Published in final edited form as:

Crit Rev Biochem Mol Biol. 2018 February ; 53(1): 29–48. doi:10.1080/10409238.2017.1394262.

Sources of spontaneous mutagenesis in bacteria

Jeremy W. Schroeder^{a,*}, Ponlkrit Yeesin^a, Lyle A. Simmons^b, and Jue D. Wang^{a,*}

^aDepartment of Bacteriology, University of Wisconsin – Madison, Madison, WI, USA

^bDepartment of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI, USA

Abstract

Mutations in an organism's genome can arise spontaneously, i.e., in the absence of exogenous stress and prior to selection. Mutations are often neutral or deleterious to individual fitness, but can also provide genetic diversity driving evolution. Mutagenesis in bacteria contributes to the already serious and growing problem of antibiotic resistance. However, the negative impacts of spontaneous mutagenesis on human health are not limited to bacterial antibiotic resistance. Spontaneous mutations also underlie tumorigenesis and evolution of drug resistance. To better understand the causes of genetic change and how they may be manipulated in order to curb antibiotic resistance or the development of cancer, we must acquire a mechanistic understanding of the major sources of mutagenesis. Bacterial systems are particularly well-suited to studying mutagenesis because of their fast growth rate and the panoply of available experimental tools but, efforts to understand mutagenic mechanisms can be complicated by the experimental system employed. Here we review our current understanding of mutagenic mechanisms in bacteria and describe the methods used to study mutagenesis in bacterial systems.

Keywords

spontaneous mutagenesis; mutation reporter; evolution; fluctuation test; selection; mutation accumulation

Introduction

Mechanisms of adaptive genetic change such as CRISPR/Cas systems have received much attention lately as mechanisms by which bacteria modify their genome sequence to evolve resistance in response to selective pressures such as phage infection [reviewed in (Barrangou and Marraffini, 2014)]. However, more than 70 years ago Salvador Luria and Max Delbruck recognized in their classical paper that *Escherichia coli* mutants resistant to phage could arise “spontaneously”, i.e., prior to phage exposure (Luria and Delbruck, 1943). By demonstrating the importance of spontaneous mutations in generating genetic diversity prior

*For correspondence: Jeremy Schroeder: jwschroeder3@wisc.edu, phone, (608) 263-1072; Jue Wang: wang@bact.wisc.edu, phone, (608) 263-0307.

Ponlkrit Yeesin: yeesin@wisc.edu, phone, (608) 263-1072

Lyle Simmons: lasimm@umich.edu, phone, (734) 647-2016

Declaration of interest: The authors declare they have no conflict of interest to report.

to selective pressure, Luria and Delbruck ushered in the era of bacterial genetics (Luria and Delbruck, 1943, Brock, 1990). In addition to phage resistance, spontaneous mutagenesis can also result in genetic change important for antibiotic resistance, host evasion, and adaptation to new environments, thus driving bacterial evolution. Although much attention is devoted to the effects of environmental and behavioral factors on mutagenesis in bacteria and the development of cancer in humans, approximately 66% of cancer driver mutations in humans are attributable to random spontaneous mutagenesis that occurs during normal DNA replication (Tomasetti *et al.*, 2017). This underscores the importance of understanding the fundamental mechanisms of spontaneous mutagenesis in biology.

In this review, we describe experimental designs for examining spontaneous mutagenesis and mutagenic mechanisms. We also discuss recent progress that has provided a deeper understanding of the cellular processes generating spontaneous mutations. Spontaneous mutations are distinct from mutations induced by environmental stresses such as UV irradiation and genotoxic agents (Ciccio and Elledge, 2010) that are covered by previous Critical Reviews articles (Foster, 2007, Galhardo *et al.*, 2007, Robleto *et al.*, 2007). For readers interested in details of the chemistry behind the myriad types of DNA damage that can cause mutagenesis, we recommend (Friedberg *et al.*, 2006).

Detection of spontaneous mutations

Mutations are heritable changes to an organism's genotype, which is the term used to describe the entirety of the genetic information in an organism's genome (Table 1). The word "mutation" is often used in reference to local changes within a gene, including base substitutions and small insertions and deletions. Mutations also include intergenic changes, large genome rearrangements, and movement of mobile genetic elements. For a period of time, studies of mutagenesis were primarily focused on point mutations, often referring to base pair substitutions, and sometimes one base pair insertions and deletions (Coulondre and Miller, 1977, Cupples and Miller, 1989, Cupples *et al.*, 1990). Base pair substitutions include transitions and transversions. Transitions refer to substitution of a pyrimidine for another pyrimidine (Table 1). Transversions represent substitution of a pyrimidine for a purine or *vice versa* (Table 1). Functionally, base pair substitutions within the open reading frame of a gene can be synonymous, in which the amino acid sequence of the gene product is unchanged, or they can be nonsynonymous, in which the amino acid encoded at the affected position of the open reading frame is altered. Nonsynonymous substitutions can also result in premature translation termination in the case when a codon for an amino acid is substituted with a stop codon. One base-pair insertions and deletions shift the frame of the gene in which they occur and are almost certain to alter gene function (Table 1). Point mutations within genes are not the only important mutations. In addition to the open reading frame of genes, mutations in cis-regulatory elements are also important because they alter gene expression. In addition to one base-pair insertions and deletions, "small" and large insertions and deletions also occur, with important consequences for gene function. There is no objective standard threshold for which insertions and deletions are considered to be small. For instance, recent work has used insertion and deletion lengths of less than 5, 10, and 28 as a cutoff for considering insertions and deletions to be small (Lee *et al.*, 2012, Long *et al.*, 2015a, Long *et al.*, 2015b).

Spontaneous mutations occur in the absence of strong selective pressure. Although the mutation rate may in some cases be affected by growth conditions and environmental factors, spontaneous point mutations are consistently found to occur at a rate of 10^{-10} to 10^{-9} per nucleotide per generation for many bacteria and growth conditions (Table 2). As a result, mutants are always the extreme minority within a population, making mutants and their corresponding mutations difficult to detect using bulk approaches. Mutations are also made difficult to detect biochemically due to the fact that replacement of one nucleotide with another does not give rise to a new chemical signature in the mutated DNA that can be conveniently detected *in vitro*. Therefore, genetic experiments are designed to detect rare mutants. In this section we will discuss several mutation detection methods: the widely used fluctuation test, the increasingly popular mutation accumulation experiment, and the newly emerging maximum-depth sequencing approach. Considerations for use of each method are discussed below.

Fluctuation tests and mutation reporter genes

The fluctuation test is the workhorse experiment for estimating spontaneous mutation rates in bacteria (Luria and Delbruck, 1943). The fluctuation test is useful in bacterial systems in that selection is applied to large bacterial populations to enable estimation of low mutation rates. In a fluctuation test, parallel cultures from small inocula are first grown in the absence of selection so that mutations may spontaneously arise (Figure 1A). Mutants that arise in culture proliferate exponentially along with the rest of the cells during the growth phase, often constituting a tiny fraction of each culture. After the growth phase, selection is applied under which only selectable mutants grow to form colonies on solid media (Figure 1A). Mutations allowing survival and growth during selection arise in mutation reporter genes, for which mutations affecting function of the gene product give rise to a selectable phenotype (Table 1, Figure 1B). Measuring the distribution of the number of mutant colonies arising from each culture in a fluctuation test enables estimation of mutation rate, which is the probability of mutagenesis per genome duplication.

Distribution of mutants per culture in fluctuation tests—Commonly applied methods for calculating mutation rate from fluctuation tests include the p_0 method, which derives the mutation rate from the proportion of cultures that do not generate any selectable mutants (Luria and Delbruck, 1943), the Lea-Coulson method, which calculates mutation rate from the median number of mutants per culture (Lea and Coulson, 1949), and the MSS method, which uses a maximum-likelihood approach to determine the most likely number of mutants per culture given the observed distribution of mutants per culture (Sarkar *et al.*, 1992). We encourage readers interested in methods of calculating mutation rate to consult these reviews (Rosche and Foster, 2000, Foster, 2006) and the following resource (Hall *et al.*, 2009).

In order to accurately infer mutation rate using mutation reporter genes, the assumptions of the fluctuation test must be valid (Luria and Delbruck, 1943, Lea and Coulson, 1949, Rosche and Foster, 2000, Foster, 2006). 1) Mutants should be selectively neutral during growth prior to selection. 2) Mutants must not be allowed to arise during selection. Because mutations will be inherited by the offspring of mutants, a mutant will proliferate to form a clone of

mutants within its culture (Figure 1A). Therefore, cultures in which mutants arose early have vastly more mutants than the mean number of mutants per culture (Figure 1A and 1C). Such cultures are termed “jackpots”, and they cause the distribution of mutants per culture in a fluctuation test to follow the Luria-Delbruck distribution when the assumptions of the test are met. When the assumptions are not met, however, the distribution of mutants per culture will be shifted toward a Poisson distribution.

When the distribution of the number of mutants per culture is shifted toward Poisson, this may indicate fitness defects of mutants in bulk culture, or post-selection mutagenesis in addition to spontaneous mutagenesis during clonal growth (Cairns *et al.*, 1988). Mutants that arise during selection typically form colonies after several days on selectable medium, but may not be distinguishable from slow-growing mutants that arose in culture prior to selection. Post-selection mutagenesis is particularly prone to occur when selecting for growth of mutants rather than for their survival (Figure 1A). Under conditions allowing post-selection mutagenesis, the total number of mutants observed in each culture is the sum of the number of mutants that arose spontaneously during growth in culture and the number that arose after selection was imposed (Figure 1A) (Cairns *et al.*, 1988). It is because of this that post-selection mutagenesis causes estimates of spontaneous mutation rate to be erroneously high.

Estimates of spontaneous mutation rate can be erroneously low due to phenotypic lag (Demerec, 1946, Newcombe, 1948, Newcombe and Scott, 1949, Kissling *et al.*, 2013). Phenotypic lag is the term for the delay in time between a mutant genotype arising and the display of the resulting phenotype. Phenotypic lag could emerge for a number of reasons when using mutation reporters. For instance, if the mutant allele is recessive to the wild type allele, sufficient time will be required to dilute or degrade the wild type protein after mutation of the reporter gene so that the mutant phenotype presents itself (Demerec, 1946). Similarly, in organisms with multiple copies of the genome or multi-fork replication, sufficient generations must transpire before a recessive allele will be expressed as a phenotype (Witkin, 1951, Kissling *et al.*, 2013). Even if the mutant allele is dominant, sufficient time will be required to express the gene product and yield an observable phenotype, resulting in phenotypic lag.

You get what you select for – types of mutation assays—There are many mutation reporters available for use in a fluctuation test. They differ in the type of selection that is applied, and their target size, which is the number of mutations that are observable by a mutation assay (Table 1). Depending on the mutation reporter used, selection for mutants can be for either their growth or survival. These aspects of mutation reporter genes may affect the validity of the two major assumptions of fluctuation tests listed above. With that in mind, we discuss several mutation reporter genes in this section.

Mutation assays can be classified into two groups: reverse and forward (Lieb, 1951, Skopek *et al.*, 1978). Reverse mutation assays rely on mutations in the reporter gene (reverse mutation reporter) that yield a functional product during selection (Figure 1B). In contrast, forward mutation assays generate a selectable phenotype upon acquiring loss-of-function mutations in the reporter gene (forward mutation reporter) (Figure 1B).

Reverse mutation assays: Reverse mutation assays detect mutations that alter a non-functional variant of a gene so that it becomes functional and encodes an active gene product after mutagenesis (Figure 1B). A fairly small target size and low diversity of mutation types, usually base pair substitutions, are observable in reverse mutation assays. A classic reverse mutation assay is the reversion of a mutant *lacZ* gene encoding β -galactosidase, which allows lactose utilization. Mutations causing reversion to an active *lacZ* gene product can be selected for growth on lactose as the only carbon source (Cupples and Miller, 1989, Cupples *et al.*, 1990). In 1989, Cupples and Miller developed a series of *lacZ* variants, each designed specifically to detect only one base pair substitution at a defined location (Cupples and Miller, 1989). Thus, with a target size of one, the Cupples and Miller reversion system is highly specific and the identity of the mutations yielding selectable mutants are known without the need for sequencing the mutation reporter gene. The *lacZ* locus was later modified to enable detection of insertions and deletions (Cupples *et al.*, 1990). These represent important improvements in reversion assays for detection of specific mutations.

