course, a reversal of the bias becomes beneficial (including a return to the unbiased state), and should occur frequently over long evolutionary periods.

Evolutionary transitions in DNA repair enzymes indicate mutation spectrum reversals in most bacterial lineages

Next, we inferred the long-term dynamics of mutation spectra by mapping evolutionary transitions in 11 bacterial DNA repair enzymes whose loss changes the Tv or GC→AT bias in *E. coli* (Table S16; limited data indicate consistent effects across species, Table S17). Identifying gene orthologues in 1093 extant bacterial genomes, we used ancestral reconstruction to infer enzyme gains and losses across the bacterial phylogeny (following (*28*); Fig 5A). Broadly consistent with previous work (*29*), we found frequent evolutionary transitions in most enzymes. Using the predicted set of enzymes for hypothetical ancestors (nodes), we estimated their Tv and GC→AT bias relative to WT *E. coli* and traced the change in relative bias over time (Fig. 5B, Fig. S13A). Over 80% lineages experienced a bias reversal during their evolutionary history, but few reinforced their ancestral bias through successive enzyme gain or loss events (Fig. 5C) – distinct from the expectation derived from stochastic simulations with similar enzyme transition rates (Fig. 5C, Fig. S13B). Notably, many more lineages experienced single reversals in Tv bias and double reversals in GC→AT bias than expected by chance (Tv bias: χ^2=805.3, $p<2x10^{-16}$; GC→AT bias: χ^2=453.5, $p<2x10^{-16}$; Fig. 5C, Fig. S13B). Hence, our results cannot be attributed to peculiarities of the enzyme set, tree topology, or number of evolutionary transitions; but likely arise from the impact of DNA repair enzymes on mutation rate and/or spectrum. While it is difficult to separate these effects, note that enzymes with a weak effect on the mutation rate but a strong impact on the spectrum are also highly dynamic (e.g. *ung* and *mutM*, whose impact on mutation spectrum we report here for the first time; Table S16). The relative biases estimated here should not be taken literally, since we cannot account for factors such as other repair enzymes whose effects on the mutation spectrum are unknown. Nevertheless, 33 additional genes from multiple DNA repair modules also show frequent evolutionary transitions (Fig. S14). Thus, sustained shifts in mutation spectra likely occur frequently, with long-term impacts on the pool of genetic variants available for selection.

Figure 5. Phylogenetic analysis shows frequent evolutionary transitions in DNA repair genes and reversals in mutation spectra. (A) Patterns of gene presence/absence in extant taxa, used to map evolutionary transitions in DNA repair genes on the bacterial phylogeny (total transitions are shown in the key), and infer the magnitude of Tv and GC→AT biases relative to WT *E. coli* in extant taxa (heatmaps). (B) Inferred trajectories of mutational biases across 1093 lineages (C) Observed number

of lineages showing spectrum shifts due to successive evolutionary transitions in repair enzymes (% values calculated as fraction across all lineages; blue=observed, black=expected values derived from stochastic forward simulations; Chi-square tests show comparisons).



Conclusions

Our work shows that pervasive mutational biases may evolve by genetic drift, whereas shifts in the strength and direction of bias could evolve under selection. Thus, a species' evolutionary history may be sufficient to explain the evolution of genomic bias, without invoking specific underlying mechanisms or selection favouring particular mutational classes. Our model also explains recent observations that GC-biased mutations in protein-coding genes are more deleterious in GC-rich genomes, and vice versa (*30*); such symmetrical results are otherwise difficult to explain. By recasting the puzzle of mutational biases as a problem of searching a vast and dynamic mutational space, our model allows generalization across any axis of the mutation spectrum. The rarity of species with unbiased mutation spectra (*17*) is thus counterintuitive but not surprising, since transitions from unbiased to biased spectra may occur easily and frequently via changes in DNA repair function. Hence, a biased spectrum should be the norm rather than the exception.

