

# Mutation bias and adaptation in bacteria

James S. Horton\* and Tiffany B. Taylor

## Abstract

Genetic mutation, which provides the raw material for evolutionary adaptation, is largely a stochastic force. However, there is ample evidence showing that mutations can also exhibit strong biases, with some mutation types and certain genomic positions mutating more often than others. It is becoming increasingly clear that mutational bias can play a role in determining adaptive outcomes in bacteria in both the laboratory and the clinic. As such, understanding the causes and consequences of mutation bias can help microbiologists to anticipate and predict adaptive outcomes. In this review, we provide an overview of the mechanisms and features of the bacterial genome that cause mutational biases to occur. We then describe the environmental triggers that drive these mechanisms to be more potent and outline the adaptive scenarios where mutation bias can synergize with natural selection to define evolutionary outcomes. We conclude by describing how understanding mutagenic genomic features can help microbiologists predict areas sensitive to mutational bias, and finish by outlining future work that will help us achieve more accurate evolutionary forecasts.

## INTRODUCTION

Each bacterial genome boasts an arsenal of genes that provide robustness against changing environments, and bacterial cells can acquire access to yet more phenotypes through the horizontal transfer of genetic material [1]. However, when exposed to new selective challenges, novel adaptive phenotypes are often generated through the process of mutation. With typically large population sizes and rapid generational times, bacterial populations are hotbeds for adaptation through mutation, allowing them to rapidly evade extinction against numerous threats.

The adaptive capability of microbes through mutation is well documented and well understood – it is a key reason why we struggle to eradicate pandemics [2] and why antimicrobial resistance is such a potent threat [3]. As such, many researchers are dedicated to elucidating the consequences of mutation, such as how gene regulatory networks are re-organized and how functions of protein products are altered following a mutational change. Complementary to these efforts is the study of the mutation process itself. Because if we can understand the causes of mutations, we will be able to anticipate their arrival. This will provide us with the ability to form accurate short-term evolutionary forecasts – allowing us to be proactively poised to deal with the consequences of mutation as soon as they occur.

Mutation is the primary means of generating genetic variation and is often described as a stochastic force. This means that the process of mutation relies on chance events. However, chance does not mean all mutations are equally likely to happen. Instead, mutations are biased. They occur at different rates, with certain types of mutations being more common than others (e.g. a transition substitution bias [4]), and certain locations in the genome being more mutable than others (e.g. those farther from the replication origin [5]). As such the raw material that generates genetic variation and facilitates adaptation is provided unevenly, with certain mutations appearing in the population more often than others [6]. These effects, while commonly overlooked, can have profound impacts on the course of adaptation and evolution.

The relationship between mutation bias and bacterial adaptation can sometimes be antagonistic, as the types of mutations that are selectively beneficial may be in direct conflict with those promoted by mutation bias. For example, bacterial genomes may universally be biased to mutate from GC → AT, and yet some bacterial genomes continue to be GC-rich [7]. Similarly in viruses, Rice *et al.* found that the SARS-CoV-2 genome exhibits selection against uracil content despite changes to uracil

Received 02 May 2023; Accepted 11 October 2023; Published 09 November 2023

**Author affiliations:** <sup>1</sup>Milner Centre for Evolution, Department of Life Sciences, University of Bath, BA2 7AY, UK.

**\*Correspondence:** James S. Horton, jsh51@bath.ac.uk

**Keywords:** mutation bias; adaptation; bacterial evolution; mutation hotspot; stress response; evolutionary forecasting.

One supplementary table is available with the online version of this article.

001404 © 2023 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

being promoted by mutation bias [8]. However, when mutation rates are higher at a nucleotide position that can confer an adaptive advantage, mutation bias can synergize with selection to make certain outcomes more likely than others. These mutational biases can be incredibly local and specific, with adaptive outcomes being defined by localized mutational ‘hotspots’ within single genes [9]. Cano *et al.* have recently shown that known mutational biases are reflected in adaptive outcomes for *Saccharomyces cerevisiae*, *Escherichia coli* and *Mycobacterium tuberculosis* [10]. This study highlights that the uneven production of mutations owing to bias can synergize with selection and help to steer the course of evolution.

With recent research revealing the impact of mutation bias on adaptation within the bacterial domain [10, 11], there is an opportunity for microbiologists from various disciplines to consider what role mutation bias plays in the evolution of their organism(s) of interest. In this review, we will describe the features that cause mutation bias in bacterial genomes and discuss how we can use this knowledge to identify areas more prone to mutation. We will also outline the contexts when bias can be a defining factor in bacterial adaptation and highlight how the evolution of pathogenic microbes in the clinic are susceptible to agents of mutation bias. We will conclude by discussing future research avenues that will facilitate the discovery and consequence of bias throughout the bacterial genome.

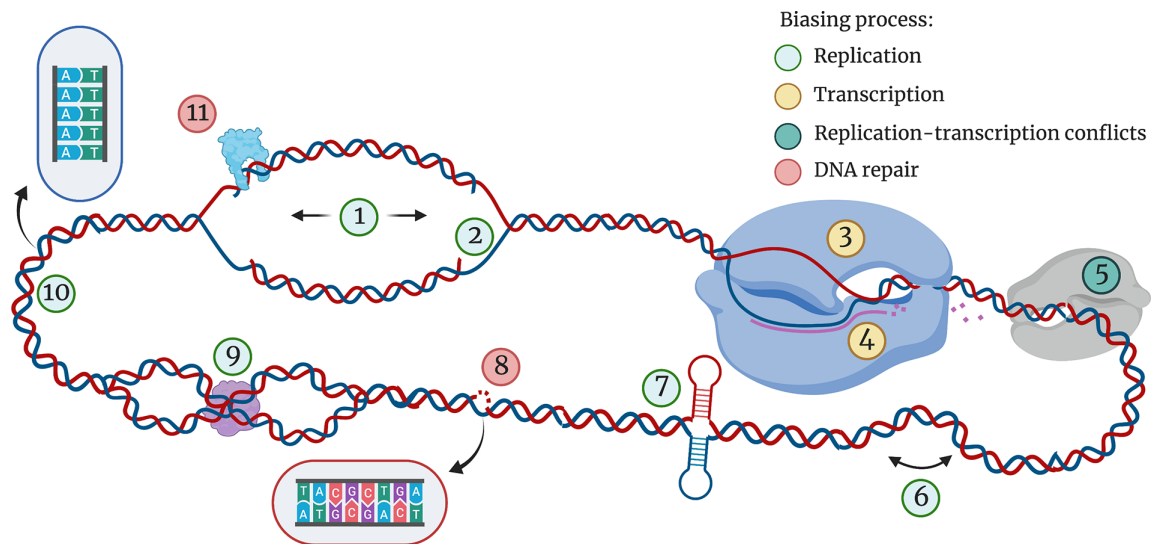
## THE HEART OF THE MATTER – LOCAL AND GENOME-WIDE CAUSES OF MUTATION BIAS

Throughout this work we will often use the terms mutation rate, mutation bias and mutational hotspot. It is worth describing these terms first in more detail. Mutation rate is used to describe mutation in absolute, discrete terms. For example, *P. aeruginosa* has been quantified to have a mutation rate of  $9.30 \times 10^{-11}$  per nucleotide position per generation [12, 13]. This description provides the average genomic rate of mutation and therefore allows us to calculate the expected number of genomic mutations over time with consideration to population size. The genomic mutation rate multiplied by the number of replicating genomes in a population determines how often mutations are expected to occur, which is described as mutation supply. Mutation bias describes the relative differences in mutation rates between positions or mutation types. For example, a null assumption is that transversion mutations should occur at a ratio of 2:1 versus transitions, because 8/12 of the possible base substitutions are transversions. (This is an over-simplification [14], but suits our purposes for this example.) But when we see that transitions are more common than this in sequencing data, and we have controlled for the effects of selection [14, 15], we can state that there is a mutation bias toward transitions as they are relatively over-represented [14]. When considering adaptation, mutation bias details instances when the list of mutational targets (the total pool of mutations capable of producing a phenotype of interest) is skewed to produce certain mutations more than would be expected by chance. Biases mean that rates of mutation fluctuate throughout the genome, resulting in mutational coldspots that mutate less than the genome average, and mutational hotspots that are typically used to describe positions that mutate at much higher rates than the genome average [16–19]. Mutation rates are the product of many interacting mechanisms operating in tandem (Fig. 1), of which each introduces its own bias. Therefore, if we observe a change in mutation rate due to a mutation or an environmental trigger, this inevitably incurs a change in bias because not all mutagenic mechanisms will be equally affected.

Mutation is the result of chance mistakes that happen as genetic material is being replicated, transcribed, recombined and repaired (Fig. 1). The DNA–protein interactions driving these mutagenic mechanisms are therefore the harbingers of change. In its inactive or dormant state DNA exists as a stable, canonical duplex, which can also be modified by epigenomic features that help to protect DNA against environmental damage [20]. However, DNA does not perpetually exist in a dormant state. It is constantly unravelled and unwound to be used as a template for gene expression and to create copies of itself. These modifications expose DNA – they open the duplex into single strands, upon which toxins in the environment such as radicals and radiation can cause damage to the vulnerable sequence [21, 22].

Many mutational mechanisms are affected by the process of replication (Fig. 1). Bacteria possess a single origin of replication (OriC) and replicate DNA bi-directionally (theta replication) symmetrically outward from the origin to the replication terminus. Not only do bacterial genomes have a designated start and end point for replication, they also have strand-associated patterns where the leading strand and lagging strand are processed differently, with the leading strand copied continuously and the lagging strand copied in segments [23]. Both replication timing and the replicated strand are associated with mutation bias (Fig. 1). The fidelity of replication has been documented to be higher closer to the replication origin in *P. fluorescens*, *E. coli*, *S. enterica*, *V. fischeri* and *V. cholerae* [5, 24–26]. Further, lagging strand synthesis has been associated with higher fidelity in *E. coli* [27], but the lagging strand has also been observed to be more vulnerable to cytosine deamination [28]. Therefore, mutations can be disproportionately generated across large swathes of the genome, with mutation rates increasing around a megabase from the replication origin and changing depending on replicative strand.