Mutations in certain wild type genes can lead to antibiotic resistance phenotypes. Although these mutations are not reversion events *per se*, these assays are considered reverse mutation assays because the mutations yielding selectable mutants result in a gene product that is active under selection (Figure 1B). For example, fluoroquinolone antibiotics target DNA gyrase. Specific nonsynonymous substitutions in *gyrA*, the gene encoding DNA gyrase, result in gyrase variants that are functional in the presence of fluoroquinolone antibiotics (Yoshida *et al.*, 1990, Becket *et al.*, 2012). This enables the use of *gyrA* as a mutation reporter. The antibiotic rifampicin targets RNA polymerase. Certain non-synonymous substitutions in *rpoB*, which encodes the β subunit of RNA polymerase, render RNA polymerase uninhibited by rifampicin (Campbell *et al.*, 2001, Garibyan *et al.*, 2003). This allows estimation of mutation rate by selecting for resistance to rifampicin. Use of *rpoB* and *gyrA* as mutation reporters is simple and can be performed in many bacteria without the necessity of genome manipulations to set up a specific reporter system (Richardson and Stojiljkovic, 2001, Wang *et al.*, 2001, Koskiniemi *et al.*, 2010, Rock *et al.*, 2015, Deiham *et al.*, 2017).

Forward mutation assays: Forward mutation assays often have a much larger target size and diversity of detectable mutations (Figure 1B). This is because forward mutation assays detect mutations resulting in loss of function of the reporter gene rather than detecting mutations resulting in gain of function of the gene, as is the case for reverse mutation assays. Luria and Delbruck's fluctuation experiment was a forward mutation assay, as loss-of-function mutations in phage receptor genes render *E. coli* resistant to phage T1 (Hantke and Braun, 1978, Kadner *et al.*, 1980). A classical forward mutation assay with a more precisely defined target for mutagenesis is *lacI*. Loss-of-function mutations in *lacI*, which encodes the repressor of the *lac* operon, allow constitutive expression of the *lac* operon. This phenotype can be selected by requiring cells to grow on plates supplemented with phenyl- β -D-galactoside, which is not an inducer of *lac* operon expression but can be used as a carbon source if the *lac* operon is expressed (Coulondre and Miller, 1977, Schaaper *et al.*, 1986, Schaaper and Dunn, 1991). Loss-of-function mutations in *lacI* can occur throughout the gene's open reading frame and cis-regulatory elements, and selectable mutants can also arise

due to mutations in *lacO*, which is a cis-regulatory element for the *lac* operon and a binding site for LacI. For this reason, the target size when selecting for Lac⁺ mutants in the *lacI* forward mutation assay is much larger than that of the *lacZ* reverse mutation assay. However, the problem of selection for growth and mutagenesis during selection still exists.

More recently, a forward mutation assay was designed to select for survival of mutants and to ensure that mutations were generated prior to selection in *E. coli* (Yoshiyama *et al.*, 2001). The forward mutation reporter used was the *rpsL* gene, which encodes the small ribosomal subunit protein S12. A multicopy plasmid encoding wild type *rpsL* was allowed to replicate and accumulate mutations in *E. coli*. Owing to the multicopy nature of the plasmid, mutations in plasmid-encoded *rpsL* were recessive and did not cause fitness defects in culture. The plasmid population was then purified and used to transform *E. coli* cells in single copy (Yoshiyama *et al.*, 2001). The cells into which the plasmid library was transformed harboured a chromosomally-encoded recessive allele of *rpsL* yielding streptomycin resistance. Therefore, cells transformed with a plasmid encoding wild type S12 expressed the dominant, wild type *rpsL* phenotype and were sensitive to streptomycin. In contrast, cells receiving a plasmid encoding inactive S12 were streptomycin resistant. Thus, transformants that were streptomycin resistant were inferred to have been transformed with a plasmid in which loss-of-function mutations occurred in *rpsL* (Fujii *et al.*, 1999, Yoshiyama *et al.*, 2001).

A fluctuation test to detect spontaneous mutations on the chromosome of *Bacillus subtilis* was recently developed to use the *thyP3* gene, which encodes thymidylate synthetase, as a forward mutation reporter. A native, temperature sensitive thymidylate synthase, ThyB(ts) (Neuhard *et al.*, 1978) was used as a phenotypic buffer during growth at its permissive temperature. This ensured that loss-of-function mutations in *thyP3* did not cause fitness defects in culture prior to selection (Sankar *et al.*, 2016). Therefore, mutations in *thyP3* arose during growth at the permissive temperature for ThyB(ts) and were selected for survival of trimethoprim treatment at the nonpermissive temperature for ThyB(ts) (Sankar *et al.*, 2016). This allowed thymidylate synthase to be used as a forward mutation reporter that accurately measured mutation rate in a fluctuation test without confounding fitness defects in culture and by selecting for survival of mutants rather than for their growth.

Mutation accumulation experiments

Although mutation reporters are easily used in bacterial systems, they are constrained by the mutations that can yield selectable mutants, and thus have a limited target size. Reductions in the cost of whole genome resequencing have led to a burst of mutation accumulation experiments in which mutations are detected by resequencing entire genomes. In comparison to mutation reporters, mutation accumulation experiments have a much larger target size: the entire mappable genome.

Mutation accumulation experiments allow detection of any mutation that is not extremely deleterious, despite possible fitness defects (Barrick and Lenski, 2013). This is achieved by subjecting a population to single-cell bottlenecks at regular intervals (Figure 1D). Serially bottlenecking populations in mutation accumulation experiments allows a lineage to accumulate mutations in a manner that is essentially neutral with respect to evolution.

Accordingly, mutations which would have been eliminated from the population due to a fitness disadvantage instead become fixed in the population after bottlenecking (Lind and Andersson, 2008). This experimental design is particularly useful for studying bacterial mutagenesis, because introducing the bottleneck is as simple as performing colony purification by passaging single bacterial colonies from one plate to another (Figure 1D). This, in combination with the dramatic reduction in whole genome resequencing cost over the past decade, has enabled several studies to make use of mutation accumulation experiments followed by whole genome resequencing to investigate mechanisms of spontaneous mutagenesis in bacteria (Lind and Andersson, 2008, Lee *et al.*, 2012, Foster *et al.*, 2015, Long *et al.*, 2015b, Rock *et al.*, 2015, Sung *et al.*, 2015, Bhagwat *et al.*, 2016, Lee *et al.*, 2016, Schroeder *et al.*, 2016, Dillon *et al.*, 2017). A large spectrum of distinct mutation types is observable using the mutation accumulation approach, from nonsynonymous single-nucleotide variants and small insertions/deletions within genes, which are also detectable by reporters, to silent mutations and intergenic mutations, which often are not detectable by reporters. In addition, large structural rearrangements of the chromosome and movement of mobile genetic elements such as insertion sequences are detectable in the mutation accumulation approach (Lee *et al.*, 2012, Lee *et al.*, 2016).

Maximum-depth sequencing

Maximum-depth sequencing is a recently developed approach to infer spontaneous mutation rates of specific regions of interest in a genome (Jee *et al.*, 2016). High-throughput sequencing libraries are prepared using bulk genomic DNA from a culture of cells and a primer specific to the region of interest to target sequencing to that position within the genome (Figure 1E). The primer contains an adapter sequence with a type of barcode termed a “unique molecular identifier”, which can be used to determine which reads represent the sequence of a single, shared original molecule (Jee *et al.*, 2016). Genomic regions of interest are directly tagged with a unique molecular identifier and linearly amplified. After high-throughput sequencing, a consensus sequence is determined for each pool of reads corresponding to a single region of interest and unique molecular identifier (Figure 1E). The error rate for consensus sequence calling is inversely proportional to the sequencing depth for each region of interest/unique molecular identifier pool. Therefore, by increasing sequencing depth, error rate for consensus calling can be reduced indefinitely. Even lethal mutations are detectable by this method, highlighting its use for detecting deleterious mutations. While maximum-depth sequencing cannot be used to study mutagenesis genome-wide, it can be performed much more quickly than the mutation accumulation approach.

Mechanisms of spontaneous mutagenesis

In this section, we discuss insights into spontaneous mutagenic mechanisms gained through use of the methods described above. While mechanisms that prevent or correct spontaneous mutagenesis are well understood (Lenhart *et al.*, 2012), less is known about the mechanisms by which normal cellular processes promote spontaneous mutagenesis. These processes include DNA replication, transcription, recombination, and metabolic processes that generate reactive species that damage DNA, i.e., reactive oxygen species. These processes may interact. For example, mutagenesis can be promoted by DNA replication and

transcription as well as by conflicts between the two processes. Here we review our current understanding of the major paths toward spontaneous mutagenesis in bacteria (Figure 2).

DNA replication errors

DNA polymerases that replicate the genome are extremely accurate in terms of their ability to correctly pair nucleotides, but they do make errors. These errors include mispairing of noncognate dNTPs in violation of Watson-Crick rules and incorporation of noncanonical nucleotide substrates such as ribonucleotides or oxidized dGTP (8-oxo-dGTP, or 8-oxo-G). In addition, DNA replication can also be mutagenic by primer-template slippage or template switching, resulting in insertions, duplications, and deletions to the genome. Each mutagenic mechanism is discussed below.

Base pairing errors—DNA polymerases that replicate bacterial genomes form base pairing errors that cause mutations if they are not corrected before the next round of DNA replication by the proofreading activity of the replicative DNA polymerase or the DNA mismatch repair system (MMR). Base pairing error generation, evasion of proofreading activity, and evasion of MMR each contribute to mutagenesis, but the mechanisms and factors affecting each are somewhat poorly understood.

Leading- and lagging strand replication: In *S. cerevisiae* base pair substitutions are more common during lagging strand replication due to extension of RNA primers by the error-prone DNA polymerase, Pol α (Lujan *et al.*, 2014, Reijns *et al.*, 2015). Errors made by Pol α are proofread by Pol δ (Pavlov *et al.*, 2006), but DNA-binding proteins such as nucleosomes and transcription factors may discourage such proofreading, providing a mechanism for errors made by Pol α during lagging strand replication to remain in the genome and become mutations in the following round of DNA replication (Reijns *et al.*, 2015).

Considerably less is understood about mechanisms promoting differences in base pairing error rates between leading- and lagging strand replication in bacteria. Available literature suggests that *E. coli* Pol III may be more accurate in base pairing of nucleotides in the lagging strand, but the mechanism for this is unknown (Fijalkowska *et al.*, 1998). Mutation accumulation studies have shown a bias in distribution of transitions between the leading- and lagging strands (Lee *et al.*, 2012, Sung *et al.*, 2015, Schroeder *et al.*, 2016), but it remains unclear whether the errors responsible for the mutations were produced during leading- or lagging strand replication. DNA damage in the lagging-strand template of *B. subtilis* may enlist an error-prone translesion DNA polymerase, PolY1, to enable mutagenic synthesis of the lagging strand past the damaged site (Million-Weaver *et al.*, 2015).

However, loss of PolY1 in *B. subtilis* does not alter the spontaneous mutation rate in a mutation reporter (Duigou *et al.*, 2004). Overexpression of PolY1, however, does increase mutagenesis, but strand bias of this effect has not been demonstrated (Duigou *et al.*, 2004).

It is therefore unclear the extent to, and mechanism by which PolY1 affects spontaneous mutagenesis in *B. subtilis*.

Rare tautomer hypothesis: Rare tautomeric forms of bases in DNA are a mechanism for transitions, as proposed by the rare tautomer hypothesis (Watson and Crick, 1953, Topal and Fresco, 1976). Under normal circumstances, a base pairing error will cause movement of the mismatched 3' terminus of the nascent strand from the polymerase active site of DNA polymerase to its exonuclease active site so that proofreading may occur (Fernandez-Leiro *et al.*, 2017). However, if a base is present in its rare tautomeric form, the geometry of base-pairing will be very close to that of correctly paired DNA (Figure 3A), allowing the 3' terminus to remain in the polymerase active site (Kool, 2002, Harris *et al.*, 2003, Kim *et al.*, 2005). It is difficult to test the rare tautomer hypothesis due to the low tautomeric constant of the bases found in DNA. However, by using the nucleoside analog dP (6-(2-deoxy- β -D-ribofuranosyl)-3,4-dihydro-6H,8H-pyrimido[4,5-c][1,2]oxazin-2-one), which readily tautomerizes to pair with either adenine or guanine, it was found that the tautomeric constant of the P base is highly correlated with its ability to be included as a mispair in nascent DNA (Harris *et al.*, 2003). This work was carried out using proofreading-deficient Klenow fragment, so the effect of tautomerization on proofreading activity could not be discerned. The first direct evidence in support of the rare tautomer hypothesis came in 2011, when the structure of a DNA polymerase with a C·A mispair in the active site was solved with the C in its tautomeric form (Wang *et al.*, 2011). This study represents strong evidence that base pairing errors with similar geometry to cognate Watson-Crick pairing can be generated when a base is present as a rare tautomer. It is poorly understood whether tautomerization of nucleobases can be affected by factors such as local sequence context *in vivo*. It is also not known to what extent rare tautomeric bases remain in DNA after their incorporation compared to the frequency with which they switch back to their canonical tautomeric form. In addition, further study will be necessary to determine whether rare tautomers primarily affect incorporation of base pairing errors into DNA by DNA polymerases versus evasion of proofreading and MMR.