Our results suggest that shifts in mutation spectra and mutation rates are deeply entangled due to their strong association with mutator genotypes. The evolutionary dynamics of mutators are governed by their genetic load and supply of beneficial mutations (*25*, *31*, *32*). In a new environment (away from fitness optima), these parameters are primarily driven by mutators' high mutation rate (*33*) – particularly under strong selection (*34*) – allowing them to hitchhike with beneficial mutations (*35*). We show that the beneficial supply rate is strongly influenced by the mutation spectrum, especially in well-adapted populations under weak selection. Further, the advantage of a mutator genotype due to its increased mutation rate should be enhanced if it also reverses the ancestral bias; but diminished otherwise. Thus, mutation rate and spectrum may jointly govern the rise, persistence, and fall of mutators under selection (*25*, *36*). These effects may explain why mutators with relatively small increases in mutation rate are abundant in natural bacterial populations (*37*), and deserve further attention (*30*). Instances where spectrum shifts occur without rate change – e.g. during stress-induced mutagenesis (*16*, *20*) or the loss of repair enzymes like *ung* and *mutM* – offer a chance to untangle their evolutionary impacts.

Together with studies showing that mutation biases are pervasive and contribute to the genetic basis of adaptation (see Introduction) under diverse conditions (*38*, *39*), our results demonstrate that mutation spectra may be key to driving adaptation and innovation under myriad scenarios. At a point where the beneficial mutation supply is limited by the existing spectrum, an antiparallel (opposite direction on the same axis) or orthogonal jump (on a different axis of the spectrum) could allow further sampling of new beneficial mutations, facilitating rapid adaptation. Our phylogenetic analysis likely underestimated such evolutionary shifts in mutation spectra, which may occur frequently on shorter timescales via horizontal transfer and recombination (*40*), and drive polymorphism in spectra across natural bacterial isolates (*37*, *41*). We predict multiple cascading

effects of shifts in mutation spectra, including a reduction in the waiting time for beneficial mutations, decreased likelihood of mutational meltdown in populations evolving under drift, and distinct genetic pathways of adaptation. We hope that future work will test these predictions.

References and Notes:

1. P. L. Foster, H. Lee, E. Popodi, J. P. Townes, H. Tang, Determinants of spontaneous mutation in the bacterium *Escherichia coli* as revealed by whole-genome sequencing. *Proc. Natl. Acad. Sci. U.S.A.*, 201512136–10 (2015).
2. A. A. M. Al Mamun *et al.*, Identity and function of a large gene network underlying mutagenic repair of DNA breaks. *Science*. **338**, 1344–1348 (2012).
3. R. Maharjan, T. Ferenci, Mutational signatures indicative of environmental stress in bacteria. *Molecular Biology and Evolution*. **32**, 380–391 (2014).
4. D. Agashe, N. Shankar, The evolution of bacterial DNA base composition. *J. Exp. Zool.* **322**, 517–528 (2014).
5. H. Long *et al.*, Evolutionary determinants of genome-wide nucleotide composition. *Nat. ecol. evol.* **48**, 582 (2018).
6. J. Bohlin, O. Brynildsrud, T. Vesth, E. Skjerve, D. W. Ussery, Amino acid usage is asymmetrically biased in AT-and GC-rich microbial genomes. *PLoS ONE*. **8**, e69878 (2013).
7. A. Stoltzfus, L. Y. Yampolsky, Climbing Mount probable: Mutation as a cause of nonrandomness in evolution. *Journal of Heredity*. **100**, 637–647 (2009).
8. M. A. Streisfeld, M. D. Rausher, Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. *Evolution*. **65**, 629–642 (2010).
9. D. R. Rokyta, P. Joyce, S. B. Caudle, H. A. Wichman, An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nat Genet*. **37**, 441–444 (2005).
10. A. Stoltzfus, D. M. McCandlish, Mutational biases influence parallel adaptation. *Molecular Biology and Evolution*. **34**, 2163–2172 (2017).
11. J. F. Storz *et al.*, The role of mutation bias in adaptive molecular evolution: insights from convergent changes in protein function. *Phil Trans Roy Soc*. **374**, 20180238–10 (2019).
12. A. Couce, J. R. Guelfo, J. Blázquez, Mutational spectrum drives the rise of mutator bacteria. *PLoS Genetics*. **9**, e1003167 (2013).
13. J. L. Payne *et al.*, Transition bias influences the evolution of antibiotic resistance in *Mycobacterium tuberculosis*. *Plos Biol*. **17**, e3000265–23 (2019).
14. A. Bachar *et al.*, Point mutations in topoisomerase I alter the mutation spectrum in *E. coli* and impact the emergence of drug resistance genotypes. *Nucleic Acids Research*. **48**, 761–769 (2019).
15. S. M. Leighow, C. Liu, H. Inam, B. Zhao, J. R. Pritchard, Multi-scale predictions of drug resistance epidemiology identify design principles for rational drug design. *CellReports*. **30**, 3951–3963.e4 (2020).
16. T. Ferenci, Irregularities in genetic variation and mutation rates with environmental stresses. *Environ Microbiol*. **21**, 3979–3988 (2019).
17. V. Katju, U. Bergthorsson, Old trade, new tricks: Insights into the spontaneous mutation process from the partnering of classical mutation accumulation experiments with high-throughput genomic approaches. *Genome Biology and Evolution*. **11**, 136–165 (2018).
18. L. Y. Yampolsky, A. Stoltzfus, Bias in the introduction of variation as an orienting factor in evolution. *Evol. Dev.* **3**, 73–83 (2001).
19. R. P. Maharjan, T. Ferenci, A shifting mutational landscape in 6 nutritional states: Stress-induced mutagenesis as a series of distinct stress input–mutation output relationships. *Plos Biol*. **15**, e2001477–22 (2017).
20. D. Agashe, The road not taken: Could stress-specific mutations lead to different evolutionary paths? *Plos Biol*. **15**, e2002862–5 (2017).
21. C. A. Macken, A. S. Perelson, Protein evolution on rugged landscapes. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6191–6195 (1989).
22. M. Sane, J. J. Miranda, D. Agashe, Antagonistic pleiotropy for carbon use is rare in new mutations. *Evolution*. **72**, 2202–2213 (2018).
23. A. Eyre-Walker, P. D. Keightley, The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8**, 610–618 (2007).