While genome organization and strandedness can cause mutation rates to differentiate throughout the genome, mutational biases can fluctuate on a local level due to interactions between the replication complex and the local nucleotide context. For example, mononucleotide repeats (e.g. CCCCC) can cause the polymerase to slip and generate an insertion or deletion (indel) [12, 16, 29], while inverse repeat sequences (e.g. ATCGAC-n-GTCGAT) cause the generation of secondary structures that cause the polymerase to stall and produce a polymorphism [30] or double-stranded break [31]. The local nucleotide neighbourhood



**Fig. 1.** Features affecting regional and localized mutation bias throughout the bacterial genome. Mutation bias primarily results from replication [157], transcription [158], collisions between replication and transcription complexes [159], and repair enzyme activity [160], as denoted by the key. Labelled features are described per the following format: *n*. Mechanism that causes mutation bias. Location in the genome impacted by mechanism. 1. DNA polymerase stalling. Affected by genomic position/replication timing; polymerase fidelity is typically higher close to *Oric* [5, 24–26]. 2–3. Cytosine deamination. 2. More frequent on lagging ssDNA strand [28]. 3. More frequent on non-transcribed ssDNA strand [34–36]. 4. Transcription-coupled repair efficacy. Higher on transcribed strand [37]. 5. Head-on collisions between RNA polymerase and the replisome. Occur at genes encoded on the lagging strand [161–163], however the extent to which these events drive mutation is contested [41–43, 45]. 6. Deletion following recombination events. Occur at homologous direct repeats [114–117]. 7. Hairpin or cruciform secondary DNA structures. Formed at palindromic and quasi-palindromic sequences [109–112]. 8. Mismatch repair efficacy. Affected by nucleotide triplet of target sequence [5, 29, 44, 54]. 9. Nucleoid-associated proteins. DNA with high superhelical density may be more mutable during replication [20, 59, 60] but bound DNA is likely less sensitive to homologous recombination [58] and small mutagenic compounds [20]. 10. DNA polymerase slippage. Occur more frequently at homopolymer tracts [12, 29, 113]. 11. Presence of repair enzymes. Absence of mismatch repair generates ‘mutator’ alleles that often cause transition biases [5, 118, 119]; presence of repair enzymes increases mutagenesis at alternate non-canonical DNA secondary structures [164]. Figure created with BioRender.com.

will often affect replication fidelity disproportionately depending on which strand the nucleotides are situated on. Homopolymeric tracts of guanine cause substitution mutations more frequently when on the leading strand [17], and mutations from inverse repeat DNA hairpins may more readily form on the lagging strand as it exists as a single strand for longer [23, 32]. However, more recent research has found that long inverted repeats can form hairpins on both strands, forming instead what is called a cruciform structure [31]. Local nucleotide context has been implicated in many mutational biases operating over small areas (e.g. small indels and single nucleotide polymorphisms) and so has been described as the major determinant of biases in bacterial systems [33].

Transcription is the other major process resulting in transient single-stranded DNA. Like replication, the two DNA strands are processed differently during transcription. RNA polymerase copies from one strand only – the transcribed strand – while the non-transcribed strand remains unbound as free single-stranded DNA during this process. As a result, the non-transcribed strand is more vulnerable to cytosine deamination [34–36]. Similarly, components of nucleotide excision repair, which correct damage endured during transcription, operate alongside RNA polymerase and so repair with higher fidelity on the transcribed strand [37, 38].

As transcription does not stop during periods of replication, there is a possibility that the two polymerase complexes can collide. If a gene is oriented co-directionally with replication, then DNA polymerase and RNA polymerase will process DNA in the same direction. Most genes, including most essential genes, are transcribed co-directionally with the replisome [39]. However as the replisome progresses faster than the RNA polymerase complex, co-directional collisions can still occur and cause genome instability [40]. Genes that are transcribed antagonistically to the replisome instead result in head-on collisions between the polymerase complexes [39]. This causes more potent genome instability, and head-on genes have been found to exhibit increased mutation rates in *B. subtilis* [41].

However, it remains unclear to what extent head-on collisions act as sources of spontaneous mutagenesis. Studies on *B. subtilis* and *E. coli* have found that patterns showing increased mutation in head-on genes can instead be explained by local nucleotide features that cause bias [42, 43]. Schroeder *et al.* found that nucleotide composition bias, i.e. that some nucleotides are found more on one strand than the other, can explain the increased rate of base substitutions found in head-on genes in *B. subtilis* [43]. (A rebuttal study however argued that this was only the case in laboratory evolved strains and did not explain substitutions

in natural populations [44].) Similarly, Foster *et al.* observed that a subset of antagonistically oriented tRNA genes contained mononucleotide runs of length 5–8 bp, which are hotspots for indels, and this explained the overall enrichment for indels in head-on genes in *E. coli* [42]. The same study also found no positive correlation between rates of mutation in head-on genes and levels of expression, which would increase the opportunity for collision [42]. Sankar *et al.* designed an experimental system to contrast head-on and co-directional mutagenesis in *B. subtilis* [45]. They introduced the *thyP3* gene in both orientations with respect to replisome progression and placed the gene under control of an inducible promoter. They then contrasted the rate and spectrum of mutations observed in *thyP3* using an assay that captured loss-of-function mutations in the gene. Indel mutations were enriched at the promoter sequence and 5' half of the gene when it was co-directional with replication and enriched at the 3' half of the gene when in the head-on orientation. However, neither orientation nor induction of transcription affected base substitution rates in the coding region, but the head-on orientation did increase substitution rates in the promoter [45]. Overall, while head-on genes cause genomic instability in bacteria and may increase the rate of indel mutations and *cis*-regulatory substitutions, it is unclear whether they are a potent mutagenic force [39].

The final major source of mutational bias comes from DNA repair. Mismatch repair, which repairs errors incurred during replication, is a prominent enforcer of potent biases. Mismatch proteins or their equivalents are ubiquitous in bacteria [46]. Yet mutator variants, which are defective for mismatch repair, often evolve in bacterial pathogens that are under strong selection for novel adaptive mutations [47]. These include *S. enterica* pathogens [48], uropathogenic *E. coli* isolates [49] and strains of *P. aeruginosa* evolving in the cystic fibrosis lung [50]. Mutators have been widely reported to showcase strong biases. In the case of the highly conserved repair enzymes *mutS* and *mutL*, their inactivation drives a sharp increase in the rate of transition mutations and small indels [51]. In contrast, *E. coli* mutants defective in *mutT* or *mutY*, which remove oxidated guanine (8-oxo-G) that cause G:A mispairs [51], instead exhibit a bias toward A:T → C:G and C:G → A:T transversions, respectively [52, 53].

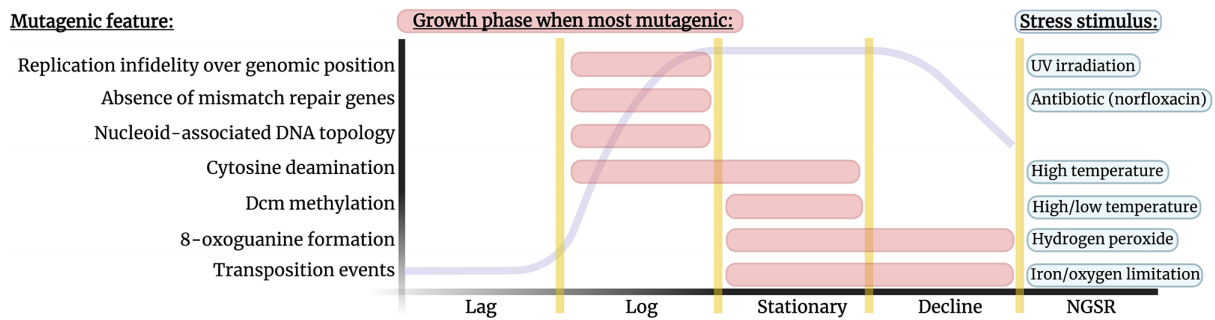
There is also a relationship between mutation frequencies found in mutator lines and local nucleotide sequence. Different rates of mutation have been observed depending on the nucleotide triplet, where the rate of mutation of a focal nucleotide varies depending on its immediate nucleotide neighbours [54]. While the most mutable nucleotide triplet can differ depending on the species, this trend has been reported for numerous mutator lines including in *B. subtilis* [44], *E. coli* [29], *P. fluorescens* [5] and *P. aeruginosa* [12]. Foster *et al.* highlighted that mismatch repair efficiency in *E. coli* can fluctuate by around 43-fold depending on the nucleotide triplet, showing that the fidelity of mismatch repair enzymes is affected by local nucleotide sequence [29]. Therefore, the major sources of mutational biases in the genome are driven by the interaction between the local nucleotide substrate and the protein complex interacting with it, whether that may be the replication complex, the transcription complex or repair enzymes.

## SETTING THE SCENE – DRIVERS OF MUTATION BIAS OUTSIDE THE BACTERIAL GENOME

So far we have examined the DNA–protein interactions operating over local and broad genomic regions that cause elevations in mutation rates and amplify mutational biases. And yet we can evolve clonal genomes and not always observe the same mutation types. For example, in the study introduced above, Foster *et al.* noted that the type of transition mutations observed in mutator lines changed depending on both growth media and incubation temperature [29]. Why? Because mutagenic mechanisms do not operate consistently *ad infinitum*. Instead, the rates of genomic replication, the expression of its genes, and the exposure to toxic agents each vary depending on external triggers from the environment [55]. This means that genomic features associated with causing mutation can in many instances be described as having mutagenic potential, but this potential will change depending on the phase of the bacterial growth cycle and the presence of other stressors. Therefore, to gain a deeper understanding of which features are driving mutation within our genomes of interest, we must consider them in tandem with growth-associated mutagenesis [56] and stress-induced mutagenesis [57].

Some examples of growth phase-sensitive mutagenic mechanisms are shown in Fig. 2. The traits depicted in this figure reflect the literature and therefore are based on traits observed primarily in model laboratory species, especially *E. coli* (see Table S1, available in the online version of this article). Note therefore that there are genomic backgrounds for which some of these principles do not hold. In some cases, mutagenic mechanisms operate during all phases of growth, but become more mutagenic during a particular growth phase. For example, fluctuating replication fidelity may drive bias any time replication occurs, but this will be felt more keenly during exponential phase when replication occurs most frequently in most bacteria. In other cases, genomic features only become mutagenic in certain growth phases, whereas they can suppress mutation in others. For example, nucleoid-associated proteins that bind to DNA and alter its topology actively protect bound DNA from mutation during periods of stress and exposure to toxic agents [20, 58]. But multiple studies have found evidence that bound and superhelical DNA structures increase mutagenesis during replication, likely due to their interference with replicative machinery [20, 59, 60]. Similarly, Abundiz-Yañez *et al.* have recently found that in *B. subtilis*, functional diadenylate cyclases suppress mutagenesis in growing cells, but during nutrient stress of the stationary phase this trend is reversed and the cyclases instead promote mutation [56].

The time spent in each growth phase and the likelihood of encountering other stressors will fluctuate substantially depending on the colonized *in vivo* niche or the conditions of the laboratory experiment. Environments in clinical samples are more complex, with colonized organs such as the gut or the cystic fibrosis lung containing a myriad of heterogenous niches that vary from patient



**Fig. 2.** Mechanisms of mutation bias are affected by the bacterial growth cycle and other environmental stressors. A selection of mutagenic features (left) vary in their mutagenicity depending on the bacterial growth phase (heightened mutagenicity highlighted by red bars). The mutagenicity of each feature can additionally be amplified by non-general stress response (NGSR) triggers, as shown on the right. The effect of environment and growth phase on the listed features reflect studies performed on common model organisms, however there will be genomic backgrounds for which these trends do not apply. For example, while many bacteria most actively replicate during the logarithmic phase, the cyanobacteria *Synechococcus elongatus* performs intensive replication during the lag phase [165], meaning replication infidelity may be most mutagenic during an earlier phase. Full details of the publications informing this figure are provided in Table S1. Figure created with BioRender.com.

to patient [61, 62]. In contrast, lab-based systems provide more control over growth and stress factors. These factors will change depending on whether we use chemostats or turbidostats, rich or poor nutrient broths and agars, agitated or static cultures. Each choice will affect the generational times spent in each of the growth phases and the evolving population's subjection to stress. Turbidostats for example can ensure that there is always an excess of nutrients facilitating exponential growth, whereas chemostats can prolong the stationary phase for many generations [63]. The stationary phase is associated with the general stress response, where gene expression changes to protect the cell [64] at the expense of increased mutation rate [65]. A key driver of stress-induced mutation during this phase is the upregulation of the sigma factor RpoS [66]. This protein increases mutation rate during stationary phase but can do so unevenly, promoting certain mutation types more than others (e.g. G:C → T:A rates increase but G:C → C:G rates do not [67]) and therefore drives mutation bias.