Sequence context: Local sequence context affects the rate at which base pairing errors are produced, proofread, and corrected by MMR. One mechanism by which sequence context can affect the accuracy of DNA polymerases is by base stacking interactions between neighboring nucleotides in DNA. Specifically, purines 5' to a given position stack more stably than pyrimidines [for reviews, see (Goodman *et al.*, 1993, Hunter, 1993, Kool, 2001)]. In agreement with this, a mutation accumulation study using MMR-deficient *S. cerevisiae* demonstrated that proofreading-deficient DNA polymerases were more likely to misinsert thymidine across from cytidine or guanosine when a purine was at the 3' end of the primer strand (Lujan *et al.*, 2014). Although this has not been tested in bacterial systems, base stacking interactions are an intrinsic property of all DNA. It is therefore reasonable to extrapolate the observations made in *S. cerevisiae* to bacterial systems.

Proofreading of errors made by DNA polymerases is also affected by local sequence context (Sinha, 1987). High GC content 5' or 3' to the mismatched nucleotide in nascent DNA makes proofreading less likely (Petruska and Goodman, 1985, Carver *et al.*, 1994). The mechanism for this is thought to be the stability of base pairing of GC-rich DNA relative to AT-rich DNA. More stable base pairing would decrease the ability of the mismatched 3' terminus to enter the exonuclease (proofreading) active site of DNA polymerase (Carver *et*

et al., 1994). Effects of sequence context on proofreading activity *in vivo* can be estimated by observing the relative mutation rate in differing contexts in cells without MMR, because most mutations in the absence of MMR are due to replicative DNA polymerase errors (Schaaper and Dunn, 1987, Schaaper, 1993). Mutation accumulation studies conducted using bacteria lacking MMR have demonstrated that neighboring G-C base pairs strongly increase the base pair substitution rate at a given genome position *in vivo* (Lee *et al.*, 2012, Sung *et al.*, 2015, Schroeder *et al.*, 2016).

The effect of local sequence context on MMR efficiency has been more difficult to estimate, because it requires comparison of substitution rates in various contexts from MMR-deficient and MMR-proficient lines, which accumulate few mutations. However, one study showed that MMR is particularly efficient at correcting mismatches neighbored by G-C base pairs (Lee *et al.*, 2012). Now that ample mutation accumulation experiments have been performed in MMR-proficient *E. coli* (Lee *et al.*, 2012, Foster *et al.*, 2015, Bhagwat *et al.*, 2016), the question of how local sequence context affects MMR efficiency genome-wide can be more rigorously examined.

Use of 8-oxo-dGTP as a substrate for DNA replication—The nucleotide pool is affected by endogenous reactive oxygen species generated by respiration. Of the bases in DNA, guanine is the most frequent target for oxidation, forming 8-oxo-7,8-dihydroguanine (8-oxo-G) (Cadet *et al.*, 2008, Kanvah *et al.*, 2010). Genetic and biochemical evidence indicates that DNA polymerase III incorporates 8-oxo-dGTP into DNA and is inefficient at proofreading 8-oxo-dGMP after its incorporation (Yamada *et al.*, 2012). 8-oxo-dGMP can also be generated when dGMP already in genomic DNA reacts with reactive oxygen species. The bases in nucleotides can adopt *anti*- or *syn* conformations. When 8-oxo-G is in the *anti* conformation it will pair with cytidine. However, it more often adopts the *syn* conformation and will readily pair with adenosine (Figure 3B) (Kasai *et al.*, 1991, Kouchakdjian *et al.*, 1991, McAuley-Hecht *et al.*, 1994). Structural evidence has suggested that when the newly synthesized DNA is rigid within the active site of a DNA polymerase, 8-oxo-G may be more mutagenic than when the newly synthesized DNA is flexible (Krahn *et al.*, 2003). This is because 8-oxo-G is in the *syn* conformation that is required for its pairing with adenosine when the newly replicated double-stranded DNA is rigid (Krahn *et al.*, 2003).

Mutations caused by 8-oxo-G are either A·T → C·G transversions or G·C → T·A transversions, depending on how 8-oxo-G enters genomic DNA. When 8-oxo-dGTP is incorporated across from adenosine during DNA replication, the following round of DNA replication may incorporate cytidine across from 8-oxo-G if the base has switched to the *anti* conformation, causing A·T → C·G transversions. Conversion of guanine to 8-oxo-G at a G·C base pair within DNA can lead to insertion of adenosine instead of cytidine when *syn*-8-oxo-G is used as a template for the next round of replication, causing G·C → T·A transversions (Miller, 1996). The oxidized guanine (GO) system limits mutagenesis by sanitizing nucleotide pools of 8-oxo-dGTP, removing oxidized 8-oxo-G bases in double-stranded DNA, and removing adenine paired with 8-oxo-G in double-stranded DNA (Michaels and Miller, 1992, Lenhart *et al.*, 2012). Work performed in *E. coli* used mutation accumulation in genetic backgrounds eliminated for several DNA repair pathways to determine the extent to which different mutagenic mechanisms affect spontaneous mutation

in unstressed conditions. Strikingly, this work showed that, other than the mismatch repair pathway, loss of the GO system was the only pathway out of many tested that showed an increase in base pair substitution rates (Foster *et al.*, 2015), with transversion rates increasing in the absence of the GO system.

Ribonucleotide misincorporation—Ribonucleotides enter genomic DNA via their misincorporation by DNA polymerases during DNA replication (Wang *et al.*, 2012, Yao *et al.*, 2013). Ribonucleotide misincorporation is by far the most frequent error made by DNA polymerases, with the error rate largely driven by the inequity of the cellular nucleotide pools, as NTPs vastly outnumber dNTPs (Traut, 1994, Buckstein *et al.*, 2008, Ferraro *et al.*, 2010, Nick McElhinny *et al.*, 2010a, Nick McElhinny *et al.*, 2010b, Reijns *et al.*, 2012, Yao *et al.*, 2013). For instance, the *E. coli* DNA polymerase III core with the clamp loader complex associated incorporates ribonucleotides into nascent DNA *in vitro* at a rate that would result in ≈ 2000 misinsertions per genome replication event *in vivo* (Yao *et al.*, 2013). This is in stark contrast to the rate of base pairing errors produced by *E. coli* DNA replication, during which about 7.9 genome replication events occur between single base pairing errors escaping proofreading (Lee *et al.*, 2012). Despite their high frequency of misincorporation into genomic DNA and their susceptibility to hydrolysis of the sugar-phosphate backbone, ribonucleotides covalently embedded in genomic DNA have a rather mild effect on bacterial mutation rate, yielding a small increase in base pair substitutions in *B. subtilis* (Yao *et al.*, 2013).

Ribonucleotide excision repair (RER) is the biochemical pathway responsible for replacing ribonucleotides in DNA. Loss of RER in *Saccharomyces cerevisiae* leads to an increased rate of 2–5 nucleotide deletions (Nick McElhinny *et al.*, 2010a, Lujan *et al.*, 2014). These deletions are dependent on *S. cerevisiae* topoisomerase 1, encoded by the *TOP1* gene (Kim *et al.*, 2011), and exhibit a strand-biased distribution (Williams *et al.*, 2013, Cho *et al.*, 2015). Bacteria lack a clear homolog of yeast topoisomerase 1, and although loss of RER causes a mild increase in mutation rate, the nature of the changes are unknown (Yao *et al.*, 2013). One explanation for the low mutability of ribonucleotides misincorporated by DNA polymerase during bacterial replication is that nucleotide excision repair (NER) is able to remove ribonucleotides in the absence of RER (Vaisman *et al.*, 2013, Vaisman *et al.*, 2014). Removal of ribonucleotides by NER in *E. coli* is error-free in contrast to their removal by topoisomerase 1 in yeast (Yao *et al.*, 2013). Despite a growing volume of research on bacterial RER, the overall mechanism of RER and the mechanisms of increased mutagenesis due to unrepaired ribonucleotides in bacterial DNA are poorly understood.

Strand misalignment—The mutagenic mechanisms discussed above involve errors made during DNA replication. In contrast to mutations arising from misinsertion of nucleotides during DNA replication, a class of mutations termed “encoded errors” arises due to accurate insertion of nucleotides using a misaligned primer-template pair [reviewed in (Lovett, 2004)]. During normal DNA replication the DNA polymerase can move from using one template to another. This type of strand misalignment is termed template switching, and it can lead to a variety of mutations, including base pair substitutions, insertions and deletions, and large structural rearrangements (Yoshiyama *et al.*, 2001, Dutra and Lovett, 2006, Seier *et*

al., 2011, Anand *et al.*, 2014, Tsaponina and Haber, 2014). Quasipalindromic sequences promote template switching by forming a hairpin structure during DNA replication such that one arm of the quasipalindrome serves as the template for the other arm (Ripley, 1982). Quasipalindromes are mutation hotspots in bacteria due to the increased likelihood of template switching events at hairpins (Viswanathan *et al.*, 2000, Dutra and Lovett, 2006, Seier *et al.*, 2011, Seier *et al.*, 2012). In addition, replication fork stalling, due to either complex DNA structures in eukaryotes (Lee *et al.*, 2007) or replication-transcription conflict in bacteria (Sankar *et al.*, 2016) can promote template switching.

Long homopolymer runs are hotspots for small insertions and deletions in bacteria (Streisinger *et al.*, 1966, Fujii *et al.*, 1999, Lee *et al.*, 2012, Schroeder *et al.*, 2016) due to a type of primer-template misalignment termed “strand slippage” during DNA replication (Streisinger *et al.*, 1966, Canceill *et al.*, 1999, Fujii *et al.*, 1999). Small insertions and deletions are enriched intergenically in *B. subtilis* due to intergenic enrichment of homopolymer runs (Schroeder *et al.*, 2016). The relative paucity of homopolymer runs in coding DNA likely reflects purifying selection against genes containing frameshift-promoting homopolymers.

DNA deamination

A major source of base substitutions is from DNA damage by deamination of the bases in nucleotides. For example, cytosine deamination leads to uracil. If uridine in DNA goes unrepaired, adenosine will be inserted into the nascent DNA strand when replication uses uridine as a template, causing a G·C → A·T transition. Nucleobases in single-stranded DNA are more susceptible to deamination than those in double-stranded DNA. For instance, cytosine deamination occurs with a rate constant of approximately 1×10^{-10} per cytosine site per second in single-stranded DNA (Figure 3C) (Frederico *et al.*, 1990). This is in stark contrast to its rate in double-stranded DNA, which is 7×10^{-13} per cytosine site per second (Frederico *et al.*, 1990). This represents about a 140-fold lower rate of cytosine deamination in double-stranded DNA compared to single-stranded DNA. Therefore, by transiently generating regions of single-stranded DNA, transcription and DNA replication strongly increase the likelihood that nucleobases will be deaminated. It is unclear what the relative contributions of transcription and DNA replication are to overall mutagenesis by their generation of single-stranded DNA. Our current understanding of the contribution of replication and transcription to this mutagenic mechanism is discussed below.

Single-stranded DNA generated during DNA replication—DNA replication is semidiscontinuous and generates transient patches of single-stranded DNA in the lagging-strand template throughout the genome. These single-stranded DNA patches are a major source of spontaneous cytosine deamination in bacteria (Bhagwat *et al.*, 2016). Given the approximate rate of spontaneous cytosine deamination in single-stranded DNA such as that found in the lagging-strand template, replication-associated deamination of cytosine to uracil could be a significant contributor to G·C → A·T transitions in bacteria. The time from cytosine deamination in the lagging-strand template to replication over the resulting uridine is extremely short, and the uridine-containing template will be quickly used as a template for insertion of adenosine into the nascent lagging strand. Uracil-DNA glycosylase (Ung)

removes uracil from DNA, leaving an abasic site. If an abasic site is produced in the lagging-strand template by Ung, the replicative DNA polymerase would stall at the site. A translesion DNA polymerase would then be required to synthesize DNA over the abasic site. In that event, the translesion DNA polymerase is likely to insert adenosine across from the abasic site, resulting in a G·C → A·T transition (Lawrence *et al.*, 1990, Reuven *et al.*, 1999, Tang *et al.*, 1999).