24. E. Firnberg, J. W. Labonte, J. J. Gray, M. Ostermeier, A comprehensive, high-resolution map of a gene's fitness landscape. *Molecular Biology and Evolution*. **31**, 1581–1592 (2014).
25. S. Wielgoss *et al.*, Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proceedings of the National Academy of Sciences*. **110**, 222–227 (2013).
26. H. A. Orr, The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution*. **52**, 935–949 (1998).
27. A. Stoltzfus, Mutation-biased adaptation in a protein NK model. *Molecular Biology and Evolution*. **23**, 1852–1862 (2006).
28. G. D. Diwan, D. Agashe, Wobbling forth and drifting back: The evolutionary history and impact of bacterial tRNA modifications. *Molecular Biology and Evolution*. **35**, 2046–2059 (2018).
29. J. A. Eisen, P. C. Hanawalt, A phylogenomic study of DNA repair genes, proteins, and processes. *Mutation Research*. **435**, 171–213 (1999).
30. A. Couce, O. Tenaillon, Mutation bias and GC content shape antimutator invasions. *Nat Commun*. **10**, 3114–9 (2019).
31. K. Jain, A. James, Fixation probability of a nonmutator in a large population of asexual mutators. *Journal of Theoretical Biology*. **433**, 85–93 (2017).
32. Y. Raynes, P. D. Sniegowski, Experimental evolution and the dynamics of genomic mutation rate modifiers. *Heredity*. **113**, 375–380 (2014).
33. J. A. G. M. de Visser, C. W. Zeyl, P. J. Gerrish, J. L. Blanchard, R. E. Lenski, Diminishing returns from mutation supply rate in asexual populations. *Science*. **283**, 404–406 (1999).
34. O. Tenaillon, B. Toupance, H. Le Nagard, F. Taddei, B. Godelle, Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. *Genetics*. **152**, 485–493 (1999).
35. C. F. Gentile, S.-C. Yu, S. A. Serrano, P. J. Gerrish, P. D. Sniegowski, Competition between high- and higher-mutating strains of *Escherichia coli*. *Biology Letters*. **7**, 422–424 (2010).
36. L. Notley-McRobb, S. Seeto, T. Ferenci, Enrichment and elimination of mutY mutators in *Escherichia coli* populations. *Genetics*. **162**, 1055–1062 (2002).
37. I. Matic *et al.*, Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science*. **277**, 1833–1834 (1997).
38. A. V. Cano, J. L. Payne, Mutation bias interacts with composition bias to influence adaptive evolution. *bioRxiv*. **6** (2020), pp. 2157–48.
39. K. Gomez, J. Bertram, J. Masel, Mutation bias can shape adaptation in large asexual populations experiencing clonal interference. *bioRxiv*. **17** (2020), pp. e3000265–10.
40. E. Denamur *et al.*, Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell*. **103**, 711–721 (2000).
41. J. E. LeClerc, B. Li, W. L. Payne, T. A. Cebula, High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science*. **274**, 1208–1211 (1996).
42. T. Baba *et al.*, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*. **2**, 2460–11 (2006).
43. L. C. Thomason, N. Costantino, D. L. Court, *E. coli genome manipulation by P1 transduction* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2001), vol. 2.
44. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. **25**, 1754–1760 (2009).
45. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).
46. D. C. Koboldt *et al.*, VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*. **25**, 2283–2285 (2009).
47. N. F. Delaney *et al.*, Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*. *PLoS ONE*. **8**, e62957 (2013).
48. L.-M. Chevin, On measuring selection in experimental evolution. *Biology Letters*. **7**, 210–213 (2010).
49. S. R. Eddy, A new generation of homology search tools based on probabilistic inference. *Genome Inform*. **23**, 205–211 (2009).
50. N. Segata, D. Börnigen, X. C. Morgan, C. Huttenhower, PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun*. **4**, 2304–11 (2013).
51. J. P. Bollback, SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*. **7**, 88 (2006).
52. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*. **3**, 217–223 (2011).