As nutrient-deficiency encountered by cells in the stationary phase drives stress-induced mutagenesis, environments that facilitate highly dense microbial populations experience greater nutrient limitation and thus exhibit a higher stress-induced mutational burden [68]. This suggests that growth rate and eventual population density should be positively correlated with mutation rate, but studies investigating both bacteria and yeast have found that they are instead often negatively correlated [69, 70]. This has been demonstrated by Krašovec *et al.* to be owed to a phenomenon termed density-associated mutation rate plasticity (DAMP) [70]. A highly mutagenic compound – oxidated guanine (8-oxo-G) – increases in abundance in the stationary phase (Fig. 2) but is combated by expression of MutT in *E. coli* [70], which prevents the misincorporation of 8-oxo-G opposite adenine in newly synthesized DNA strands [53]. Highly dense populations enjoy an increase in the number of 8-oxo-G scavengers and so exhibit lower mutagenesis from 8-oxo-G. Therefore there is a trade-off between mutations resulting from growth-associated population density and stress-induced mutations resulting from nutrient limitation [68].

Whether bacteria augment their environment during colonization of a new niche will also affect growth-associated and stress-induced mutagenesis. Agitating cultures coated with anti-adhesive compounds help to prevent the formation of biofilms [71], but biofilms are facilitated in static broth cultures [72] and on solid growth media. The formation of biofilms alters the local environmental niche for cells within the superstructure, within which the lack of nutrients and increase in toxin concentration can enable the SOS response and promote mutation [73]. In *E. coli*, this non-general stress response increases the expression of specialized DNA polymerases (Pol II, Pol IV and Pol V) that can perform translesion synthesis but are error prone [74]. A recent review by Joseph and Badrinarayanan summarized how translesion synthesis can promote mutational biases, as different mutation types were reported depending on the type of lesion and the error prone polymerase interacting with it [75]. Taken together, these reports highlight how simple changes to the experimental regime can have strong ramifications on which mutagenic mechanisms will be most active during evolution and which mutational biases they will promote.

In addition to variations of nutrients and spatial niches, bacteria are faced with selective challenges that can be highly threatening to population survival. These include the presence of toxins, which can trigger non-general stress responses and promote different types of mutation (Fig. 2). But perhaps a more notable mutagenic environmental trigger is the challenge of antibiotics, which is a common application in both clinical and laboratory strains. Antibiotics have long been known to cause stress that increases mutation rate in bacteria [76]. A study from Long *et al.* highlighted that certain antibiotics – in this case, norfloxacin – can also promote mutational biases, with G:C→T:A transversions becoming more frequent in *E. coli* with increasing dosage of the antibiotic [77]. Therefore, even the treatment method can itself alter bias and change the resultant mutational spectrum.

## THE SILENT PARTNER – WHEN MUTATION BIAS CAN GUIDE ADAPTATION IN THE LAB AND THE CLINIC

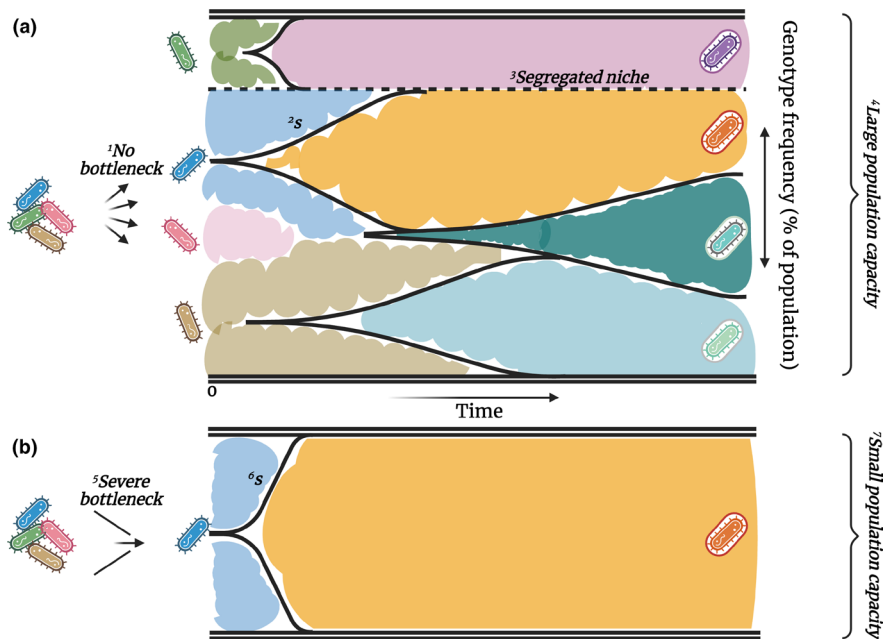
In evolving populations where standing genetic variation exists, adaptive alleles compete for dominance. Either the fittest of these genotypes reaches fixation (i.e. the genotype rises to a high frequency) through natural selection [78], or a population bottleneck reduces diversity such that chance becomes a stronger influencer [79]. Alternatively, multiple genotypes can persist and co-exist in the population [80–82], providing an opportunity for a secondary mutant descendent to evolve and achieve fixation [83]. Such competition between standing genetic variation provides opportunity for a ‘soft’ selective sweep, coined by authors Hermisson and Pennings [84]. The authors described soft sweeps as more likely to be observed when the potential fitness gain from mutations is modest, or, if the potential fitness gain is high, the mutational supply (or starting mutational diversity) is also high. In contrast, a ‘hard’ selective sweep involves the *de novo* arrival into the population of a single selectively advantageous allele that sweeps to dominance before a rival adaptive mutation appears [84]. In this latter ‘origin-fixation’ model [85], where the first adaptive mutation to occur will reach fixation, mutation bias can become a defining adaptive force [86, 87].

The context-dependency of mutation bias influence can be well illustrated by Cano *et al.* [10], who found that known biases are reflected in adaptive outcomes for three diverse species of microbes [10]. However, the authors also showed that using mutation bias as a predictive tool for adaptive evolution becomes less reliable as mutational supply and mutational target size increases [10]. This finding can be explained by re-visiting Hermisson and Penning’s earlier assertions [84]; as we increase the rate at which new mutations arrive in the population, or a larger pool of mutations become adaptive, we increase the chance of multiple adaptive genotypes co-existing in the population before one reaches fixation. This shifts the adaptive dynamics from facilitating a ‘hard’ to a ‘soft’ selective sweep, lessening the impact of mutation bias. Therefore, for mutation bias to be at its most influential, we search for instances when adaptive mutations occurring at biased rates can undergo a hard sweep to dominance.

As displayed in Fig. 3, for mutation bias to be at its most impactful we search for instances where (1) there is little to no standing adaptive genetic variation when selection is imposed, (2) an adaptive mutant can reach fixation before a competitor appears (i.e. mutation supply is low relative to the selective benefit of an adaptive mutation), (3) competition can take place across the population as sub-populations are not segregated in isolated niches. And that at least a portion of the genomic positions exhibiting biased mutation can prove adaptive. Determining how ubiquitous hard selective sweeps [criteria (1) and (2)] are in nature has been the subject of fierce debate [78, 88, 89], but is safe to say these criteria will not be satisfied under all adaptive scenarios. We can, however, find opportunities that meet these requirements during instances of microbial infection, where mutation bias has been reported [4, 90] and is sometimes imposed by the clinical treatment itself [91].

During infection, populations undergo severe bottlenecks that limit mutational diversity [92] as they pass through various defensive mechanisms including mucosal barriers [93] and organ filters [94]. A depletion of genetic diversity does not in itself guarantee a hard selective sweep (and by extension, a defining role for mutation bias). First, populations may begin as commensal bacteria and acquire renewed genetic diversity while evolving under purifying selection, before switching to a pathogenic phenotype [95]. Second, subsequent rounds of bottlenecks render populations susceptible to genetic drift [96], which can introduce more chance dependence into an adaptive outcome. Third, clinical isolates can evolve mutator phenotypes to cope with the selective burden imposed by host defences [97]. This increases the rate at which a beneficial (or deleterious) mutation will be realized, but also provides more opportunity for multiple adaptive genotypes to be simultaneously present in the population and compete for fixation [98]. However, modelling and experimental data has shown that when mutation supply is maintained at wild-type levels, the early arrival of an adaptive mutation can dominate and keep mutational diversity low [99]. Therefore, non-mutator, obligate pathogens undergoing limited numbers of bottlenecks can readily satisfy criteria (1) and (2). This means that microbial adaptation in the clinic can be highly guided by mutation bias, permitted genomic biases operate at adaptive positions [4, 90, 91].

While mutation bias in clinical isolates can be difficult to measure, a researcher can assume much greater control of adaptive variables in the laboratory. A recent study by Schenk *et al.* demonstrated the increased influence of bias when adaptive mutations can sweep to fixation before a rival mutation appears [requirement (2) above] [100]. They observed that in their *E. coli* strains, mutations resulting from structural variations (SVs) exhibited higher rates than single nucleotide polymorphisms (SNPs). This meant that SVs were biased to occur sooner than SNPs in their populations, despite SVs on average yielding lower fitness than SNP mutations [100]. The *E. coli* strains were propagated for approximately 500 generations in cultures facilitating either small or large population sizes. The authors observed that adaptive mutations in smaller populations were more likely to undergo a hard sweep to fixation. They also found that SVs were more likely to occur earlier in both population sizes, but the biased mutational class was more likely to reach fixation in smaller populations [100]. Owing to the number of cells, larger population sizes have a larger mutational supply ( $\mu$ , Fig. 3) and it takes longer for an adaptive mutant to outcompete inferior genotypes and reach fixation (affecting  $s$ , Fig. 3). The study above demonstrates that even when a mutational class is biased to occur, the ability for this bias to determine adaptive outcomes is greatly influenced by the population dynamics (Fig. 3).



**Fig. 3.** The role of mutation bias in determining adaptive outcomes is influenced by initial genetic diversity, the selective benefit ( $s$ ) of novel mutations, the rate of mutation supply ( $\mu$ ), the number of isolated adaptive niches, and total population capacity. Shown in the two panels are toy Muller plots, which display how unique genotypes (individual colours) change in frequency (height of colour on y axis) over time (x axis). Colours that emerge after timepoint 0 represent evolved genotypes. Both scenarios show evolution to a novel environmental niche and include a genotype that exhibits potent mutation bias (blue), i.e. one specific adaptive genomic position exhibits a high relative localized mutation rate, which allows it to mutate rapidly (near timepoint 0) to a specific descendent genotype (orange). However, the orange genotype only dominates in one scenario. (a) Mutation bias does not play a dominant role in final genetic diversity: <sup>1</sup>No bottlenecking means there are multiple genotypes initially evolving in the new niche. The biased genotype (blue) evolves rapidly, but the <sup>2</sup>relative difference in fitness ( $s$ ) between the descendent genotype (orange), its ancestor (blue), and other competing genotypes is not marked enough for it to rapidly dominate the population. This allows alternative adaptive mutants to appear from other genotypes, and for a member of the ancestor population to acquire a mutation (at rate  $\mu$ ) to a less common, but fitter genotype (teal) before the population goes extinct. The inability of the orange genotype to dominate is compounded by <sup>3</sup>physically separated niches that allow alternate mutants to evolve in isolation (purple), <sup>4</sup>and a large population capacity that increases the opportunity for multiple adaptive mutations to appear. (b) Mutation bias is the defining force in deciding the adaptive outcome: <sup>5</sup>A severe bottleneck reduces starting genetic diversity to one genotype. Upon evolving at the biased position, <sup>6</sup>the strong relative increase in fitness, compounded by <sup>7</sup>a smaller population capacity without segregated niches, allows the evolved genotype to rapidly dominate the population pool before the wild-type genotype (blue) can mutate at rate  $\mu$  to the more fit alternative genotype (teal in panel a). Scenario (b) is common in microbial evolution experiments, whereas clinical isolates will evolve in contexts somewhere between these two extremes. Figure created with BioRender.com.