Mutation accumulation studies of *E. coli*, *B. subtilis* and *Mesoplasma florem* have shown that G·C → A·T transitions are much more likely when cytidine is in the lagging-strand template than when guanosine is in the lagging-strand template (Sung *et al.*, 2015, Schroeder *et al.*, 2016). Cytosine deamination in the lagging-strand template may play a role in this strand bias. Deamination of cytosine in the lagging-strand template has also been proposed to contribute to the nearly universal enrichment of guanosine compared to cytosine in the lagging-strand template (GC skew) of bacteria (Lobry, 1996, Bhagwat *et al.*, 2016).

Single-stranded DNA generated during transcription—RNA polymerase uses one strand of DNA as its template strand to pair with the newly synthesized RNA, leaving the non-template strand unpaired (Gnatt *et al.*, 2001). Since the non-template strand becomes locally exposed to solvent as single-stranded DNA during transcription, it is prone to damage by various mechanisms, including cytosine deamination to uracil. As discussed, uridine pairs with adenosine when DNA is replicated, causing G·C → A·T transitions (Beletskii and Bhagwat, 1996). Mutation accumulation experiments in *Salmonella typhimurium* showed that G·C → A·T transitions were made more likely in coding sequences of Ung-deficient cells if the cytosine were in the non-template strand for transcription and if the gene were highly expressed, thereby exposing it during transcription and making it more likely to undergo deamination (Lind and Andersson, 2008). Most bacterial genes are transcribed co-directionally with DNA replication (Rocha, 2004), so the lagging-strand template for replication and the non-template strand for transcription are often the same physical strand of DNA. We find it likely, therefore, that strand-biased occurrence of G·C → A·T transitions (Sung *et al.*, 2015, Schroeder *et al.*, 2016) arises due to a combination of cytosine deamination in both the lagging-strand template at the replication fork (Bhagwat *et al.*, 2016) and single-stranded DNA present in the non-template strand during transcription (Beletskii and Bhagwat, 1996, Lind and Andersson, 2008). These strand-biased mutagenic mechanisms are in addition to differences in leading- and lagging strand replication fidelity (Fijalkowska *et al.*, 1998).

In addition to promoting cytosine deamination in coding sequences, transcription encourages premutagenic DNA damage in promoter regions. During transcription initiation, the promoter sequences are melted but are not base paired with RNA because the promoter region is upstream of the transcription start site. Thus, both the template strand and the non-template strand sequences may be susceptible to deamination during initiation before promoter escape by RNA polymerase. Forward mutation assays have shown that the conserved T₋₇ position in the -10 element of promoters for head-on genes is a hotspot for A·T → G·C transitions in both *B. subtilis* (Sankar *et al.*, 2016) and *E. coli* (Yoshiyama *et al.*, 2001). During transcription initiation, T₋₇ of the non-template strand is sequestered inside the housekeeping sigma factor during transcription initiation, leaving the complementary

adenosine (A₇) unpaired as single-stranded DNA (Feklistov and Darst, 2011). Adenine in single-stranded DNA can be spontaneously deaminated to hypoxanthine (Figure 3C), which readily pairs with cytidine in the next round of DNA replication, causing A·T → G·C transitions. This mechanism may explain other base substitutions observed in cis-regulatory elements of genes in forward mutation assays (Schaaper *et al.*, 1986, Yoshiyama *et al.*, 2001, Sankar *et al.*, 2016). The fact that promoter substitutions have not been more commonly observed in mutation assays is likely due to the prevalent use of reversion assays that only detect coding substitutions. In addition, promoter mutations are likely to cause major gene expression defects, and are therefore strongly selected against in essential genes and may be difficult to observe in mutation accumulation studies. However, the importance of such mutations should not be underestimated due to their fundamental impact in altering gene expression. Collectively, these studies show that the non-transcribed strand is prone to chemical modifications that have the potential to increase spontaneous mutagenesis of both coding sequences and regulatory regions of highly transcribed genes.

Replication-transcription conflict

DNA replication and transcription must both use the same DNA template. DNA replication forks move much faster than transcription in bacteria, leading to unavoidable conflicts between these two fundamental processes. There are two types of replication-transcription conflicts, depending upon gene orientation relative to DNA replication (Rocha, 2004). Co-directional conflict arises in genes for which the coding strand for transcription corresponds to the leading-strand template for replication, whereas head-on conflicts occur when the coding strand of a gene is the lagging-strand template for replication. Both types of conflict can lead to replication fork stalling, homologous recombination, and genome instability (Mirkin *et al.*, 2006, Srivatsan *et al.*, 2010, Dutta *et al.*, 2011, Merrikh *et al.*, 2011). We will focus on the mechanisms by which replication-transcription conflicts generate spontaneous mutations within the promoter and the coding region of a gene.

Head-on conflict between replication and transcription impedes replication fork progression and has been proposed to be more mutagenic than co-directional conflict (Vilette *et al.*, 1995, Vilette *et al.*, 1996, Mirkin and Mirkin, 2005, Srivatsan *et al.*, 2010). In *B. subtilis*, increased nonsynonymous substitution rates in the coding sequence of head-on genes have been observed and proposed to be generated by replication-transcription conflict through the recruitment of the error-prone translesion polymerase PolY1 (Paul *et al.*, 2013, Million-Weaver *et al.*, 2015). However, a role for PolY1 in spontaneous mutagenesis in *B. subtilis* remains unclear (Duigou *et al.*, 2004). Other studies using either a forward mutation assay or a mutation accumulation approach did not find strong evidence for replication-transcription conflict in generating base pair substitutions in coding sequences (Sankar *et al.*, 2016, Schroeder *et al.*, 2016). Therefore, the contribution of replication-transcription conflict on spontaneous base pair substitution rates in coding sequences has remained unclear.

Replication-transcription conflicts contribute strongly to insertions and deletions within a gene's coding sequence. Insertions and deletions were found to coincide with the locations where replication and RNA polymerase first encounter each other, implicating replication fork stalling due to conflicts with transcription as a trigger for template switching by DNA

replication or double-strand breaks that result in insertions and deletions (Dutta *et al.*, 2011, Sankar *et al.*, 2016). In addition, collision of the replication fork with either the transcription initiation complex or transcriptional repressors can occur (Mirkin *et al.*, 2006), generating insertions and deletions in promoter regions (Vilette *et al.*, 1992, Vilette *et al.*, 1996, Sankar *et al.*, 2016).

In addition, T → C transitions at the promoter increase when transcription is driven head-on with replication (Yoshiyama *et al.*, 2001, Sankar *et al.*, 2016). One explanation for this is that head-on conflicts interfere with transcription initiation at a step after promoter open complex formation, either by promoting abortive initiation or prolonging single-stranded DNA exposure. If promoter substitutions are a major source of mutagenesis due to head-on replication-transcription conflict, it may explain why essential genes are highly enriched to be transcribed co-directionally with replication, as promoter mutations in head-on genes will result in strong purifying selection against a head-on orientation for essential genes.

Spontaneous induction of the SOS response

Although the SOS response is often thought of as being induced by externally imposed stress such as exposure to ultraviolet light and DNA-damaging agents, approximately 1% of exponentially growing *E. coli* are undergoing the SOS response in the absence of exogenous stress (Pennington and Rosenberg, 2007). Three distinct DNA polymerases, Pol IV, Pol V, and Pol II are upregulated during the *E. coli* SOS response. Pals IV and V are Y-family DNA polymerases (Jarosz *et al.*, 2007). Pol IV is encoded by the gene *dinB*. Pol IV is accurate when replicating over adducts on the N² position of guanine nucleotides (Jarosz *et al.*, 2006, Yuan *et al.*, 2008), but Pals IV and V are both error-prone on undamaged DNA (Goodman and Woodgate, 2013). Increased expression of Pals IV and V, such as during SOS in *E. coli*, yields an elevated mutation rate (MacLean *et al.*, 2013). Pol IV makes single-nucleotide deletions due to its ability to extend misaligned primer-template pairs (Wagner *et al.*, 1999). To a lesser extent, Pol IV also produces base pair substitutions (Wagner *et al.*, 1999, Tang *et al.*, 2000). Pol V is encoded by the genes *umuC* and *umuD*. Pol V has been proposed to be primarily responsible for the increased mutation rate during the SOS response; it efficiently incorporates nucleotides across from abasic sites and UV lesions such as thymidine dimers in DNA and produces transversions (Steinborn, 1978, Tang *et al.*, 1999, Tang *et al.*, 2000, Curti *et al.*, 2009, Robinson *et al.*, 2015). Pol II is encoded by the *polB* gene. Pol II has exonucleolytic proofreading activity, and is therefore quite accurate and does not contribute greatly to SOS-induced mutagenesis (Banach-Orlowska *et al.*, 2005). Rather, Pol II may correct errors made by Pals IV and V or remove mismatches at the nascent strand 3' end prior to recruitment of Pals IV and V for extension, thereby reducing mutagenesis during SOS (Banach-Orlowska *et al.*, 2005). Pals IV, V, and II can displace Pol III at the replication fork during SOS (Sutton and Duzen, 2006, Furukohri *et al.*, 2008, Curti *et al.*, 2009, Kath *et al.*, 2014, Kath *et al.*, 2016), so it is reasonable to expect that spontaneous induction of the SOS response may contribute to mutagenesis during unstressed conditions in *E. coli*. However, despite SOS induction in about 1% of exponentially growing *E. coli* in the absence of exogenous stress (Pennington and Rosenberg, 2007), *E. coli* lacking Pals IV and V, or all three SOS-induced DNA polymerases have no detectable change in base pair substitution or insertion/deletion rate in mutation accumulation experiments,

suggesting they may not provide a substantial source of mutagenesis during normal growth (Foster *et al.*, 2015).

The requirements for SOS induction differ substantially between *B. subtilis* and *E. coli* (Simmons *et al.*, 2009). In *B. subtilis* and many other bacteria, two replicative DNA polymerase subunits, PolC and DnaE, are involved in Pol III function (Dervyn *et al.*, 2001, Foster *et al.*, 2003, Titok *et al.*, 2006). In *B. subtilis*, DnaE is required to extend RNA primers during DNA replication prior to switching with the more processive and proofreading capable PolC that synthesizes the rest of the lagging strand (Sanders *et al.*, 2010). DnaE is error-prone due to its lack of proofreading activity, and expression of DnaE is induced during SOS (Le Chatelier *et al.*, 2004). PolY1 (YqjH) and PolY2 (YqjW) in *B. subtilis* are homologous to *E. coli* Pol IV and UmuC, respectively (Sung *et al.*, 2003). *B. subtilis* does not have a clear Pol II homolog. PolY1 is constitutively expressed in *B. subtilis* (Duigou *et al.*, 2004). PolY2 is SOS-induced and is required for UV-induced mutagenesis (Duigou *et al.*, 2004). Although overexpression of either PolY1 or Y2 increased spontaneous mutagenesis, deletion of *yqjH*, *yqjW*, or both genes, had no detectable effect on spontaneous mutagenesis (Duigou *et al.*, 2004). It is currently unclear whether spontaneous SOS induction during normal growth plays a role in spontaneous mutagenesis in *B. subtilis*.

Homologous recombination

Homologous recombination can cause genome rearrangements via crossover between non-allelic ribosomal RNA operons on sister chromosomes (Reams *et al.*, 2014). Additionally, homologous recombination between similar genes within a genome causes base pair substitutions in a process termed “gene conversion” (Abdulkarim and Hughes, 1996, Paulsson *et al.*, 2017). However, mutagenic mechanisms of homologous recombination remain poorly characterized. This is largely attributable to the methods commonly used to detect mutations. Mutation reporters often yield information preferentially on base pair substitutions and small insertions and deletions and will not detect large structural changes in a genome that are often associated with homologous recombination. Extraction of information on large genome rearrangements in mutation accumulation studies is not frequently performed, but it could reveal more details on the rates and mechanisms of mutagenesis due to homologous recombination. Recently, mutation accumulation methodology has been extended to examine the movement of a specific class of mobile genetic element, insertion sequences, in *E. coli*.

Insertion sequences

Insertion sequences (ISs) are a class of mobile genetic element present in diverse bacteria. ISs usually bear only the open reading frame encoding the transposase that encourages their mobilization [reviewed in (Siguier *et al.*, 2006, Darmon and Leach, 2014)]. IS elements represent a previously underappreciated source of spontaneous mutagenesis in bacteria because their mobilization within a genome has been difficult to detect using classical genetic methods. A recent analysis of IS dynamics used whole genome resequencing data from 520 *E. coli* mutation accumulation lines from various studies (Lee *et al.*, 2016). Strikingly, the rate of insertion of IS elements is within an order of magnitude of the rates of base pair substitution and insertion/deletion per genome per generation in *Deinococcus*

radiodurans and *E. coli* (Lee *et al.*, 2014, Long *et al.*, 2015a, Lee *et al.*, 2016, Shewaramani *et al.*, 2017). ISs insert at a rate of 3.5×10^{-4} per genome per generation, and homologous-recombination-mediated deletions between ISs occur at a rate of 4.5×10^{-5} per genome per generation in *E. coli*. By comparison, the base pair substitution rate in *E. coli* is $\approx 1 \times 10^{-3}$ per genome per generation (Foster *et al.*, 2015).