53. W. Sung *et al.*, Asymmetric context-dependent mutation patterns revealed through mutation-accumulation experiments. *Molecular Biology and Evolution*. **32**, 1672–1683 (2015).
54. M. M. Dillon, W. Sung, R. Sebra, M. Lynch, V. S. Cooper, Genome-wide biases in the rate and molecular spectrum of spontaneous mutations in *Vibrio cholerae* and *Vibrio fischeri*. *Molecular Biology and Evolution*. **34**, 93–109 (2017).
55. W. Wei *et al.*, Mutation landscape of base substitutions, duplications, and deletions in the representative current cholera pandemic strain. *Genome Biology and Evolution*. **10**, 2072–2085 (2018).
56. J. R. Dettman, J. L. Sztepanacz, R. Kassen, The properties of spontaneous mutations in the opportunistic pathogen *Pseudomonas aeruginosa*. *BMC Genomics*. **17**, 27 (2016).
57. H. Long *et al.*, Background mutational features of the radiation-resistant bacterium *Deinococcus radiodurans*. *Molecular Biology and Evolution*. **32**, 2383–2392 (2015).
58. W. Sung, M. S. Ackerman, S. F. Miller, T. G. Doak, M. Lynch, Drift-barrier hypothesis and mutation-rate evolution. *Proceedings of the National Academy of Sciences*. **109**, 18488–18492 (2012).
59. S. Kucukyildirim *et al.*, The rate and spectrum of spontaneous mutations in *Mycobacterium smegmatis*, a bacterium naturally devoid of the postreplicative mismatch repair pathway. *G3: Genes| Genomes| Genetics*. **6**, 2157–2163 (2016).
60. A. Castañeda-García *et al.*, Specificity and mutagenesis bias of the mycobacterial alternative mismatch repair analyzed by mutation accumulation studies. *Sci Adv*. **6**, eaay4453 (2020).
61. N. Takemoto, I. Numata, M. Su'etsugu, T. Miyoshi-Akiyama, Bacterial EndoMS/NucS acts as a clamp-mediated mismatch endonuclease to prevent asymmetric accumulation of replication errors. *Nucleic Acids Research*. **46**, 6152–6165 (2018).
62. G. I. Lang, L. Parsons, A. E. Gammie, Mutation rates, spectra, and genome-wide distribution of spontaneous mutations in mismatch repair deficient yeast. *G3: Genes| Genomes| Genetics*. **3**, 1453–1465 (2013).
63. A. Farlow *et al.*, The spontaneous mutation rate in the fission yeast *Schizosaccharomyces pombe*. *Genetics*. **201**, 737–744 (2015).
64. B. A. Niccum, H. Lee, W. MohammedIsmail, H. Tang, P. L. Foster, The spectrum of replication errors in the absence of error correction assayed across the whole genome of *Escherichia coli. Genetics*. **209**, 1043–1054 (2018).
65. H. Lee, E. Popodi, H. Tang, P. L. Foster, Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proceedings of the National Academy of Sciences*. **109**, E2774–E2783 (2012).
66. P. L. Foster *et al.*, Determinants of base-pair substitution patterns revealed by whole-genome sequencing of DNA mismatch repair defective *Escherichia coli*. *Genetics*. **209**, 1029–1042 (2018).
67. I. J. Fijalkowska, R. M. Schaaper, Mutants in the Exo I motif of *Escherichia coli dnaQ*: Defective proofreading and inviability due to error catastrophe. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2856–2861 (1996).
68. M. M. Dillon, N. P. Rouillard, B. Van Dam, R. Gallet, V. S. Cooper, Diverse phenotypic and genetic responses to short-term selection in evolving *Escherichia coli* populations. *Evolution*. **70**, 586–599 (2016).
69. H. Long, S. F. Miller, E. Williams, M. Lynch, Specificity of the DNA Mismatch Repair System (MMR) and Mutagenesis Bias in Bacteria. *Molecular Biology and Evolution*. **35**, 2414–2421 (2018).
70. M. Blasius, U. Hübscher, S. Sommer, *Deinococcus radiodurans*: What belongs to the survival kit? *Critical Reviews in Biochemistry and Molecular Biology*. **43**, 221–238 (2008).
71. P. Hsieh, Molecular mechanisms of DNA mismatch repair. *Mutation Research*. **486**, 71–87 (2001).