In some cases, efforts have been made to mimic clinical settings to uncover which variables are key to driving adaptation, including bottlenecks [100, 101]. Garoff *et al.*, for example, found that soft selective sweeps were prevalent in evolving populations of *E. coli* when a large population was transferred through bottlenecks [101]. However, when smaller populations were transferred through bottlenecks, initially soft sweeps eventually ‘hardened’ [79] into hard selective sweeps as genetic diversity was sequentially lost [101]. Interestingly, the mutations fixed in the latter regime more closely matched mutations observed in clinical pathogens [101]. This provides another hint that in-host adaptation grants a context where mutation bias can be highly influential, as well as demonstrating that clinical evolutionary dynamics can be reflected in laboratory experiments.

It is common practice to initiate evolution experiments with clonal isolates [72, 102], which clears the runway for a hard selective sweep (Fig. 3b). The opportunity for a hard sweep is compounded if the initial microbial population is poorly adapted for the selective conditions, meaning there is ample opportunity for single-step mutations granting a large fitness increase at the onset of the experiment [102, 103]. Longer experimental designs provide opportunities for multiple sweeps as well as multiple rounds of bottlenecks [82, 104]; however, population bottlenecks can be avoided by performing longer experiments in a chemostat, which maintains population diversity [63]. In contrast, shorter experimental designs that isolate the first adaptive mutant to appear, such as when screening for a one-step mutation granting antibiotic resistance [105], can artificially enforce a ‘fixation event’ and further increase the prominence of mutation bias. Therefore, while heavily dependent on the experimental design, many laboratory experiment outcomes, i.e. those initiated with clonal populations evolved over short periods in a novel selective niche, are equipped to be strongly influenced by mutation bias.

## A HOTSPOT IN A HAYSTACK – UTILIZING GENOMIC SIGNATURES TO FIND POTENTIAL BIAS

Determining whether bias is playing a role in the emergence of adaptive mutations requires detailed information about the mechanisms (Fig. 1) and environmental conditions (Fig. 2) driving the biases, as well as the evolutionary context to determine whether they are likely to be impactful (Fig. 3). But once armed with the mechanistic information we are primed to predict the genomic locations exhibiting bias. It has been observed that regions of biased mutation are not conserved even between close homologues [86, 106], and so identifying the mechanisms driving bias offer the best window into forming predictions of biased mutation. Model laboratory bacteria possess well-annotated genomes that contain accurate nucleotide sequence, gene locations and orientations, and a full complement of confirmed or predicted gene functions. As such they are excellent candidates to form mutational forecasts by searching genomic sequence for markers of bias.

Consider, for example, that we wished to predict whether a mutational hotspot at a specific nucleotide position existed within a locus of interest (Fig. 4). For diagnosing intra-locus mutation bias we can initially look extremely locally, narrowing our search window to just the nucleotides immediately flanking each focal base of interest, relying on the knowledge that the nucleotide triplet has a significant effect on base substitution rate [5, 29, 44]. Following this, if the base of interest is a cytosine in an enterobacterial genome, we can expand our window of interest slightly wider, scanning for a CCWGG (W is A or T) sequence flanking our focal nucleotide. This screens for a dcm methylation target site to see if the cytosine is more prone to deamination [107]. We next widen the window to a local neighbourhood of bases, inclusive of perhaps 50 nucleotides either side of our focal nucleotide. Here we can search for quasi-palindromic sequences that can facilitate intra-strand bonding and enable template-switching, which causes small insertions and deletions or single-base substitutions [108–111]. And we can search for homopolymeric tracts that either induce slippage events causing deletions [12, 29, 112] or a base substitution if the base of interest is a thymine preceded by a tract of guanine [17]. For these latter elements, a pipeline dedicated to recognizing hotspots using simple sequence repeats has already been developed [113].

If we wished to predict mutagenicity across an entire locus, we can widen our view to consider the genomic context of the target gene. Gene strandedness (encoding on the positive or negative strand) informs which strand of the gene is the coding and which the non-coding strand during transcription. It also informs which is the lagging strand and which is the leading strand during replication. Both of these elements have a bearing on strand-related biases across the locus. We may also consider the gene of interest's proximity to recombination-causing flanking direct repeats, which can cause the locus to be rearranged or deleted [114–117], and their proximity to nucleoid-associated protein target sites, which determine a locus' vulnerability to acquiring polymorphisms depending on the phase of growth [20, 58–60]. Finally, if we wish to establish broad genomic trends of bias, then we may look at the differential 'wave-like' patterns of mutation rates, where rates are often lower near the origin and terminus of replication and higher at intermediate distances [5, 24–26]. And we may evaluate the functionality or presence of mismatch repair genes [5, 118, 119].

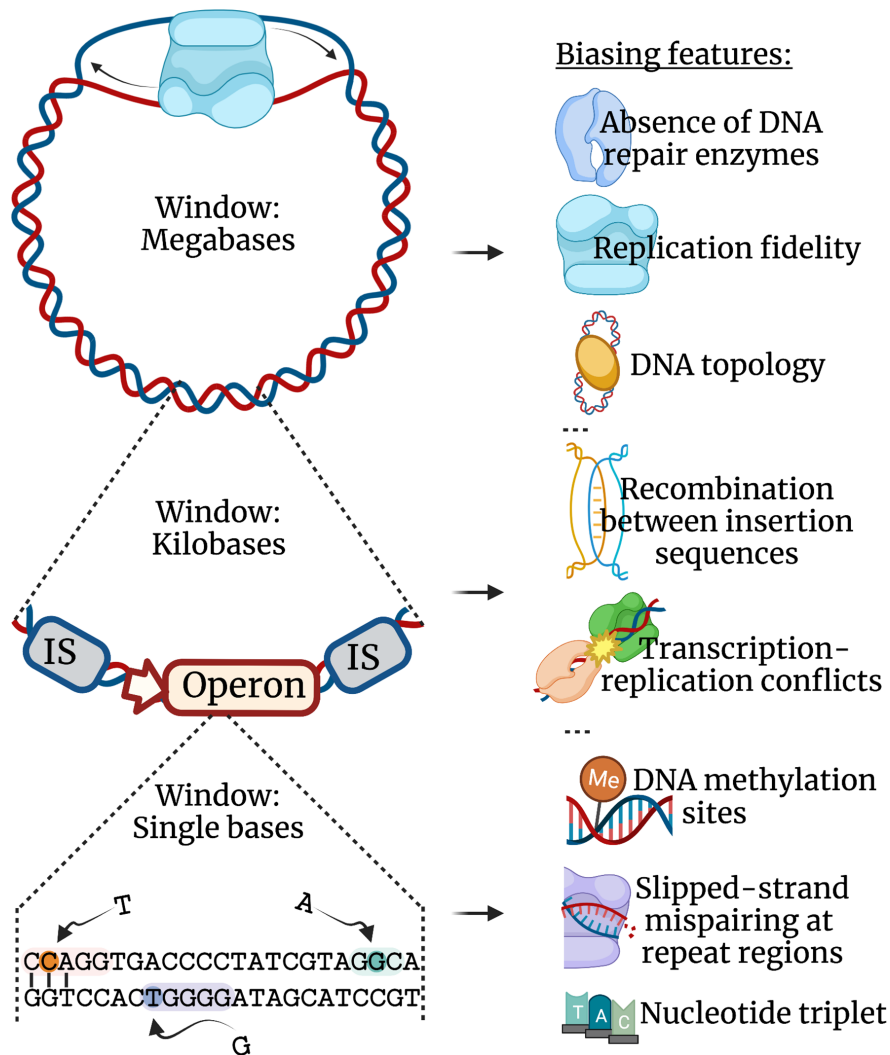
Blokzijl *et al.* released a package in 2018 that uses human mutation data to expose trends in mutational patterns, which can then be used to infer the underlying mechanisms [120]. But there is also the opportunity to forecast mutation from a reference genomic sequence using our knowledge of mutational mechanisms (Fig. 4). These features can be incorporated into pipelines that allow us to identify sequence that is sensitive [121], or resilient [122], to mutation. And by doing so, we can enrich our understanding as to when mutation bias corroborates with selection to guide the adaptation of bacteria.

## FUTURE WORK – EXPERIMENTAL METHODS TO RECOGNIZE AND QUANTIFY INTERACTING FORMS OF BIAS

Many of the mutagenic mechanisms driving mutation bias are now understood, yet we remain limited in our ability to identify likely sites of mutation bias that facilitate accurate forecasts of adaptation. One reason for this is that we must first acquire a generalizable and holistic understanding about how mutagenic mechanisms interact with one another. These mechanisms do not operate in isolation, but in tandem. For example, we know that methylated cytosines can become hotspots for mutagenesis following deamination [108], and there is an increased likelihood of deamination on the non-transcribed DNA strand [123]. Therefore, the two likely compound the mutagenicity of sites on which they both operate. Additionally, in human cancer cells, it is understood that both mismatch repair and transcription-coupled nucleotide excision repair drive strand-specific composition biases of mononucleotide repeats [124].