It was determined that nine of the fourteen IS families in *E. coli* MG1655 were active, as five of the families showed no evidence of movement in any mutation accumulation line. IS *IA* and IS *5A* families were the most active, accounting for $\approx 80\%$ of all insertions and deletions due to ISs. Another mutation accumulation study compared the mutation rates and spectra of *E. coli* mutation accumulation lines passaged under either aerobic or anaerobic conditions. The major finding of this study was that the insertion rate of a specific IS element, IS *150*, increased under anaerobic conditions (Shewaramani *et al.*, 2017). The mechanism by which anaerobic growth promotes IS *150* mobilization is currently unclear. Mutation accumulation studies have thus provided evidence for ISs being a common source of spontaneous mutagenesis in the *E. coli* genome.

Concluding remarks

Spontaneous mutagenesis is a major source of genetic diversity leading to evolution of species and antibiotic resistance in bacteria. Recent progress in spontaneous mutagenesis has demonstrated that spontaneous mutations arise from endogenous cellular processes such as DNA replication, transcription and replication-transcription conflict. Local sequence context heavily influences base pair substitution rate due to a combination of effects on base pair misinsertion and proofreading by DNA polymerases, and DNA mismatch repair efficiency. In addition, spontaneous mutagenesis extends far beyond point mutations in coding sequences of genes, with many mutations occurring in regulatory elements and intergenic regions. Finally, movement of and homologous recombination between insertion sequences provides a substantial source of spontaneous mutagenesis in bacteria. Further work on spontaneous mutagenesis in bacteria will continue to provide the foundation for a deeper understanding of mutagenic mechanisms and their contributions to evolution, antibiotic resistance and cancer.

Acknowledgments

We would like to thank members of the Wang laboratory for their helpful feedback on this review.

LAS was supported by NIH R01 GM107312. JDW was supported in part by NIH R01 GM084003. JDW is an HHMI Faculty Scholar.

Biographies

Jeremy Schroeder is a postdoctoral fellow at the University of Wisconsin – Madison working to understand mechanisms promoting diversity within bacterial populations. Dr. Schroeder received his Ph.D. in Molecular, Cellular, and Developmental Biology at the University of Michigan.

Ponlkrit Yeesin is a Royal Thai Scholar and a graduate student in the Microbiology Doctoral Training Program at the University of Wisconsin – Madison. His research focuses on mechanisms of DNA replication and recombination in Gram-positive bacteria.

Lyle Simmons is an Associate Professor in the Department of Molecular, Cellular, and Developmental Biology at the University of Michigan. Dr. Simmons received his Ph.D. in Biochemistry and Molecular Biology from Michigan State University and was a postdoctoral fellow in the Department of Biology at MIT.

Jue Wang is a Professor in the Department of Bacteriology at the University of Wisconsin – Madison. Dr. Wang received her Ph.D. in Biochemistry from University of California, San Francisco and was a postdoctoral fellow in the Department of Biology at MIT.

References

- Abdulkarim F, Hughes D. Homologous recombination between the *tuf* genes of *Salmonella typhimurium*. *J Mol Biol.* 1996; 260:506–22. [PubMed: 8759317]
- Anand RP, Tsaponina O, Greenwell PW, Lee CS, Du W, Petes TD, Haber JE. Chromosome rearrangements via template switching between diverged repeated sequences. *Genes Dev.* 2014; 28:2394–406. [PubMed: 25367035]
- Banach-Orlowska M, Fijalkowska IJ, Schaaper RM, Jonczyk P. DNA polymerase II as a fidelity factor in chromosomal DNA synthesis in *Escherichia coli*. *Mol Microbiol.* 2005; 58:61–70. [PubMed: 16164549]
- Barrangou R, Marraffini LA. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell.* 2014; 54:234–44. [PubMed: 24766887]
- Barrick JE, Lenski RE. Genome dynamics during experimental evolution. *Nat Rev Genet.* 2013; 14:827–39. [PubMed: 24166031]
- Becket E, Tse L, Yung M, Cosico A, Miller JH. Polynucleotide phosphorylase plays an important role in the generation of spontaneous mutations in *Escherichia coli*. *J Bacteriol.* 2012; 194:5613–20. [PubMed: 22904280]
- 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; 93:13919–24. [PubMed: 8943036]
- Bhagwat AS, Hao W, Townes JP, Lee H, Tang H, Foster PL. Strand-biased cytosine deamination at the replication fork causes cytosine to thymine mutations in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 2016; 113:2176–81. [PubMed: 26839411]
- Brock, TD. The emergence of bacterial genetics. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press; 1990.
- Buckstein MH, He J, Rubin H. Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*. *J Bacteriol.* 2008; 190:718–26. [PubMed: 17965154]
- Cadet J, Douki T, Ravanat JL. Oxidatively generated damage to the guanine moiety of DNA: mechanistic aspects and formation in cells. *Acc Chem Res.* 2008; 41:1075–83. [PubMed: 18666785]
- Cairns J, Overbaugh J, Miller S. The origin of mutants. *Nature.* 1988; 335:142–5. [PubMed: 3045565]
- Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA. Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell.* 2001; 104:901–12. [PubMed: 11290327]
- Canceill D, Viguera E, Ehrlich SD. Replication slippage of different DNA polymerases is inversely related to their strand displacement efficiency. *J Biol Chem.* 1999; 274:27481–90. [PubMed: 10488082]

- Carver TE Jr, Hochstrasser RA, Millar DP. Proofreading DNA: recognition of aberrant DNA termini by the Klenow fragment of DNA polymerase I. *Proc Natl Acad Sci U S A*. 1994; 91:10670–4. [PubMed: 7938011]
- Cho JE, Kim N, Jinks-Robertson S. Topoisomerase I-dependent deletions initiated by incision at ribonucleotides are biased to the non-transcribed strand of a highly activated reporter. *Nucleic Acids Res*. 2015; 43:9306–13. [PubMed: 26271994]
- Ciccio A, Elledge SJ. The DNA damage response: making it safe to play with knives. *Mol Cell*. 2010; 40:179–204. [PubMed: 20965415]
- Coulondre C, Miller JH. Genetic studies of the lac repressor. III. Additional correlation of mutational sites with specific amino acid residues. *J Mol Biol*. 1977; 117:525–67. [PubMed: 609095]
- Cupples CG, Cabrera M, Cruz C, Miller JH. A set of lacZ mutations in *Escherichia coli* that allow rapid detection of specific frameshift mutations. *Genetics*. 1990; 125:275–80. [PubMed: 2199309]
- Cupples CG, Miller JH. A set of lacZ mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc Natl Acad Sci U S A*. 1989; 86:5345–9. [PubMed: 2501784]
- Curti E, McDonald JP, Mead S, Woodgate R. DNA polymerase switching: effects on spontaneous mutagenesis in *Escherichia coli*. *Mol Microbiol*. 2009; 71:315–31. [PubMed: 19019142]
- Darmon E, Leach DR. Bacterial genome instability. *Microbiol Mol Biol Rev*. 2014; 78:1–39. [PubMed: 24600039]
- Deiham B, Douraghi M, Adibhesami H, Yaseri M, Rahbar M. Screening of mutator phenotype in clinical strains of *Acinetobacter baumannii*. *Microb Pathog*. 2017; 104:175–179. [PubMed: 28111327]
- Demerec M. Induced Mutations and Possible Mechanisms of the Transmission of Heredity in *Escherichia Coli*. *Proc Natl Acad Sci U S A*. 1946; 32:36–46. [PubMed: 16578190]
- Dervyn E, Suski C, Daniel R, Bruand C, Chapuis J, Errington J, Janniere L, Ehrlich SD. Two essential DNA polymerases at the bacterial replication fork. *Science*. 2001; 294:1716–9. [PubMed: 11721055]
- Dettman JR, Sztepanacz JL, Kassen R. The properties of spontaneous mutations in the opportunistic pathogen *Pseudomonas aeruginosa*. *BMC Genomics*. 2016; 17:27. [PubMed: 26732503]
- Dillon MM, Sung W, Lynch M, Cooper VS. The Rate and Molecular Spectrum of Spontaneous Mutations in the GC-Rich Multichromosome Genome of *Burkholderia cenocepacia*. *Genetics*. 2015; 200:935–46. [PubMed: 25971664]
- Dillon MM, Sung W, Sebra R, Lynch M, Cooper VS. Genome-Wide Biases in the Rate and Molecular Spectrum of Spontaneous Mutations in *Vibrio cholerae* and *Vibrio fischeri*. *Mol Biol Evol*. 2017; 34:93–109. [PubMed: 27744412]
- Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci U S A*. 1991; 88:7160–4. [PubMed: 1831267]
- Duigou S, Ehrlich SD, Noirot P, Noirot-Gros MF. Distinctive genetic features exhibited by the Y-family DNA polymerases in *Bacillus subtilis*. *Mol Microbiol*. 2004; 54:439–51. [PubMed: 15469515]
- Dutra BE, Lovett ST. Cis and trans-acting effects on a mutational hotspot involving a replication template switch. *J Mol Biol*. 2006; 356:300–11. [PubMed: 16376936]
- Dutta D, Shatalin K, Epshtein V, Gottesman ME, Nudler E. Linking RNA polymerase backtracking to genome instability in *E. coli*. *Cell*. 2011; 146:533–43. [PubMed: 21854980]
- Feklistov A, Darst SA. Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell*. 2011; 147:1257–69. [PubMed: 22136875]
- Fernandez-Leiro R, Conrad J, Yang JC, Freund SM, Scheres SH, Lamers MH. Self-correcting mismatches during high-fidelity DNA replication. *Nat Struct Mol Biol*. 2017; 24:140–143. [PubMed: 28067916]
- Ferraro P, Franzolin E, Pontarin G, Reichard P, Bianchi V. Quantitation of cellular deoxynucleoside triphosphates. *Nucleic Acids Res*. 2010; 38:e85. [PubMed: 20008099]
- Fijalkowska IJ, Jonczyk P, Tkaczyk MM, Bialoskorska M, Schaaper RM. Unequal fidelity of leading strand and lagging strand DNA replication on the *Escherichia coli* chromosome. *Proc Natl Acad Sci U S A*. 1998; 95:10020–5. [PubMed: 9707593]

- Ford CB, Lin PL, Chase MR, Shah RR, Iartchouk O, Galagan J, Mohaideen N, Ioerger TR, Sacchettini JC, Lipsitch M, Flynn JL, Fortune SM. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat Genet.* 2011; 43:482–6. [PubMed: 21516081]
- Foster KA, Barnes MH, Stephenson RO, Butler MM, Skow DJ, Lamarr WA, Brown NC. DNA polymerase III of *Enterococcus faecalis*: expression and characterization of recombinant enzymes encoded by the *polC* and *dnaE* genes. *Protein Expr Purif.* 2003; 27:90–7. [PubMed: 12509989]
- Foster PL. Methods for determining spontaneous mutation rates. *Methods Enzymol.* 2006; 409:195–213. [PubMed: 16793403]
- Foster PL. Stress-induced mutagenesis in bacteria. *Crit Rev Biochem Mol Biol.* 2007; 42:373–97. [PubMed: 17917873]
- 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. [PubMed: 26460006]
- Frederico LA, Kunkel TA, Shaw BR. A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry.* 1990; 29:2532–7. [PubMed: 2185829]
- Friedberg, EC., Walker, GC., Siede, W., Wood, RD., Schultz, RA., Ellenberger, T. DNA repair and mutagenesis. 2. American Society for Microbiology Press; 2006.
- Fujii S, Akiyama M, Aoki K, Sugaya Y, Higuchi K, Hiraoka M, Miki Y, Saitoh N, Yoshiyama K, Ihara K, Seki M, Ohtsubo E, Maki H. DNA replication errors produced by the replicative apparatus of *Escherichia coli*. *J Mol Biol.* 1999; 289:835–50. [PubMed: 10369765]
- Furukohri A, Goodman MF, Maki H. A dynamic polymerase exchange with *Escherichia coli* DNA polymerase IV replacing DNA polymerase III on the sliding clamp. *J Biol Chem.* 2008; 283:11260–9. [PubMed: 18308729]
- Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a stress response and the regulation of evolvability. *Crit Rev Biochem Mol Biol.* 2007; 42:399–435. [PubMed: 17917874]
- Garibyan L, Huang T, Kim M, Wolff E, Nguyen A, Nguyen T, Diep A, Hu K, Iverson A, Yang H, Miller JH. Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst).* 2003; 2:593–608. [PubMed: 12713816]
- Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science.* 2001; 292:1876–82. [PubMed: 11313499]
- Goodman MF, Creighton S, Bloom LB, Petruska J. Biochemical basis of DNA replication fidelity. *Crit Rev Biochem Mol Biol.* 1993; 28:83–126. [PubMed: 8485987]
- Goodman MF, Woodgate R. Translesion DNA polymerases. *Cold Spring Harb Perspect Biol.* 2013; 5:a010363. [PubMed: 23838442]
- Hall BM, Ma CX, Liang P, Singh KK. Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics.* 2009; 25:1564–5. [PubMed: 19369502]
- Hantke K, Braun V. Functional interaction of the *tonA/tonB* receptor system in *Escherichia coli*. *J Bacteriol.* 1978; 135:190–7. [PubMed: 353030]
- Harris VH, Smith CL, Jonathan Cummins W, Hamilton AL, Adams H, Dickman M, Hornby DP, Williams DM. The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: support for the rare tautomer hypothesis of substitution mutagenesis. *J Mol Biol.* 2003; 326:1389–401. [PubMed: 12595252]
- Hughes D, Andersson DI. Carbon starvation of *Salmonella typhimurium* does not cause a general increase of mutation rates. *J Bacteriol.* 1997; 179:6688–91. [PubMed: 9352917]
- Hunter CA. Sequence-dependent DNA structure. The role of base stacking interactions. *J Mol Biol.* 1993; 230:1025–54. [PubMed: 8478917]
- Jarosz DF, Beuning PJ, Cohen SE, Walker GC. Y-family DNA polymerases in *Escherichia coli*. *Trends Microbiol.* 2007; 15:70–7. [PubMed: 17207624]