Acknowledgements:
We thank Shyamsunder Buddh, Brian Charlesworth, Deborah Charlesworth, Joachim Krug, Krushnamegh Kunte, Saurabh Mahajan, Christopher Marx, and Mukund Thattai for discussion; Joshua Miranda for assistance with MA experiments; and the NGS facility at the Bangalore Life Sciences Cluster. **Funding:** We acknowledge funding and support from the National Centre for Biological Sciences (NCBS-TIFR), the Council for Scientific and Industrial Research India (Senior Research Fellowship to MS), the University Grants Commission of India (Senior Research Fellowship to GDD), the Department of Science and Technology India (KVPY Fellowship to BAB),

the Natural Sciences and Engineering Research Council of Canada (LMW), and the Wellcome Trust (GDD is supported by grant 210585/B/18/Z to Robert B. Russell; and DBT/Wellcome Trust India Alliance grant IA/I/17/1/503091 to DA). We also thank the International Centre for Theoretical Sciences (ICTS) for supporting the Bangalore School on Population Genetics and Evolution (code: ICTS/popgen2020/01), where this collaboration was initiated. **Competing interests:** Authors declare no competing interests. **Author contributions:** MS designed and conducted experiments, analysed data, and drafted the manuscript. GDD designed and conducted phylogenetic analyses. BAB conducted experiments and analysed data. LMW designed and conducted simulation analyses. DA conceived and designed experiments and phylogenetic analyses, analysed data, obtained funding, and wrote the manuscript. **Data and materials availability:** All data used for experimental and phylogenetic analysis are available as Supplementary Information files. Genome sequences will be deposited in GenBank, and simulation code is available on Github (https://github.com/lmwahl/MutationSpectrum).

List of Supplementary Materials:

Materials and Methods
Figures S1-S14
Tables S1-S17
References 42-71