Features can therefore combine to raise the rates of mutation, but they can also counteract one another. Quasi-palindromic inverse repeats, for example, can cause a localized increase in mutation rate due to their interference with polymerase complexes [30]. Yet inverse repeats also allow single-stranded DNA to form intrastrand bonds that renders the DNA less exposed and susceptible to environmental damage [125]. In another example, a localized mutational hotspot may increase the mutation rate of a specific transversion mutation in an adaptive gene. In a wild-type genomic background this hotspot mutation may dominate the observed mutation pool following adaptation. However, placing this hotspot in a mutator context where transition mutations become more common can shift the observed spectrum to favour the realization of transitions,

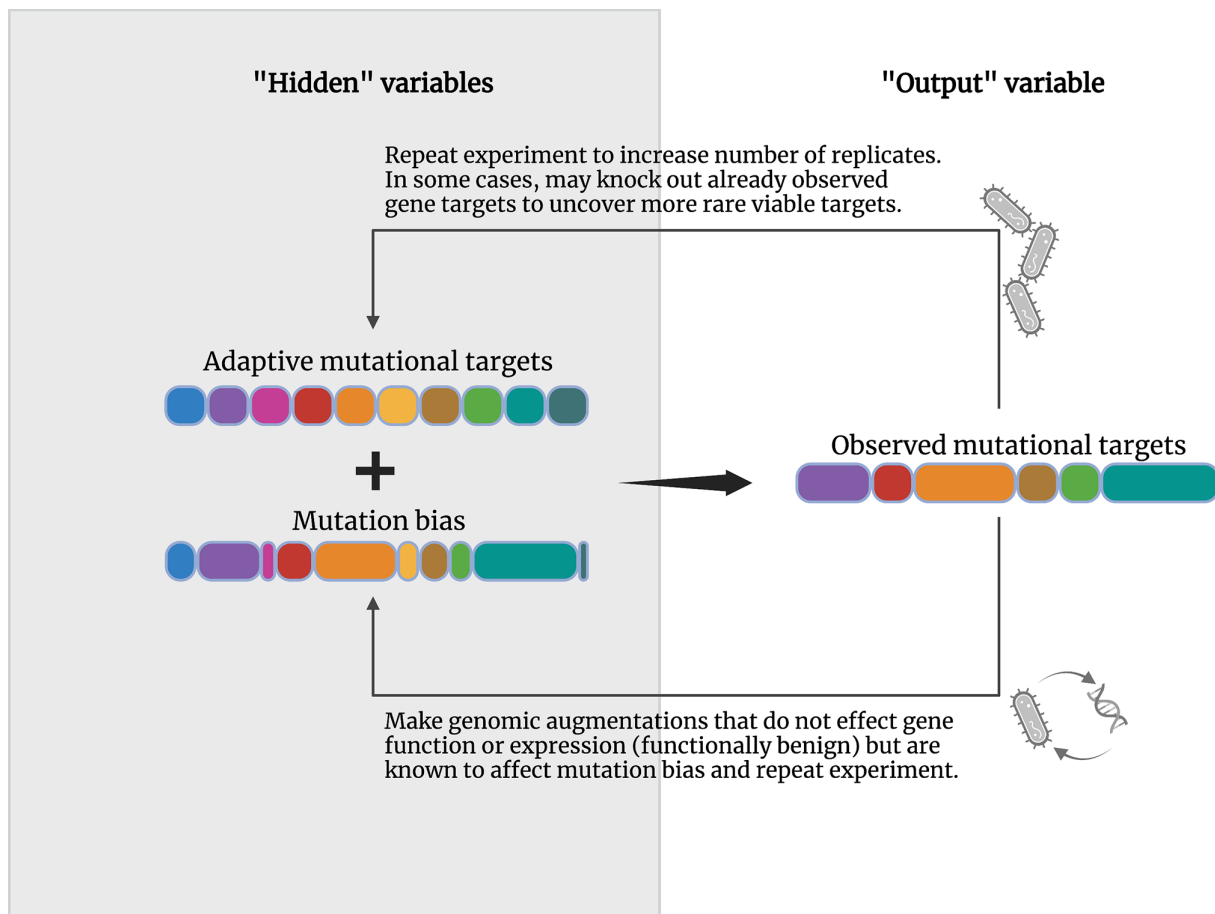




**Fig. 4.** Genomic sequence data can be used to identify areas sensitive to exhibiting biased mutation. When analysing sequence data for mutagenic regions, features can be searched for at different scales, from bias operating across genomic regions to those operating on individual nucleotides. Examples of such features are listed on the right. For example, the 'wave-like' pattern of replication fidelity, which is correlated with genomic region (and likely affected by regions of DNA topology [166]) will fluctuate on the scale of the megabase. Therefore, an assembled genome with an annotated origin of replication can highlight genomic regions less sensitive to mutation. Likewise, the absence of coding regions for mismatch repair enzymes will cause a substitution bias (often a transition bias) throughout the entire genome, which can be identified through alignment searches. At the other end of the scale, if the window of interest is zoomed in to consider local nucleotide neighbourhoods, then small nucleotide motifs such as methylation target sites and homopolymer tracts can increase the rate and bias of mutation at single nucleotide positions. Such narrow investigations can be particularly insightful when examining the open reading frames and promoter regions of known adaptive gene targets. Figure created with BioRender.com.

despite the hotspot's continued presence [19]. We therefore must be able to recognize and quantify the interplay of genomic features to make predictors of mutational bias robust across genomic backgrounds.

So how can we recognize new forms of bias? When we evolve isolates in the laboratory or sequence evolved isolates from clinical samples, we rarely obtain a complete pool of the mutations that can facilitate a phenotype. Instead we observe a sample from the potential pool of mutations that can offer a phenotype (Fig. 5). The result of this is that we can readily observe the effect of selection, i.e. the enrichment of a particular mutation, but we are unable to recognize if there has been any input from mutation bias. This is because without knowing the complete pool of adaptive mutational targets, how can we know if the mutations we observe represent the full pool, or are just the mutations that are biased to occur, or are just a sub-set observed simply through chance?



**Fig. 5.** How to investigate mutation bias in laboratory experimental systems. Following a directed evolution experiment, sequencing techniques provide a researcher with a collection of observed mutational targets (i.e. the 'output' variable). However, the probability of observing a given mutation is defined by the number of possible mutations that can grant a phenotype under selection, and the mutational biases that operate on the positions within this pool (i.e., the 'hidden' variables). Experimental techniques that can elucidate the input of mutation bias and the full list of mutational targets are listed beside the lower and upper arrows. Figure created with BioRender.com.

Elucidating biases at play from the 'bottom-up', i.e. without any *a priori* knowledge of operating biases, will be crucial for uncovering and quantifying the potency of novel bias-driving interactions. In evolution experiments, biases in mutation during adaptation often manifest as instances of parallel or repeatable evolution [86, 87, 106]. Parallel evolution describes occasions where independently evolving populations acquire the same mutation (this makes the evolutionary outcome repeatable, hence why the terms are sometimes used interchangeably). In an experimental context, this involves repeating an identical experiment using clonal samples and repeatedly observing the same mutation or subset of mutations following adaptation [86, 87, 106]. Observations of parallel evolution are not always owed to mutation bias [126]; selection can act on a very low number of positions [127] and therefore a limited mutational target size makes some degree of parallelism inevitable. However, sometimes parallel evolution can be driven by a mutational hotspot [128], where a given mutation occurs at a higher rate than elsewhere. This means that while a pool of mutations are possible, we repeatedly observe mutations fixed at the hotspot site due to a strong mutation bias [86]. Such instances of parallelism are fine targets for using experimental methods to investigate agents of mutation bias (Fig. 5).

## **FUTURE WORK – DOES NATURAL SELECTION ENFORCE, REMOVE OR MOSTLY IGNORE AGENTS OF BIAS?**

An alternative way to anticipate regions of the genome that will be more mutable than others is to consider the role played by natural selection. As most new mutations are deleterious [129], high mutation rates across the genome will be unfavourable to most organisms. The necessary suppression or control of mutation rates explains the ubiquitous evolution of DNA repair enzymes [46];

or the evolution of alternative DNA repair pathways for those organisms that did not evolve mismatch repair [130, 131]. However, there may be more nuance to the role of selection than enforcing non-discriminatory DNA repair enzymes. In a recent study looking at the flowering plant *Arabidopsis thaliana*, Monroe *et al.* reported that mutational biases in the genome reflect mutations fixed in natural evolving populations [122]. They observed that gene bodies and essential genes, which are better protected by epigenomic modifiers such as methylation and histone modifications, exhibit lower mutation rates relative to non-essential loci and intergenic regions [122]. Critics of this work have identified that many of the called mutations in this dataset are products of sequencing errors, which are enriched to occur in regions of perceived higher mutability [132, 133]. However, if the findings remain genuine despite these errors – as the authors argue [134] – this work suggests that selection can suppress mutation rates unequally throughout the genome.

If genuine, these observations can be explained by a theory that is likewise applicable to microbes, proposed by Lynch [135]. Lynch's drift-barrier hypothesis posits that purifying selection (which removes deleterious mutations from a population) works to lower mutation rates, but this is counteracted by the random percolations associated with genetic drift [135, 136]. While this hypothesis does not allow for mutation rates to be specifically reduced for individual genes, if the length of the sequence affected is large enough (i.e. if we consider all essential genes and gene bodies together), selection may be able to drive mutation down over certain genomic regions [122]. Essential genes and others subject to strong purifying selection (meaning that a mutation is almost always strongly deleterious) will therefore be more likely to evolve lower mutation rates. Therefore, we might predict that loci outside of these regions will exhibit relatively higher mutation rates.

A more controversial argument is that selection actively enforces higher mutation rates when these will be beneficial. Above, we discussed growth-associated and stress-associated mutagenesis that result in higher mutation rates. There is an argument that the higher mutation rates might be an advantageous part of the stress response, as they act to elevate rates when the cell is stressed so that mutation can be utilized for evolutionary escape. MacLean *et al.* [57] have argued that stress-induced mutagenesis is likely owed to a combination of factors [57], including genetic drift and positive selection for the reduction of damage from stressors. This positive selection indirectly also selects for increased mutation rate as the same proteins determine both traits [57]. The capacity for mutagenic mechanisms to hitchhike alongside positively selected traits is known as second-order selection [137]. This phenomenon also explains how mutator genotypes lacking mismatch repair can be indirectly selected for under conditions where mutation is necessary for cell survival, as the mutator genotypes are more likely to produce an adaptive mutation [138]. A recent study by Sane *et al.* has also suggested that the change in mutational bias in mutator genotypes provides access mutational types that appear less often in wild-type backgrounds [11]. They argue that these under-sampled mutation types are more likely to be adaptive [11], therefore a change in mutation bias can be indirectly supported by a second-order selection effect. However, direct selection for higher mutation rates or changes in mutation bias have yet to be conclusively demonstrated (although they have recently been implicated in yeast [69]).

Aside from stress-associated mutation and mutators, which affect mutation rates throughout the genome, others have argued that DNA sequence is sometimes selectively primed to mutate at specific positions. Mutable DNA sequences capable of targeted mutation include homopolymer or tandem repeats of nucleotides, which readily undergo reversible insertions and deletions at a high rate [113]. These have been described to readily occur in genes termed 'contingency loci' by Moxon *et al.* [139], who argued that higher rates of mutation are found in genes capable of introducing phenotypic variation in changing environments, such as those encountered within a host [140]. When placed in a promoter sequence or open reading frame, a single insertion or deletion event can remove the in-frame expression of the contingency locus. This mutation may be useful initially but later cease to be useful, as could be the case in a fluctuating environment. But the elevated indel rate at the tract allows the sequence to later revert the mutation and recover its wild-type genotype and phenotype. More broadly, the process of bacterial populations rapidly generating reversible phenotypic diversity is known as phase variation. Pathogens have been noted to employ phase variation to cope with bottlenecks and the loss of genetic diversity when infecting hosts, and phase-variable loci are typically found in cell-surface genes such as adhesins and pili [141, 142]. Phase variation does not require mutagenesis from contingency loci, as changes in gene expression can be achieved instead through epigenomic modifications [143]. However, several studies have demonstrated that localized mutagenic positions can be favoured under selection [16, 144–146]. These include a recent study that detailed how *B. fragilis* utilizes reversible DNA inversions to control the expression of capsular polysaccharides as a means to cope with periods of inflammation [147]. However, whether these positions have been maintained by selection in microbial pathogens is less clear [140].