- Jarosz DF, Godoy VG, Delaney JC, Essigmann JM, Walker GC. A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. *Nature*. 2006; 439:225–8. [PubMed: 16407906]
- Jee J, Rasouly A, Shamovsky I, Akivis Y, Steinman SR, Mishra B, Nudler E. Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing. *Nature*. 2016; 534:693–6. [PubMed: 27338792]
- Kadner RJ, Heller K, Coulton JW, Braun V. Genetic control of hydroxamate-mediated iron uptake in *Escherichia coli*. *J Bacteriol*. 1980; 143:256–64. [PubMed: 6249788]
- Kanvah S, Joseph J, Schuster GB, Barnett RN, Cleveland CL, Landman U. Oxidation of DNA: damage to nucleobases. *Acc Chem Res*. 2010; 43:280–7. [PubMed: 19938827]
- Kasai H, Chung MH, Jones DS, Inoue H, Ishikawa H, Kamiya H, Ohtsuka E, Nishimura S. 8-Hydroxyguanine, a DNA adduct formed by oxygen radicals: its implication on oxygen radical-involved mutagenesis/carcinogenesis. *J Toxicol Sci*. 1991; 16(Suppl 1):95–105. [PubMed: 1656061]
- Kath JE, Chang S, Scotland MK, Wilbertz JH, Jergic S, Dixon NE, Sutton MD, Loparo JJ. Exchange between *Escherichia coli* polymerases II and III on a processivity clamp. *Nucleic Acids Res*. 2016; 44:1681–90. [PubMed: 26657641]
- Kath JE, Jergic S, Heltzel JM, Jacob DT, Dixon NE, Sutton MD, Walker GC, Loparo JJ. Polymerase exchange on single DNA molecules reveals processivity clamp control of translesion synthesis. *Proc Natl Acad Sci U S A*. 2014; 111:7647–52. [PubMed: 24825884]
- Kim N, Huang SN, Williams JS, Li YC, Clark AB, Cho JE, Kunkel TA, Pommier Y, Jinks-Robertson S. Mutagenic processing of ribonucleotides in DNA by yeast topoisomerase I. *Science*. 2011; 332:1561–4. [PubMed: 21700875]
- Kim TW, Delaney JC, Essigmann JM, Kool ET. Probing the active site tightness of DNA polymerase in subangstrom increments. *Proc Natl Acad Sci U S A*. 2005; 102:15803–8. [PubMed: 16249340]
- Kissling GE, Grogan DW, Drake JW. Confounders of mutation-rate estimators: selection and phenotypic lag in *Thermus thermophilus*. *Mutat Res*. 2013; 749:16–20. [PubMed: 23916418]
- Kool ET. Hydrogen bonding, base stacking, and steric effects in dna replication. *Annu Rev Biophys Biomol Struct*. 2001; 30:1–22. [PubMed: 11340050]
- Kool ET. Active site tightness and substrate fit in DNA replication. *Annu Rev Biochem*. 2002; 71:191–219. [PubMed: 12045095]
- Koskiniemi S, Hughes D, Andersson DI. Effect of translesion DNA polymerases, endonucleases and RpoS on mutation rates in *Salmonella typhimurium*. *Genetics*. 2010; 185:783–95. [PubMed: 20421601]
- Kouchakdjian M, Bodepudi V, Shibutani S, Eisenberg M, Johnson F, Grollman AP, Patel DJ. NMR structural studies of the ionizing radiation adduct 7-hydro-8-oxodeoxyguanosine (8-oxo-7H-dG) opposite deoxyadenosine in a DNA duplex. 8-Oxo-7H-dG(syn).dA(anti) alignment at lesion site. *Biochemistry*. 1991; 30:1403–12. [PubMed: 1991121]
- Krahn JM, Beard WA, Miller H, Grollman AP, Wilson SH. Structure of DNA polymerase beta with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. *Structure*. 2003; 11:121–7. [PubMed: 12517346]
- Kucukyildirim S, Long H, Sung W, Miller SF, Doak TG, Lynch M. The Rate and Spectrum of Spontaneous Mutations in *Mycobacterium smegmatis*, a Bacterium Naturally Devoid of the Postreplicative Mismatch Repair Pathway. *G3 (Bethesda)*. 2016; 6:2157–63. [PubMed: 27194804]
- Lawrence CW, Banerjee SK, Borden A, Leclerc JE. T-T cyclobutane dimers are misinstructive, rather than non-instructive, mutagenic lesions. *Mol Gen Genet*. 1990; 222:166–8. [PubMed: 2233676]
- Le Chatelier E, Becherel OJ, D'alencon E, Canceill D, Ehrlich SD, Fuchs RP, Janniere L. Involvement of DnaE, the second replicative DNA polymerase from *Bacillus subtilis*, in DNA mutagenesis. *J Biol Chem*. 2004; 279:1757–67. [PubMed: 14593098]
- Lea DE, Coulson CA. The distribution of the numbers of mutants in bacterial populations. *J Genet*. 1949; 49:264–85. [PubMed: 24536673]
- Lee H, Doak TG, Popodi E, Foster PL, Tang H. Insertion sequence-caused large-scale rearrangements in the genome of *Escherichia coli*. *Nucleic Acids Res*. 2016; 44:7109–19. [PubMed: 27431326]

- Lee H, Popodi E, Foster PL, Tang H. Detection of structural variants involving repetitive regions in the reference genome. *J Comput Biol.* 2014; 21:219–33. [PubMed: 24552580]
- Lee H, Popodi E, Tang H, Foster PL. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci U S A.* 2012; 109:E2774–83. [PubMed: 22991466]
- Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell.* 2007; 131:1235–47. [PubMed: 18160035]
- Lenhart JS, Schroeder JW, Walsh BW, Simmons LA. DNA repair and genome maintenance in *Bacillus subtilis*. *Microbiol Mol Biol Rev.* 2012; 76:530–64. [PubMed: 22933559]
- Lieb M. Forward and Reverse Mutation in a Histidine-Requiring Strain of *Escherichia Coli*. *Genetics.* 1951; 36:460–77. [PubMed: 17247359]
- Lind PA, Andersson DI. Whole-genome mutational biases in bacteria. *Proc Natl Acad Sci U S A.* 2008; 105:17878–83. [PubMed: 19001264]
- Lobry JR. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol Biol Evol.* 1996; 13:660–5. [PubMed: 8676740]
- Long H, Kucukyildirim S, Sung W, Williams E, Lee H, Ackerman M, Doak TG, Tang H, Lynch M. Background Mutational Features of the Radiation-Resistant Bacterium *Deinococcus radiodurans*. *Mol Biol Evol.* 2015a; 32:2383–92. [PubMed: 25976352]
- Long H, Sung W, Miller SF, Ackerman MS, Doak TG, Lynch M. Mutation rate, spectrum, topology, and context-dependency in the DNA mismatch repair-deficient *Pseudomonas fluorescens* ATCC948. *Genome Biol Evol.* 2015b; 7:262–71.
- Lovett ST. Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences. *Mol Microbiol.* 2004; 52:1243–53. [PubMed: 15165229]
- Lujan SA, Clausen AR, Clark AB, Macalpine HK, Macalpine DM, Malc EP, Mieczkowski PA, Burkholder AB, Fargo DC, Gordenin DA, Kunkel TA. Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res.* 2014; 24:1751–64. [PubMed: 25217194]
- Luria SE, Delbruck M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics.* 1943; 28:491–511. [PubMed: 17247100]
- Lynch M. Evolution of the mutation rate. *Trends Genet.* 2010; 26:345–52. [PubMed: 20594608]
- Maclean RC, Torres-Barcelo C, Moxon R. Evaluating evolutionary models of stress-induced mutagenesis in bacteria. *Nat Rev Genet.* 2013; 14:221–7. [PubMed: 23400102]
- Mcauley-Hecht KE, Leonard GA, Gibson NJ, Thomson JB, Watson WP, Hunter WN, Brown T. Crystal structure of a DNA duplex containing 8-hydroxydeoxyguanine-adenine base pairs. *Biochemistry.* 1994; 33:10266–70. [PubMed: 8068665]
- Merrikh H, Machon C, Grainger WH, Grossman AD, Soultanas P. Co-directional replication-transcription conflicts lead to replication restart. *Nature.* 2011; 470:554–7. [PubMed: 21350489]
- Michaels ML, Miller JH. The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J Bacteriol.* 1992; 174:6321–5. [PubMed: 1328155]
- Miller JH. Spontaneous mutators in bacteria: insights into pathways of mutagenesis and repair. *Annu Rev Microbiol.* 1996; 50:625–43. [PubMed: 8905093]
- Million-Weaver S, Samadpour AN, Moreno-Habel DA, Nugent P, Brittnacher MJ, Weiss E, Hayden HS, Miller SI, Liachko I, Merrikh H. 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. [PubMed: 25713353]
- Mirkin EV, Castro Roa D, Nudler E, Mirkin SM. Transcription regulatory elements are punctuation marks for DNA replication. *Proc Natl Acad Sci U S A.* 2006; 103:7276–81. [PubMed: 16670199]
- Mirkin EV, Mirkin SM. Mechanisms of transcription-replication collisions in bacteria. *Mol Cell Biol.* 2005; 25:888–95. [PubMed: 15657418]
- Neuhard J, Price AR, Schack L, Thomassen E. Two thymidylate synthetases in *Bacillus subtilis*. *Proc Natl Acad Sci U S A.* 1978; 75:1194–8. [PubMed: 418407]