Finally, it is worth considering the mutagenic mechanisms that are 'less visible' to selection. Many drivers of mutation bias are influenced by the local nucleotide context (Fig. 1). These include homopolymer tracts that cause slipped-strand mispairing [12], methylation recognition sites that cause cytosine deamination [108] inverse repeats that result in hairpin formation and template switching [30], and nucleotide triplets that affect mismatch repair efficiency [29]. As these features rely on short tracts of nucleotides, changing a handful or sometimes a single nucleotide can affect their mutation rates drastically [86]. Further, these base-pair substitutions can be synonymous [86]. While synonymous mutations are now appreciated to be capable of causing changes to fitness [148–151], they are typically under weaker selection than non-synonymous changes [152]. As such, nucleotide neighbourhoods driving divergences in localized mutation biases may not require purifying or positive selection but may appear and degrade in a population simply through synonymous changes via the process of genetic drift.

Whether selection suppresses, elevates, or is mostly ineffectual in governing mutational biases, many questions remain unanswered. For example, there is some empirical evidence that mutation hotspots might be more common in certain types of genes. Multiple experimental evolution studies have found mutation hotspots within histidine kinases at the nexus of regulatory pathways [18, 86], or in other downstream effector genes [16, 153]. Such genes often illicit significant changes to phenotypes. But do we see mutational hotspots in these sequences because of their adaptive capacity, or simply because such hotspots are easier to find due to the phenotypic changes they provoke? It is likely that the identification of these hotspots is owed to the latter explanation. If so, then hotspots are not found in these genes because of positive selection and yet they are found relatively often. This suggests that hotspots may be fairly common, and thus must be able to evade purifying selection at least for a time. Therefore, localized differences in mutational biases may be found throughout the genome and be more widespread than current experimental evolution data suggests.

## CONCLUSION

There has been a growing interest in whether we can improve our ability to predict evolution. A deeper understanding of mutation biases and their role in defining adaptation moves us closer to this goal. In this review, we have discussed the causes and consequences of mutation biases on bacterial adaptation in the lab and in a clinical setting. We have highlighted the immediate challenge that lies ahead, which is bringing together our existing knowledge of how mechanisms of mutation bias work in isolation and use empirical investigation to unravel how they interplay. Efforts have been made to collate mutational data from experimental evolution studies [154]. Such pursuits have the potential to reveal trends that allow for the identification of mutational hotspots and other signals of mutation bias occurring in adaptive genes of interest.

We must also begin to consider the genomic context of mutation bias – for example, we cannot assume an identical mutagenic mechanism is driving mutational bias just because it is facilitating a common outcome [155]. Therefore, we must also gain evidence to define mechanisms of bias across different genomic contexts and within different genomic backgrounds. For example, mutation accumulation experiments in the Lynch laboratory identified that mutation rates are lower toward the origin and terminus of replication in *Pseudomonas fluorescens* [5]. But a mechanism explaining this effect was only implicated when they later performed a similar experiment that compared mutation rates between the core and secondary chromosomes in *Vibrio* species [25]. This demonstrates the power of using alternative bacterial species to both test the robustness and elucidate the mechanism of biasing features in bacterial genomes.

Together, this knowledge may provide the inputs necessary for the ultimate challenge of creating pipelines that will be able to determine mutability at a given site based on information from surrounding genetic features and genomic details. To achieve these ambitious goals will require a collaborative effort between empirical and bioinformatic research groups. It will also require high-throughput approaches to mutation detection, which will increase the likelihood of effectively elucidating the complete pool of mutational adaptive targets – something we are yet to achieve at the necessary scale [156]. Open sharing of data (e.g. [154]) will also be vital for progress to allow bioinformatic approaches to be implemented in a way that will be informative generally across bacteria. However, given the rapid progress that this field has made in recent years, and the growing interests of research groups in achieving predictive power of evolutionary outcomes, we are excited to follow the advances and innovative approaches to the challenges discussed.

### Funding information

This project was supported by a BBSRC NI grant (BB/T012994/1) awarded to T.B.T and supporting J.S.H, and a Royal Society Dorothy Hodgkin Research Fellowship (DH150169) awarded to and supporting T.B.T.

### Conflicts of interest

The authors declare no conflicts of interest.

### References

- Rodríguez-Beltrán J, DelaFuente J, León-Sampedro R, MacLean RC, San Millán Á. Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. *Nat Rev Microbiol* 2021;19:347–359.
- Malik JA, Ahmed S, Mir A, Shinde M, Bender O, et al. The SARS-CoV-2 mutations versus vaccine effectiveness: new opportunities to new challenges. *J Infect Public Health* 2022;15:228–240.
- Davies J, Davies D. Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* 2010;74:417–433.
- Payne JL, Menardo F, Trauner A, Borrell S, Gygli SM, et al. Transition bias influences the evolution of antibiotic resistance in *Mycobacterium tuberculosis*. *PLoS Biol* 2019;17:e3000265.
- Long H, Sung W, Miller SF, Ackerman MS, Doak TG, et al. Mutation rate, spectrum, topology, and context-dependency in the DNA mismatch repair-deficient *Pseudomonas fluorescens* ATCC948. *Genome Biol Evol* 2014;7:262–271.
- Stoltzfus A. *Mutation, Randomness, and Evolution*. Oxford University Press, 2021.
- Hershberg R, Petrov DA. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet* 2010;6:e1001115.
- Rice AM, Morales AC, Ho AT, Mordstein C, Mühlhausen S, et al. Evidence for strong mutation bias toward, and selection against, U content in SARS-CoV-2: implications for vaccine design. *Mol Biol Evol* 2021;38:67–83.
- Horton JS, Ali SUP, Taylor TB. Transient mutation bias increases the predictability of evolution on an empirical genotype–phenotype landscape. *Phil Trans R Soc B* 2023;378:20220043.

10. Cano AV, Rozhoňová H, Stoltzfus A, McCandlish DM, Payne JL. Mutation bias shapes the spectrum of adaptive substitutions. *Proc Natl Acad Sci U S A* 2022;119:1–11.
11. Sane M, Diwan GD, Bhat BA, Wahl LM, Agashe D. Shifts in mutation spectra enhance access to beneficial mutations. *Proc Natl Acad Sci U S A* 2023;120:e2207355120.
12. Dettman JR, Sztepanacz JL, Kassen R. The properties of spontaneous mutations in the opportunistic pathogen *Pseudomonas aeruginosa*. *BMC Genomics* 2016;17:1–14.
13. Katju V, Bergthorsson U. Old trade, new tricks: insights into the spontaneous mutation process from the partnering of classical mutation accumulation experiments with high-throughput genomic approaches. *Genome Biol Evol* 2019;11:136–165.
14. Stoltzfus A, Norris RW. On the causes of evolutionary transition:transversion bias. *Mol Biol Evol* 2016;33:595–602.
15. Lyons DM, Luring AS. Evidence for the selective basis of transition-to-transversion substitution bias in two RNA viruses. *Mol Biol Evol* 2017;34:3205–3215.
16. Hefetz I, Israeli O, Bilinsky G, Plaschkes I, Hazkani-Covo E, et al. A reversible mutation in a genomic hotspot saves bacterial swarms from extinction. *iScience* 2023;26:106043.
17. Cherry JL, Katz L. T residues preceded by runs of G are hotspots of T→G mutation in bacteria. *Genome Biol Evol* 2023;15:evad087.
18. Kapel N, Caballero JD, MacLean RC. Localized pnrB hypermutation drives the evolution of colistin heteroresistance. *Cell Rep* 2022;39:110929.
19. Shepherd MJ, Horton JS, Taylor TB, Hendrickson H. A near-deterministic mutational hotspot in *Pseudomonas fluorescens* is constructed by multiple interacting genomic features. *Mol Biol Evol* 2022;39:msac132.
20. Warnecke T, Supek F, Lehner B. Nucleoid-associated proteins affect mutation dynamics in *E. coli* in a growth phase-specific manner. *PLoS Comput Biol* 2012;8:e1002846.
21. Maki H. Origins of spontaneous mutations: specificity and directionality of base-substitution, frameshift, and sequence-substitution mutageneses. *Annu Rev Genet* 2002;36:279–303.
22. Selveshwari S, Lele K, Dey S. Genomic signatures of UV resistance evolution in *Escherichia coli* depend on the growth phase during exposure. *J Evol Biol* 2021;34:953–967.
23. Bikard D, Loot C, Baharoglu Z, Mazel D. Folded DNA in action: hairpin formation and biological functions in prokaryotes. *Microbiol Mol Biol Rev* 2010;74:570–588.
24. Hudson RE, Bergthorsson U, Roth JR, Ochman H. Effect of chromosome location on bacterial mutation rates. *Mol Biol Evol* 2002;19:85–92.
25. Dillon MM, Sung W, Lynch M. Periodic variation of mutation rates in bacterial genomes. *mBio* 2018;9.
26. Zhang X, Zhang X, Zhang X, Liao Y, Song L, et al. Spatial vulnerabilities of the *Escherichia coli* genome. *Genetics* 2018;210:547–558.
27. Maslowska KH, Makiela-Dzbenka K, Mo J-Y, Fijalkowska IJ, Schaaper RM. High-accuracy lagging-strand DNA replication mediated by DNA polymerase dissociation. *Proc Natl Acad Sci U S A* 2018;115:4212–4217.
28. Bhagwat AS, Hao W, Townes JP, Lee H, Tang H, et al. Strand-biased cytosine deamination at the replication fork causes cytosine to thymine mutations in *Escherichia coli*. *Proc Natl Acad Sci U S A* 2016;113:2176–2181.
29. Foster PL, Niccum BA, Popodi E, Townes JP, Lee H, et al. Determinants of base-pair substitution patterns revealed by whole-genome sequencing of DNA mismatch repair defective *Escherichia coli*. *Genetics* 2018;209:1029–1042.
30. Dutra BE, Lovett ST. Cis and trans-acting effects on a mutational hotspot involving a replication template switch. *J Mol Biol* 2006;356:300–311.
31. Lai PJ, Lim CT, Le HP, Katayama T, Leach DRF, et al. Long inverted repeat transiently stalls DNA replication by forming hairpin structures on both leading and lagging strands. *Genes Cells* 2016;21:136–145.
32. Trinh TQ, Sinden RR. Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*. *Nature* 1991;352:544–547.
33. Schroeder JW, Hirst WG, Szewczyk GA, Simmons LA. The effect of local sequence context on mutational bias of genes encoded on the leading and lagging strands. *Curr Biol* 2016;26:692–697.
34. Beletskii A, Bhagwat AS. Transcription-induced mutations: increase in C to T mutations in the nontranscribed strand during transcription in *Escherichia coli*. *Proc Natl Acad Sci U S A* 1996;93:13919–13924.
35. Davis BD. Transcriptional bias: a non-Lamarckian mechanism for substrate-induced mutations. *Proc Natl Acad Sci U S A* 1989;86:5005–5009.
36. Francino MP, Ochman H. Deamination as the basis of strand-asymmetric evolution in transcribed *Escherichia coli* sequences. *Mol Biol Evol* 2001;18:1147–1150.
37. Mellon I, Hanawalt PC. Induction of the *Escherichia coli* lactose operon selectively increases repair of its transcribed DNA strand. *Nature* 1989;342:95–98.
38. Kraithong T, Hartley S, Jeruzalmi D, Pakotiprapha D. A peek inside the machines of bacterial nucleotide excision repair. *Int J Mol Sci* 2021;22:952.
39. Schroeder JW, Sankar TS, Wang JD, Simmons LA. The roles of replication-transcription conflict in mutagenesis and evolution of genome organization. *PLoS Genet* 2020;16:e1008987.
40. Brüning J-G, Mariani KJ. Bypass of complex co-directional replication-transcription collisions by replisome skipping. *Nucleic Acids Res* 2021;49:9870–9885.
41. Paul S, Million-Weaver S, Chattopadhyay S, Sokurenko E, Merrikh H. Accelerated gene evolution through replication-transcription conflicts. *Nature* 2013;495:512–515.
42. Foster PL, Niccum BA, Lee H. DNA replication-transcription conflicts do not significantly contribute to spontaneous mutations due to replication errors in *Escherichia coli*. *mBio* 2021;12:e0250321.
43. Schroeder JW, Hirst WG, Szewczyk GA, Simmons LA. The effect of local sequence context on mutational bias of genes encoded on the leading and lagging strands. *Curr Biol* 2016;26:692–697.
44. Merrikh CN, Weiss E, Merrikh H. The accelerated evolution of lagging strand genes is independent of sequence context. *Genome Biol Evol* 2016;8:3696–3702.
45. Sankar TS, Wastuwidyaningtyas BD, Dong Y, Lewis SA, Wang JD. The nature of mutations induced by replication-transcription collisions. *Nature* 2016;535:178–181.
46. Eisen JA, Hanawalt PC. A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat Res* 1999;435:171–213.
47. Boyce KJ. Mutators enhance adaptive micro-evolution in pathogenic microbes. *Microorganisms* 2022;10:442.
48. LeClerc JE, Li B, Payne WL, Cebula TA. High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science* 1996;274:1208–1211.
49. Denamur E, Bonacorsi S, Giraud A, Duriez P, Hilali F, et al. High frequency of mutator strains among human uropathogenic *Escherichia coli* isolates. *J Bacteriol* 2002;184:605–609.
50. Oliver A, Cantón R, Campo P, Baquero F, Blázquez J. High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science* 2000;288:1251–1254.
51. Denamur E, Matic I. Evolution of mutation rates in bacteria. *Mol Microbiol* 2006;60:820–827.
52. Couce A, Rodríguez-Rojas A, Blázquez J. Bypass of genetic constraints during mutator evolution to antibiotic resistance. *Proc R Soc B* 2015;282:20142698.
53. Foster PL, Lee H, Popodi E, Townes JP, Tang H. Determinants of spontaneous mutation in the bacterium *Escherichia coli* as

- revealed by whole-genome sequencing. *Proc Natl Acad Sci U S A* 2015;112:E5990–9.
54. Sung W, Ackerman MS, Gout J-F, Miller SF, Williams E, *et al*. Asymmetric context-dependent mutation patterns revealed through mutation-accumulation experiments. *Mol Biol Evol* 2015;32:1672–1683.
  55. Ferenci T. Irregularities in genetic variation and mutation rates with environmental stresses. *Environ Microbiol* 2019;21:3979–3988.
  56. Abundiz-Yañez K, Leyva-Sánchez HC, Robleto EA, Pedraza-Reyes M. Stress-associated and growth-dependent mutagenesis are divergently regulated by c-di-AMP levels in *Bacillus subtilis* *Int J Mol Sci* 2022;24:455.
  57. MacLean RC, Torres-Barceló C, Moxon R. Evaluating evolutionary models of stress-induced mutagenesis in bacteria. *Nat Rev Genet* 2013;14:221–227.
  58. Foster PL, Hanson AJ, Lee H, Popodi EM, Tang H. On the mutational topology of the bacterial genome. *G3 Genes|Genomes|Genetics* 2013;3:399–407.
  59. Kivisaar M. Mutation and recombination rates vary across bacterial chromosome. *Microorganisms* 2020;8:25.
  60. Niccum BA, Lee H, MohammedIsmail W, Tang H, Foster PL. The symmetrical wave pattern of base-pair substitution rates across the *Escherichia coli* chromosome has multiple causes. *mBio* 2019;10:e01226-19.
  61. Pereira FC, Berry D. Microbial nutrient niches in the gut. *Environ Microbiol* 2017;19:1366–1378.
  62. La Rosa R, Johansen HK, Molin S. Adapting to the airways: metabolic requirements of *Pseudomonas aeruginosa* during the infection of cystic fibrosis patients. *Metabolites* 2019;9:234.
  63. Gresham D, Dunham MJ. The enduring utility of continuous culturing in experimental evolution. *Genomics* 2014;104:399–405.
  64. Loewen PC, Hengge-Aronis R. The role of the sigma factor sigma S (KatF) in bacterial global regulation. *Annu Rev Microbiol* 1994;48:53–80.
  65. Maharjan R, Ferenci T. Stress-induced mutation rates show a sigmoidal and saturable increase due to the RpoS sigma factor in *Escherichia coli*. *Genetics* 2014;198:1231–1235.
  66. Battesti A, Majdalani N, Gottesman S. The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* 2011;65:189–213.
  67. Maharjan R, Ferenci T. Mutational signatures indicative of environmental stress in bacteria. *Mol Biol Evol* 2015;32:380–391.
  68. Krašovec R, Richards H, Gifford DR, Belavkin RV, Channon A, *et al*. Opposing effects of final population density and stress on *Escherichia coli* mutation rate. *ISME J* 2018;12:2981–2987.
  69. Liu H, Zhang J. The rate and molecular spectrum of mutation are selectively maintained in yeast. *Nat Commun* 2021;12:4044.
  70. Krašovec R, Richards H, Gifford DR, Hatcher C, Faulkner KJ, *et al*. Spontaneous mutation rate is a plastic trait associated with population density across domains of life. *PLoS Biol* 2017;15:e2002731.
  71. Friedlander A, Nir S, Rechtes M, Shemesh M. Preventing biofilm formation by dairy-associated bacteria using peptide-coated surfaces. *Front Microbiol* 2019;10:1405.
  72. Paul B, Rainey & Michael Travisano. Adaptive radiation in a heterogeneous environment. *Nat Lett* 1998;394:69–72.
  73. Podlesek Z, Žgur Bertok D. The DNA damage inducible SOS response is a key player in the generation of bacterial persister cells and population wide tolerance. *Front Microbiol* 2020;11:1785.
  74. Kivisaar M. Mechanisms of stationary-phase mutagenesis in bacteria: mutational processes in pseudomonads. *FEMS Microbiol Lett* 2010;312:1–14.
  75. Joseph AM, Badrinarayanan A. Visualizing mutagenic repair: novel insights into bacterial translesion synthesis. *FEMS Microbiol Rev* 2020;44:572–582.
  76. Revitt-Mills SA, Robinson A. Antibiotic-induced mutagenesis: under the microscope. *Front Microbiol* 2020;11:585175.
  77. Long H, Miller SF, Strauss C, Zhao C, Cheng L, *et al*. Antibiotic treatment enhances the genome-wide mutation rate of target cells. *Proc Natl Acad Sci U S A* 2016;113:E2498–505.
  78. Messer PW, Petrov DA. Population genomics of rapid adaptation by soft selective sweeps. *Trends Ecol Evol* 2013;28:659–669.
  79. Wilson BA, Petrov DA, Messer PW. Soft selective sweeps in complex demographic scenarios. *Genetics* 2014;198:669–684.
  80. Avalos A, Pan H, Li C, Acevedo-Gonzalez JP, Rendon G, *et al*. A soft selective sweep during rapid evolution of gentle behaviour in an Africanized honeybee. *Nat Commun* 2017;8:1550.
  81. Garud NR, Messer PW, Buzbas EO, Petrov DA. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLoS Genet* 2015;11:e1005004.
  82. Avrani S, Bolotin E, Katz S, Hershberg R. Rapid genetic adaptation during the first four months of survival under resource exhaustion. *Mol Biol Evol* 2017;34:1758–1769.
  83. Miller CR, Joyce P, Wichman HA. Mutational effects and population dynamics during viral adaptation challenge current models. *Genetics* 2011;187:185–202.
  84. Hermisson J, Pennings PS. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 2005;169:2335–2352.
  85. McCandlish DM, Stoltzfus A. Modeling evolution using the probability of fixation: history and implications. *Q Rev Biol* 2014;89:225–252.
  86. Horton JS, Flanagan LM, Jackson RW, Priest NK, Taylor TB. A mutational hotspot that determines highly repeatable evolution can be built and broken by silent genetic changes. *Nat Commun* 2021;12:6092.
  87. Sackman AM, McGee LW, Morrison AJ, Pierce J, Anisman J, *et al*. Mutation-driven parallel evolution during viral adaptation. *Mol Biol Evol* 2017;34:3243–3253.
  88. Jensen JD. On the unfounded enthusiasm for soft selective sweeps. *Nat Commun* 2014;5:5281.
  89. Rousselle M, Simion P, Tilak M-K, Figuet E, Nabholz B, *et al*. Is adaptation limited by mutation? A timescale-dependent effect of genetic diversity on the adaptive substitution rate in animals. *PLoS Genet* 2020;16:e1008668.
  90. Chiner-oms Á, Berney M, Boinett C, González-candelas F, Young DB, *et al*. Genome-wide mutational biases fuel tuberculosis complex. *Nat Commun* 2019;10.
  91. Goldhill DH, Langat P, Xie H, Galiano M, Miah S, *et al*. Determining the mutation bias of favipiravir in influenza virus. *J Virol* 2019;93:e01217–18.
  92. Bergstrom CT, McElhany P, Real LA. Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. *Proc Natl Acad Sci U S A* 1999;96:5095–5100.
  93. Hannan TJ, Totsika M, Mansfield KJ, Moore KH, Schembri MA, *et al*. Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *FEMS Microbiol Rev* 2012;36:616–648.
  94. Bachta KER, Allen JP, Cheung BH, Chiu C-H, Hauser AR. Systemic infection facilitates transmission of *Pseudomonas aeruginosa* in mice. *Nat Commun* 2020;11:543.
  95. Golubchik T, Batty EM, Miller RR, Farr H, Young BC, *et al*. Within-host evolution of *Staphylococcus aureus* during asymptomatic carriage. *PLoS One* 2013;8:e61319.
  96. Didelot X, Walker AS, Peto TE, Crook DW, Wilson DJ. Within-host evolution of bacterial pathogens. *Nat Rev Microbiol* 2016;14:150–162.
  97. Couce A, Tenailon O. Mutation bias and GC content shape antimutator invasions. *Nat Commun* 2019;10:3114.
  98. Harris KB, Flynn KM, Cooper VS. Polygenic adaptation and clonal interference enable sustained diversity in