- Newcombe HB. Delayed Phenotypic Expression of Spontaneous Mutations in *Escherichia Coli*. *Genetics*. 1948; 33:447–76. [PubMed: 17247291]
- Newcombe HB, Scott GW. Factors Responsible for the Delayed Appearance of Radiation-Induced Mutants in *Escherichia Coli*. *Genetics*. 1949; 34:475–92. [PubMed: 17247328]
- Nick Mcelhinny SA, Kumar D, Clark AB, Watt DL, Watts BE, Lundstrom EB, Johansson E, Chabes A, Kunkel TA. Genome instability due to ribonucleotide incorporation into DNA. *Nature chemical biology*. 2010a; 6:774–81. [PubMed: 20729855]
- Nick Mcelhinny SA, Watts BE, Kumar D, Watt DL, Lundstrom EB, Burgers PM, Johansson E, Chabes A, Kunkel TA. Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proceedings of the National Academy of Sciences of the United States of America*. 2010b; 107:4949–54. [PubMed: 20194773]
- Paul S, Million-Weaver S, Chattopadhyay S, Sokurenko E, Merrikh H. Accelerated gene evolution through replication-transcription conflicts. *Nature*. 2013; 495:512–5. [PubMed: 23538833]
- Paulsson J, El Karoui M, Lindell M, Hughes D. The processive kinetics of gene conversion in bacteria. *Mol Microbiol*. 2017
- Pavlov YI, Frahm C, Nick Mcelhinny SA, Niimi A, Suzuki M, Kunkel TA. Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. *Curr Biol*. 2006; 16:202–7. [PubMed: 16431373]
- Pennington JM, Rosenberg SM. Spontaneous DNA breakage in single living *Escherichia coli* cells. *Nat Genet*. 2007; 39:797–802. [PubMed: 17529976]
- Petruska J, Goodman MF. Influence of neighboring bases on DNA polymerase insertion and proofreading fidelity. *J Biol Chem*. 1985; 260:7533–9. [PubMed: 3158658]
- Reams AB, Kofoed E, Duleba N, Roth JR. Recombination and annealing pathways compete for substrates in making *rrn* duplications in *Salmonella enterica*. *Genetics*. 2014; 196:119–35. [PubMed: 24214339]
- Reijns MA, Kemp H, Ding J, De Proce SM, Jackson AP, Taylor MS. Lagging-strand replication shapes the mutational landscape of the genome. *Nature*. 2015; 518:502–6. [PubMed: 25624100]
- Reijns MA, Rabe B, Rigby RE, Mill P, Astell KR, Lettice LA, Boyle S, Leitch A, Keighren M, Kilanowski F, Devenney PS, Sexton D, Grimes G, Holt IJ, Hill RE, Taylor MS, Lawson KA, Dorin JR, Jackson AP. Enzymatic removal of ribonucleotides from DNA is essential for Mammalian genome integrity and development. *Cell*. 2012; 149:1008–22. [PubMed: 22579044]
- Reuven NB, Arad G, Maor-Shoshani A, Livneh Z. The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA, and SSB and is specialized for translesion replication. *J Biol Chem*. 1999; 274:31763–6. [PubMed: 10542196]
- Richardson AR, Stojiljkovic I. Mismatch repair and the regulation of phase variation in *Neisseria meningitidis*. *Mol Microbiol*. 2001; 40:645–55. [PubMed: 11359570]
- Ripley LS. Model for the participation of quasi-palindromic DNA sequences in frameshift mutation. *Proc Natl Acad Sci U S A*. 1982; 79:4128–32. [PubMed: 7051004]
- Robinson A, McDonald JP, Caldas VE, Patel M, Wood EA, Punter CM, Ghodke H, Cox MM, Woodgate R, Goodman MF, Van Oijen AM. Regulation of Mutagenic DNA Polymerase V Activation in Space and Time. *PLoS Genet*. 2015; 11:e1005482. [PubMed: 26317348]
- Robledo EA, Yasbin R, Ross C, Pedraza-Reyes M. Stationary phase mutagenesis in *B. subtilis*: a paradigm to study genetic diversity programs in cells under stress. *Crit Rev Biochem Mol Biol*. 2007; 42:327–39. [PubMed: 17917870]
- Rocha EP. The replication-related organization of bacterial genomes. *Microbiology*. 2004; 150:1609–27. [PubMed: 15184548]
- Rock JM, Lang UF, Chase MR, Ford CB, Gerrick ER, Gawande R, Coscolla M, Gagneux S, Fortune SM, Lamers MH. DNA replication fidelity in *Mycobacterium tuberculosis* is mediated by an ancestral prokaryotic proofreader. *Nat Genet*. 2015; 47:677–81. [PubMed: 25894501]
- Rosche WA, Foster PL. Determining mutation rates in bacterial populations. *Methods*. 2000; 20:4–17. [PubMed: 10610800]
- Sanders GM, Dallmann HG, Mchenry CS. Reconstitution of the *B. subtilis* replisome with 13 proteins including two distinct replicases. *Mol Cell*. 2010; 37:273–81. [PubMed: 20122408]

- Sankar TS, Wastuwidyaningtyas BD, Dong Y, Lewis SA, Wang JD. The nature of mutations induced by replication-transcription collisions. *Nature*. 2016; 535:178–81. [PubMed: 27362223]
- Sarkar S, Ma WT, Sandri GH. On fluctuation analysis: a new, simple and efficient method for computing the expected number of mutants. *Genetica*. 1992; 85:173–9. [PubMed: 1624139]
- Schaaper RM. Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J Biol Chem*. 1993; 268:23762–5. [PubMed: 8226906]
- Schaaper RM, Danforth BN, Glickman BW. Mechanisms of spontaneous mutagenesis: an analysis of the spectrum of spontaneous mutation in the *Escherichia coli* lacI gene. *J Mol Biol*. 1986; 189:273–84. [PubMed: 3018259]
- 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 U S A*. 1987; 84:6220–4. [PubMed: 3306672]
- Schaaper RM, Dunn RL. Spontaneous mutation in the *Escherichia coli* lacI gene. *Genetics*. 1991; 129:317–26. [PubMed: 1660424]
- 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; 26:692–7. [PubMed: 26923786]
- Seier T, Padgett DR, Zilberberg G, Sutra VA Jr, Toha N, Lovett ST. Insights into mutagenesis using *Escherichia coli* chromosomal lacZ strains that enable detection of a wide spectrum of mutational events. *Genetics*. 2011; 188:247–62. [PubMed: 21441210]
- Seier T, Zilberberg G, Zeiger DM, Lovett ST. Azidothymidine and other chain terminators are mutagenic for template-switch-generated genetic mutations. *Proc Natl Acad Sci U S A*. 2012; 109:6171–4. [PubMed: 22474374]
- Shewaramani S, Finn TJ, Leahy SC, Kassen R, Rainey PB, Moon CD. Anaerobically Grown *Escherichia coli* Has an Enhanced Mutation Rate and Distinct Mutational Spectra. *PLoS Genet*. 2017; 13:e1006570. [PubMed: 28103245]
- Siguier P, Filee J, Chandler M. Insertion sequences in prokaryotic genomes. *Curr Opin Microbiol*. 2006; 9:526–31. [PubMed: 16935554]
- Simmons LA, Goranov AI, Kobayashi H, Davies BW, Yuan DS, Grossman AD, Walker GC. Comparison of responses to double-strand breaks between *Escherichia coli* and *Bacillus subtilis* reveals different requirements for SOS induction. *J Bacteriol*. 2009; 191:1152–61. [PubMed: 19060143]
- Sinha NK. Specificity and efficiency of editing of mismatches involved in the formation of base-substitution mutations by the 3'→5' exonuclease activity of phage T4 DNA polymerase. *Proc Natl Acad Sci U S A*. 1987; 84:915–9. [PubMed: 3029763]
- Skopek TR, Liber HL, Kaden DA, Thilly WG. Relative sensitivities of forward and reverse mutation assays in *Salmonella typhimurium*. *Proc Natl Acad Sci U S A*. 1978; 75:4465–9. [PubMed: 360221]
- Srivatsan A, Tehranchi A, Macalpine DM, Wang JD. Co-orientation of replication and transcription preserves genome integrity. *PLoS Genet*. 2010; 6:e1000810. [PubMed: 20090829]
- Steinborn G. Uvm mutants of *Escherichia coli* K12 deficient in UV mutagenesis. I. Isolation of uvm mutants and their phenotypical characterization in DNA repair and mutagenesis. *Mol Gen Genet*. 1978; 165:87–93. [PubMed: 362169]
- Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye M. Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday. *Cold Spring Harb Symp Quant Biol*. 1966; 31:77–84. [PubMed: 5237214]
- Sung HM, Yeaman G, Ross CA, Yasbin RE. Roles of YqjH and YqjW, homologs of the *Escherichia coli* UmuC/DinB or Y superfamily of DNA polymerases, in stationary-phase mutagenesis and UV-induced mutagenesis of *Bacillus subtilis*. *J Bacteriol*. 2003; 185:2153–60. [PubMed: 12644484]
- Sung W, Ackerman MS, Gout JF, Miller SF, Williams E, Foster PL, Lynch M. Asymmetric Context-Dependent Mutation Patterns Revealed through Mutation-Accumulation Experiments. *Mol Biol Evol*. 2015; 32:1672–83. [PubMed: 25750180]

- 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; 109:18488–92. [PubMed: 23077252]
- Sutton MD, Duzen JM. Specific amino acid residues in the beta sliding clamp establish a DNA polymerase usage hierarchy in *Escherichia coli*. *DNA Repair (Amst)*. 2006; 5:312–23. [PubMed: 16338175]
- Tang M, Pham P, Shen X, Taylor JS, O'donnell M, Woodgate R, Goodman MF. Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature*. 2000; 404:1014–8. [PubMed: 10801133]
- Tang M, Shen X, Frank EG, O'donnell M, Woodgate R, Goodman MF. UmuD'(2)C is an error-prone DNA polymerase, *Escherichia coli* pol V. *Proc Natl Acad Sci U S A*. 1999; 96:8919–24. [PubMed: 10430871]
- Titok M, Suski C, Dalmais B, Ehrlich SD, Janniere L. The replicative polymerases PolC and DnaE are required for theta replication of the *Bacillus subtilis* plasmid pBS72. *Microbiology*. 2006; 152:1471–8. [PubMed: 16622063]
- Tomasetti C, Li L, Vogelstein B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science*. 2017; 355:1330–1334. [PubMed: 28336671]
- Topal MD, Fresco JR. Complementary base pairing and the origin of substitution mutations. *Nature*. 1976; 263:285–9. [PubMed: 958482]
- Traut TW. Physiological concentrations of purines and pyrimidines. *Molecular and cellular biochemistry*. 1994; 140:1–22. [PubMed: 7877593]
- Tsaponina O, Haber JE. Frequent Interchromosomal Template Switches during Gene Conversion in *S. cerevisiae*. *Mol Cell*. 2014; 55:615–25. [PubMed: 25066232]
- Vaisman A, McDonald JP, Huston D, Kuban W, Liu L, Van Houten B, Woodgate R. Removal of misincorporated ribonucleotides from prokaryotic genomes: an unexpected role for nucleotide excision repair. *PLoS Genet*. 2013; 9:e1003878. [PubMed: 24244177]
- Vaisman A, McDonald JP, Noll S, Huston D, Loeb G, Goodman MF, Woodgate R. Investigating the mechanisms of ribonucleotide excision repair in *Escherichia coli*. *Mutat Res Fundam Mol Mech Mutagen*. 2014; 761:21–33.
- Vilette D, Ehrlich SD, Michel B. Transcription-induced deletions in *Escherichia coli* plasmids. *Mol Microbiol*. 1995; 17:493–504. [PubMed: 8559068]
- Vilette D, Ehrlich SD, Michel B. Transcription-induced deletions in plasmid vectors: M13 DNA replication as a source of instability. *Mol Gen Genet*. 1996; 252:398–403. [PubMed: 8879240]
- Vilette D, Uezst M, Ehrlich SD, Michel B. DNA transcription and repressor binding affect deletion formation in *Escherichia coli* plasmids. *EMBO J*. 1992; 11:3629–34. [PubMed: 1396563]
- Viswanathan M, Lacirignola JJ, Hurley RL, Lovett ST. A novel mutational hotspot in a natural quasipalindrome in *Escherichia coli*. *J Mol Biol*. 2000; 302:553–64. [PubMed: 10986118]
- Wagner J, Gruz P, Kim SR, Yamada M, Matsui K, Fuchs RP, Nohmi T. The *dinB* gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell*. 1999; 4:281–6. [PubMed: 10488344]
- Wang G, Wilson TJ, Jiang Q, Taylor DE. Spontaneous mutations that confer antibiotic resistance in *Helicobacter pylori*. *Antimicrob Agents Chemother*. 2001; 45:727–33. [PubMed: 11181351]
- Wang W, Hellinga HW, Beese LS. Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc Natl Acad Sci U S A*. 2011; 108:17644–8. [PubMed: 22006298]
- Wang W, Wu EY, Hellinga HW, Beese LS. Structural factors that determine selectivity of a high fidelity DNA polymerase for deoxy-, dideoxy-, and ribonucleotides. *J Biol Chem*. 2012; 287:28215–26. [PubMed: 22648417]
- Watson JD, Crick FH. Genetical implications of the structure of deoxyribonucleic acid. *Nature*. 1953; 171:964–7. [PubMed: 13063483]
- Wielgoss S, Barrick JE, Tenaillon O, Cruveiller S, Chane-Woon-Ming B, Medigue C, Lenski RE, Schneider D. Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With *Escherichia coli*. *G3 (Bethesda)*. 2011; 1:183–186. [PubMed: 22207905]
- Williams JS, Smith DJ, Marjavaara L, Lujan SA, Chabes A, Kunkel TA. Topoisomerase I-mediated removal of ribonucleotides from nascent leading-strand DNA. *Molecular Cell*. 2013; 49:1010–5. [PubMed: 23375499]

- Witkin EM. Nuclear segregation and the delayed appearance of induced mutants in *Escherichia coli*. Cold Spring Harb Symp Quant Biol. 1951; 16:357–72. [PubMed: 14942750]
- Yamada M, Shimizu M, Katafuchi A, Gruz P, Fujii S, Usui Y, Fuchs RP, Nohmi T. *Escherichia coli* DNA polymerase III is responsible for the high level of spontaneous mutations in *mutT* strains. Mol Microbiol. 2012; 86:1364–75. [PubMed: 23043439]
- Yao NY, Schroeder JW, Yurieva O, Simmons LA, O'donnell ME. Cost of rNTP/dNTP pool imbalance at the replication fork. Proc Natl Acad Sci U S A. 2013; 110:12942–7. [PubMed: 23882084]
- Yoshida H, Bogaki M, Nakamura M, Nakamura S. Quinolone resistance-determining region in the DNA gyrase *gyrA* gene of *Escherichia coli*. Antimicrob Agents Chemother. 1990; 34:1271–2. [PubMed: 2168148]
- Yoshiyama K, Higuchi K, Matsumura H, Maki H. Directionality of DNA replication fork movement strongly affects the generation of spontaneous mutations in *Escherichia coli*. J Mol Biol. 2001; 307:1195–206. [PubMed: 11292335]
- Yuan B, Cao H, Jiang Y, Hong H, Wang Y. Efficient and accurate bypass of N2-(1-carboxyethyl)-2'-deoxyguanosine by DinB DNA polymerase in vitro and in vivo. Proc Natl Acad Sci U S A. 2008; 105:8679–84. [PubMed: 18562283]

Figure 1.