- experimental *Pseudomonas aeruginosa* populations. *Mol Biol Evol* 2021;38:5359–5375.
99. Couce A, Rodríguez-Rojas A, Blázquez J. Determinants of genetic diversity of spontaneous drug resistance in bacteria. *Genetics* 2016;203:1369–1380.
  100. Schenk MF, Zwart MP, Hwang S, Ruelens P, Severing E, et al. Population size mediates the contribution of high-rate and large-benefit mutations to parallel evolution. *Nat Ecol Evol* 2022;6:439–447.
  101. Bacigalupe R, Tormo-Mas MÁ, Penadés JR, Fitzgerald JR. A multihost bacterial pathogen overcomes continuous population bottlenecks to adapt to new host species. *Sci Adv* 2019;5.
  102. Garoff L, Pietsch F, Huseby DL, Lilja T, Brandis G, et al. Population bottlenecks strongly influence the evolutionary trajectory to fluoroquinolone resistance in *Escherichia coli*. *Mol Biol Evol* 2020;37:1637–1646.
  103. Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* 2003;4:457–469.
  104. MacLean RC, Perron GG, Gardner A. Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in *Pseudomonas aeruginosa*. *Genetics* 2010;186:1345–1354.
  105. Lenski RE. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *ISME J* 2017;11:2181–2194.
  106. Vogwill T, Kojadinovic M, Furió V, MacLean RC. Testing the role of genetic background in parallel evolution using the comparative experimental evolution of antibiotic resistance. *Mol Biol Evol* 2014;31:3314–3323.
  107. Pentz JT, Lind PA. Forecasting of phenotypic and genetic outcomes of experimental evolution in *Pseudomonas protegens*. *PLoS Genet* 2021;17:e1009722.
  108. Cherry JL. Methylation-induced hypermutation in natural populations of bacteria. *J Bacteriol* 2018;200:e00371–18.
  109. Löytynoja A, Goldman N. Short template switch events explain mutation clusters in the human genome. *Genome Res* 2017;27:1039–1049.
  110. Lavi B, Levy Karin E, Pupko T, Hazkani-Covo E. The prevalence and evolutionary conservation of inverted repeats in proteobacteria. *Genome Biol Evol* 2018;10:918–927.
  111. Yoshiyama K, Maki H. Spontaneous hotspot mutations resistant to mismatch correction in *Escherichia coli*: transcription-dependent mutagenesis involving template-switching mechanisms. *J Mol Biol* 2003;327:7–18.
  112. Voineagu I, Narayanan V, Lobachev KS, Mirkin SM. Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. *Proc Natl Acad Sci U S A* 2008;105:9936–9941.
  113. Moxon R, Bayliss C, Hood D. Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu Rev Genet* 2006;40:307–333.
  114. Darling AE, Miklós I, Ragan MA. Dynamics of genome rearrangement in bacterial populations. *PLoS Genet* 2008;4:e1000128.
  115. Tillier ERM, Collins RA. Genome rearrangement by replication-directed translocation. *Nat Genet* 2000;26:195–197.
  116. Mira A, Ochman H, Moran NA. Deletional bias and the evolution of bacterial genomes. *Trends Genet* 2001;17:589–596.
  117. Raeside C, Gaffé J, Deatherage DE, Tenailon O, Briska AM, et al. Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. *mBio* 2014;5:1–13.
  118. Schaaper RM, Dunn RL. Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: the nature of in vivo DNA replication errors. *Proc Natl Acad Sci USA* 1987;84:6220–6224.
  119. Wong A, Rodrigue N, Kassen R, Guttman DS. Genomics of adaptation during experimental evolution of the opportunistic pathogen *Pseudomonas aeruginosa*. *PLoS Genet* 2012;8:e1002928.
  120. Blokzijl F, Janssen R, van Boxel R, Cuppen E. Mutational Patterns: comprehensive genome-wide analysis of mutational processes. *Genome Med* 2018;10:33.
  121. Jack BR, Leonard SP, Mishler DM, Renda BA, Leon D, et al. Predicting the genetic stability of engineered DNA sequences with the EFM calculator. *ACS Synth Biol* 2015;4:939–943.
  122. Monroe JG, Srikant T, Carbonell-Bejerano P, Becker C, Lensink M, et al. Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* 2022;602:101–105.
  123. Lind PA, Andersson DI. Whole-genome mutational biases in bacteria. *Proc Natl Acad Sci USA* 2008;105:17878–17883.
  124. Georgakopoulos-Soares I, Koh G, Momen SE, Jiricny J, Hemberg M, et al. Transcription-coupled repair and mismatch repair contribute towards preserving genome integrity at mononucleotide repeat tracts. *Nat Commun* 2020;11:1–9.
  125. Hoede C, Denamur E, Tenailon O. Selection acts on DNA secondary structures to decrease transcriptional mutagenesis. *PLOS Genet* 2006;2:e176.
  126. Bailey SF, Blanquart F, Bataillon T, Kassen R. What drives parallel evolution?: How population size and mutational variation contribute to repeated evolution. *Bioessays* 2017;39:1–9.
  127. Weinreich DM, Delaney NF, Depristo MA, Hartl DL. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 2006;312:111–114.
  128. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA* 1991;88:7160–7164.
  129. Kishony R, Leibler S. Environmental stresses can alleviate the average deleterious effect of mutations. *J Biol* 2003;2:14.
  130. Castañeda-García A, Prieto AI, Rodríguez-Beltrán J, Alonso N, Cantillon D, et al. A non-canonical mismatch repair pathway in prokaryotes. *Nat Commun* 2017;8:14246.
  131. Castañeda-García A, Martín-Bleuca I, Cebrián-Sastre E, Chiner-Oms A, Torres-Puente M, et al. Specificity and mutagenesis bias of the mycobacterial alternative mismatch repair analyzed by mutation accumulation studies. *Sci Adv* 2020;6:eaa4453.
  132. Liu H, Zhang J, Zhai W. Is the mutation rate lower in genomic regions of stronger selective constraints? *Mol Biol Evol* 2022;39:msac169.
  133. Wang L, Ho AT, Hurst LD, Yang S. Re-evaluating evidence for adaptive mutation rate variation. *Nature* 2023;619:E52–E56.
  134. Monroe JG, Murray KD, Xian W, Srikant T, Carbonell-Bejerano P, et al. Reply to: re-evaluating evidence for adaptive mutation rate variation. *Nature* 2023;619:E57–E60.
  135. Lynch M. Evolution of the mutation rate. *Trends Genet* 2010;26:345–352.
  136. Sung W, Ackerman MS, Miller SF, Doak TG, Lynch M. Drift-barrier hypothesis and mutation-rate evolution. *Proc Natl Acad Sci U S A* 2012;109:18488–18492.
  137. Tenailon O, Taddei F, Radmian M, Matic I. Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Res Microbiol* 2001;152:11–16.
  138. Swings T, Van den Bergh B, Wuyts S, Oeyen E, Voordeckers K, et al. Adaptive tuning of mutation rates allows fast response to lethal stress in *Escherichia coli*. *Elife* 2017;6:e22939.
  139. Moxon ER, Rainey PB, Nowak MA, Lenski RE. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr Biol* 1994;4:24–33.
  140. De Ste Croix M, Holmes J, Wanford JJ, Moxon ER, Oggioni MR, et al. Selective and non-selective bottlenecks as drivers of the evolution of hypermutable bacterial loci. *Mol Microbiol* 2020;113:672–681.
  141. Gor V, Ohniwa RL, Morikawa K. No change, no life? What we know about phase variation in *Staphylococcus aureus*. *Microorganisms* 2021;9:244.

142. Phillips ZN, Tram G, Seib KL, Atack JM. Phase-variable bacterial loci: how bacteria gamble to maximise fitness in changing environments. *Biochem Soc Trans* 2019;47:1131–1141.
143. Tan A, Atack JM, Jennings MP, Seib KL. The capricious nature of bacterial pathogens: phasevarions and vaccine development. *Front Immunol* 2016;7:586.
144. Bayliss CD, Field D, Moxon ER. The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*. *J Clin Invest* 2001;107:657–666.
145. Jerome JP, Bell JA, Plovovich-Jones AE, Barrick JE, Brown CT, et al. Standing genetic variation in contingency loci drives the rapid adaptation of *Campylobacter jejuni* to a novel host. *PLoS One* 2011;6:e16399.
146. Zhou K, Aertsen A, Michiels CW. The role of variable DNA tandem repeats in bacterial adaptation. *FEMS Microbiol Rev* 2014;38:119–141.
147. Carasso S, Zaatry R, Hajjo H, Kadosh-Kariti D, Ben-Assa N, et al. Inflammation and bacteriophages affect DNA inversion states and functionality of the gut microbiota. *bioRxiv*. DOI: 10.1101/2023.03.31.535104
148. Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 2009;324:255–258.
149. Kristofich J, Morgenthaler AB, Kinney WR, Ebmeier CC, Snyder DJ, et al. Synonymous mutations make dramatic contributions to fitness when growth is limited by a weak-link enzyme. *PLoS Genet* 2018;14:e1007615.
150. Lebeuf-Taylor E, McCloskey N, Bailey SF, Hinz A, Kassen R. The distribution of fitness effects among synonymous mutations in a gene under directional selection. *eLife* 2019;8.
151. Frumkin I, Lajoie MJ, Gregg CJ, Hornung G, Church GM, et al. Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proc Natl Acad Sci USA* 2018;115:E4940–E4949.
152. Rahman S, Kosakovsky Pond SL, Webb A, Hey J. Weak selection on synonymous codons substantially inflates *dN/dS* estimates in bacteria. *Proc Natl Acad Sci U S A* 2021;118:e2023575118.
153. Lind PA, Libby E, Herzog J, Rainey PB. Predicting mutational routes to new adaptive phenotypes. *eLife* 2019;8.
154. Phaneuf PV, Gosting D, Palsson BO, Feist AM. ALEdb 1.0: a database of mutations from adaptive laboratory evolution experimentation. *Nucleic Acids Res* 2019;47:D1164–D1171.
155. Rocha EPC, Touchon M, Feil EJ. Similar compositional biases are caused by very different mutational effects. *Genome Res* 2006;16:1537–1547.
156. Sun TA, Lind PA. Distribution of mutation rates challenges evolutionary predictability. *Microbiology* 2023;169:001323.
157. Ganai RA, Johansson E. DNA replication—a matter of fidelity. *Mol Cell* 2016;62:745–755.
158. Li W, Lynch M. Universally high transcript error rates in bacteria. *eLife* 2020;9:1–15.
159. Merrih H, Zhang Y, Grossman AD, Wang JD. Replication–transcription conflicts in bacteria. *Nat Rev Microbiol* 2012;10:449–458.
160. Long H, Miller SF, Williams E, Lynch M, Ruiz-Trillo I. Specificity of the DNA Mismatch Repair System (MMR) and mutagenesis bias in bacteria. *Mol Biol Evol* 2018;35:2414–2421.
161. Pomerantz RT, O'Donnell M. Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science* 2010;327:590–592.
162. Million-Weaver S, Samadpour AN, Moreno-Habel DA, Nugent P, Brittnacher MJ, et al. An underlying mechanism for the increased mutagenesis of lagging-strand genes in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 2015;112:E1096–105.
163. Lang KS, Hall AN, Merrih CN, Ragheb M, Tabakh H, et al. Replication–transcription conflicts generate R-loops that orchestrate bacterial stress survival and pathogenesis. *Cell* 2017;170:787–799.
164. Brazda V, Fojta M, Bowater RP. Structures and stability of simple DNA repeats from bacteria. *Biochem J* 2020;477:325–339.
165. Watanabe S, Ohbayashi R, Kanesaki Y, Saito N, Chibazakura T, et al. Intensive DNA replication and metabolism during the lag phase in cyanobacteria. *PLoS One* 2015;10:e0136800.
166. Maddamsetti R, Grant NA. Divergent evolution of mutation rates and biases in the long-term evolution experiment with *Escherichia coli*. *Genome Biol Evol* 2020;12:1591–1603.

**The Microbiology Society is a membership charity and not-for-profit publisher.**

**Your submissions to our titles support the community – ensuring that we continue to provide events, grants and professional development for microbiologists at all career stages.**

**Find out more and submit your article at [microbiologyresearch.org](https://microbiologyresearch.org)**