Fluctuation tests, mutation reporters, mutation accumulation experiments, and maximum-depth sequencing. (A) An illustration of the fluctuation test. Mutants (black circles) arise during exponential growth of independent cultures. If selection is lethal to non-mutants, only mutants that arose during culture growth will form colonies (black colonies). However, if selection is for growth, slow growth during selection may give rise to additional mutants that form colonies during selection (red colonies). (B) The rationale behind reverse and forward mutation reporters is illustrated. For reverse mutation reporters, a reversion assay may be performed (top), in which a functional product allows survival and growth under selection. In the other class of reverse mutation reporter (middle) an active gene product is produced before and after mutagenesis, but the mutant gene product remains functional during

selection, in contrast to non-mutant gene products, which do not function during selection. Forward mutation reporters (bottom) have a larger target size, in that any mutation eliminating expression or gene product activity results in growth or survival under selection. (C) A table of data from Luria and Delbruck's 1943 paper (Luria and Delbruck, 1943) with examples of jackpot cultures indicated by rows shaded in gray. (D) An example mutation accumulation line. Mutations accumulate in the line's genome (shown below the example mutation accumulation line) as it is passaged. After the desired number of generations, genomic DNA from each line is purified and sequenced to identify mutations that arose during the experiment. (E) The schematic workflow of maximum depth sequencing. A unique molecular identifier (UMI) is appended to the 3' end of the original molecule of the region of interest (ROI), followed by linear amplification of the ROI molecules with their associated UMIs. The consensus sequence for each ROI/UMI pool is determined and the proportion of ROI/UMI pools with mutations at a given position is determined in order to calculate mutation rate at the position. A color version of this figure is available online.

Figure 2.

Common sources of spontaneous mutagenesis in bacteria. Cellular processes that promote spontaneous mutagenesis are listed at the top of the schematic with arrows indicating the types of mutations with which they are associated.

Figure 3.

Nucleobase modifications involved in mutagenesis. (A) Base pairing interactions between rare tautomeric forms of the nucleobases in DNA are shown for comparison to canonical base pairs. (B) Interactions between deoxyadenosine and *syn*-8-oxo-dG. (C) Cytosine deamination leads to uracil, causing C·G→T·A transitions, and adenine deamination leads to hypoxanthine, yielding A·T→G·C transitions (see section: Replication-transcription conflict). Structures were illustrated using ChemDraw software.

Table 1Key definitions (Friedberg *et al.*, 2006)

Term	Definition
Mutation	A heritable change in an organisms's genome sequence
Spontaneous mutation	A mutation that occurs in the absence of exogenously applied stress or selective pressure
Mutant	An organism with one or more mutations in its genome
Genotype	The genetic information that an organism encodes in its genome
Phenotype	The ensemble of observable characteristics of an organism
Mutagenesis	The process by which mutations are produced
Mutation reporter	A gene or set of genes in which mutations may yield mutants with a selectable phenotype
Transition	A type of base pair substitution: Purine mutating to purine ($G \rightarrow A$, $A \rightarrow G$) or pyrimidine mutating to pyrimidine ($C \rightarrow T$, $T \rightarrow C$)
Transversion	A type of base pair substitution: Purine mutating to pyrimidine ($G \rightarrow T$, $G \rightarrow C$, $A \rightarrow T$, $A \rightarrow C$) or pyrimidine mutating to purine ($T \rightarrow A$, $T \rightarrow G$, $C \rightarrow A$, $C \rightarrow G$)
Insertion/deletion	Addition or removal of nucleotides from a genome
Frameshift	Insertion/deletion of $(3n \pm 1)$ nucleotides within the open reading frame of a gene, where n is an integer.
Target size	The number of unique mutations that are observable in a mutation assay. A simplified example: If one type of mutation at three distinct sites in a reporter gene can generate selectable mutants, the target size is three. If three types of mutations at three distinct sites in the gene can generate selectable mutants, the target size is nine.

Table 2

Mutation rate estimates from various bacteria, methods, and environments

Bacterium	Method ^a	Condition ^b	Mutation type	Mutation rate per nucleotide per generation ($\times 10^{-10}$) ^c	Reference
<i>Bacillus subtilis</i>	MA/WGS	LBA	BPS	3.35	(Sung <i>et al.</i> , 2015, Schroeder <i>et al.</i> , 2016)
<i>Bacillus subtilis</i>	MA/WGS	LBA	Indel	1.20	(Sung <i>et al.</i> , 2015, Schroeder <i>et al.</i> , 2016)
<i>Bacillus subtilis</i>	RifR (<i>rpoB</i>)	LB	BPS	0.34	(Lynch, 2010, Schroeder <i>et al.</i> , 2016)
<i>Bacillus subtilis</i>	TmpR (<i>thyP3</i>)	Co-directional replication-transcription	Promoter BPS	85.4×10^{-8} per locus per generation	(Sankar <i>et al.</i> , 2016)
<i>Bacillus subtilis</i>	TmpR (<i>thyP3</i>)	Head-on replication-transcription	Promoter BPS	2.68×10^{-8} per locus per generation	(Sankar <i>et al.</i> , 2016)
<i>Bacillus subtilis</i>	TmpR (<i>thyP3</i>)	Co-directional replication-transcription	Coding sequence BPS	0.26 ^d	(Sankar <i>et al.</i> , 2016)
<i>Bacillus subtilis</i>	TmpR (<i>thyP3</i>)	Head-on replication-transcription	Coding sequence BPS	0.45 ^d	(Sankar <i>et al.</i> , 2016)
<i>Burkholderia cenocepacia</i>	MA/WGS	TSA	BPS	1.33	(Dillon <i>et al.</i> , 2015)
<i>Burkholderia cenocepacia</i>	MA/WGS	TSA	Indel	0.17	(Dillon <i>et al.</i> , 2015)
<i>Deinococcus radiodurans</i>	MA/WGS	NAG	BPS	4.99	(Long <i>et al.</i> , 2015a)
<i>Deinococcus radiodurans</i>	MA/WGS	NAG	Indel	0.22	(Long <i>et al.</i> , 2015a)
<i>Escherichia coli</i>	MA/WGS	LBA	BPS	2.00	(Lee <i>et al.</i> , 2012)
<i>Escherichia coli</i>	MA/WGS	LBA	Indel	0.37	(Lee <i>et al.</i> , 2012)
<i>Escherichia coli</i>	MA/WGS	Minimal medium agar	BPS	1.94	(Foster <i>et al.</i> , 2015)
<i>Escherichia coli</i>	Long-term evolution	DMM	BPS	0.89	(Wielgoss <i>et al.</i> , 2011)
<i>Escherichia coli</i>	RifR (<i>rpoB</i>)	LB	BPS	0.33	(Lee <i>et al.</i> , 2012)
<i>Escherichia coli</i>	NalR (<i>gyrA</i>)	LB	BPS	0.21	(Lee <i>et al.</i> , 2012)
<i>Escherichia coli</i>	Histidine auxotrophy (<i>hisGDCBHAFE</i>)	Minimal medium	BPS	5.06	(Lieb, 1951, Drake, 1991)
<i>Escherichia coli</i>	Lac reversion	NR	BPS	5.5	Mean of two estimates in (Drake, 1991)

Bacterium	Method ^a	Condition ^b	Mutation type	Mutation rate per nucleotide per generation ($\times 10^{-10}$) ^c	Reference
<i>Helicobacter pylori</i>	RifR (<i>rpoB</i>), CipR (<i>gyrA</i>), ClaR (23S rRNA)	BHI-YE	BPS	27.2	(Lynch, 2010)
<i>Mesoplasma florum</i>	MA/WGS	Mycoplasma medium agar	BPS	97.8	(Sung <i>et al.</i> , 2012)
<i>Mesoplasma florum</i>	MA/WGS	Mycoplasma medium agar	Indel	23.1	(Sung <i>et al.</i> , 2012)
<i>Mycobacterium smegmatis</i>	MA/WGS	7H10 agar with glycerol and OADC	BPS	5.27	(Kucukyildirim <i>et al.</i> , 2016)
<i>Mycobacterium smegmatis</i>	MA/WGS	7H10 agar with glycerol and OADC	Indel	1.24	(Kucukyildirim <i>et al.</i> , 2016)
<i>Mycobacterium tuberculosis</i>	MA/WGS	7H10 agar	BPS	2.54	(Rock <i>et al.</i> , 2015)
<i>Mycobacterium tuberculosis</i>	MA/WGS	7H10 agar	Indel	1.98	(Rock <i>et al.</i> , 2015)
<i>Mycobacterium tuberculosis</i>	RifR (<i>rpoB</i>)	7H9 with OADC	BPS	2.1	(Ford <i>et al.</i> , 2011)
<i>Pseudomonas aeruginosa</i>	MA/WGS	LBA	BPS	0.79	(Dettman <i>et al.</i> , 2016)
<i>Pseudomonas aeruginosa</i>	MA/WGS	LBA	Indel	0.14	(Dettman <i>et al.</i> , 2016)
<i>Salmonella enterica</i>	RifR (<i>rpoB</i>)	LB	BPS	1.30	(Lynch, 2010)
<i>Salmonella enterica</i>	Lac reversion	LB	BPS	2.10	(Lynch, 2010)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	LBA, anaerobic	BPS	0.34	(Hughes and Andersson, 1997, Lynch, 2010)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	M9 agar with glucose, anaerobic	BPS	1.0	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	M9 agar with galactose, anaerobic	BPS	0.38	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	LBA, aerobic	BPS	1.0	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	M9 agar with glucose, aerobic	BPS	1.1	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	M9 agar with galactose, aerobic	BPS	0.57	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	M9 agar with galactose, anaerobic	BPS	0.38	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	LBA, aerobic	BPS	1.0	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	M9 agar with glucose, aerobic	BPS	1.1	(Hughes and Andersson, 1997)
<i>Salmonella typhimurium</i>	RifR (<i>rpoB</i>)	M9 agar with galactose, aerobic	BPS	0.57	(Hughes and Andersson, 1997)

^aMA/WGS = mutation accumulation and whole genome resequencing, RifR = rifampin resistance, TmpR = trimethoprim resistance, NalR = nalidixic acid resistance, CycR = cycloserine resistance, CipR = ciprofloxacin resistance, ClaR = clarithromycin resistance.

^bLBA = LB agar, TSA = tryptic soy agar, NAG = nutrient agar with glucose, DMM = Davis minimal medium, BHI-YE = brain heart infusion with yeast extract, NR = not reported

^cBPS rate per site per generation estimates from mutation reporters were calculated as described in (Lynch, 2010) unless otherwise noted.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

The BPS mutation rate per site per generation, μ_{BS} , was calculated using $\mu_{BS} = f_T \times f_L \times f_D$ from (Lynch, 2010), where μ_L is the per locus mutation rate, f_T , f_{BS} , and f_L are all 1, and $L = 840$. 220 unique BPSs at 173 sites were detected in the *thyP3* coding sequence by (Sankar *et al.*, 2016). Therefore, $f_D = 220/(3 \times 173) = 0.475$. Dividing by three in calculation of f_D accounts for the three possible nucleotides that could have substituted for the native nucleotide (Lynch, 2010).