Mutation—The Engine of Evolution: Studying Mutation and Its Role in the Evolution of Bacteria

Ruth Hershberg

Rachel & Menachem Mendelovitch Evolutionary Processes of Mutation & Natural Selection Research Laboratory, Department of Genetics and Developmental Biology, The Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa 31096, Israel

Correspondence: ruthersh@tx.technion.ac.il

Mutation is the engine of evolution in that it generates the genetic variation on which the evolutionary process depends. To understand the evolutionary process we must therefore characterize the rates and patterns of mutation. Starting with the seminal Luria and Delbruck fluctuation experiments in 1943, studies utilizing a variety of approaches have revealed much about mutation rates and patterns and about how these may vary between different bacterial strains and species along the chromosome and between different growth conditions. This work provides a critical overview of the results and conclusions drawn from these studies, of the debate surrounding some of these conclusions, and of the challenges faced when studying mutation and its role in bacterial evolution.

Genetic variation is a prerequisite to evolutionary change. In the absence of such variation, no subsequent change can be achieved. Genetic variation is ultimately all generated by mutation. It is therefore clear that mutation is a major evolutionary force that must be studied and understood to understand evolution. Yet, often mutation is set aside and thought of as a random generator of variation that follows very simple and predictable rules.

Many reviews of mutation deal with the molecular mechanisms of mutation and repair (e.g., Modrich 1991; Smith 1992; Lieber 2010). This work, in contrast, relates to mutation as an evolutionary force, focusing on bacteria. We will show that mutation is extremely difficult to study, that we do not know nearly enough about mutation and that recently sev-

eral of our decades-old assumptions were shown to be mistaken, in light of newly available data.

MUTATIONS VERSUS SUBSTITUTIONS

It is important to note that, in this article, we will only be considering *de novo* point mutations. We will not discuss large insertions or deletions or horizontal gene transfer events. To proceed, we must define some terms.

For the purpose of this article, we will define “DNA mutations” as single nucleotide changes in the DNA sequence of an individual organism. These will be the end result of the molecular DNA change, and of the fact that this DNA change was not repaired by the cellular repair systems. Once a mutation occurs and is present within an individual, it will either increase in

Editor: Howard Ochman

Additional Perspectives on Microbial Evolution available at www.cshperspectives.org

Copyright © 2015 Cold Spring Harbor Laboratory Press; all rights reserved

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a018077

R. Hershberg

frequency within the population, or will vanish from the population. The ultimate fate of mutations depends on a combination of natural selection and stochastic forces, such as genetic drift.

We will define “DNA substitutions” as those mutations that we can directly observe when we consider DNA sequence data. The substitutions we observe may reflect the mutations that have occurred for better or worse, depending on how natural selection has affected them. For example, if when comparing sequences we observe that a certain substitution type (e.g., C to T transitions) occurs more frequently within our data, this could either mean that this mutation type occurs more frequently, or that natural selection tends to favor this mutation type

once it occurs (Fig. 1). Note that our definition of substitutions differs somewhat from that of others that sometimes define substitutions as either mutations that have fixed (e.g., Gillespie 1998) or a specific class of base-change mutation (e.g., Graur and Li 2000).

We will define a phenotypic, or marker mutation, as a phenotypic change occurring in an individual. For example, an antibiotic resistance phenotypic mutation causes an individual bacterium to become resistant to an antibiotic. Similarly, we can define a phenotypic, or marker substitution, as a phenotypic change we are able to observe, for example, an increase in the frequency of resistant mutants within a bacterial population. Such an increase can occur because the resistance mutation occurs more frequently

A Normal levels of selection

B Relaxed selection

Figure 1. Different types of mutations (represented by differently colored arrows) occur at different frequencies (represented by arrow thickness). Selection acts as a sieve and allows only a subset of these mutations to persist and become the differences we see between genomes. Such differences are referred to as substitutions. Various types of mutations have different fitness effect distributions, and will be differently affected by selection. (A) Under normal levels of selection, selection will introduce its own biases into patterns of variation. Thus, biases in the patterns of observable substitutions between genomes are likely not to reflect mutational biases. (B) When selection is extremely relaxed, it is expected to affect patterns of variation to a much lesser extent, because it will affect only mutations with very high-fitness effects. Under such conditions, observed substitutions between genomes approximate a random sample of the mutations that have occurred. Because of this, when selection is relaxed, biases in the patterns of substitutions observed between genomes will better approximate mutational biases.

or because of natural selection favoring the resistant mutant.

Often, mutation is studied by assuming that certain types of DNA mutations (e.g., synonymous mutations) or certain marker mutations (e.g., antibiotic resistance mutations when a bacterium is not exposed to antibiotics) evolve entirely neutrally. If there is absolutely no selection acting on an observed class of substitutions, their patterns and rates will indeed be a derivative of the patterns and rates of mutation. However, as we will see later in this article, it is rare to find cases in which DNA or marker mutations are totally unaffected by selection. Determining mutational patterns and rates is therefore a tricky business that requires one to find creative ways to eliminate or minimize the effects of natural selection on observed substitutions.

LURIA AND DELBRUCK—ESTIMATING MUTATION RATES CAN BE A NOISY BUSINESS

In their seminal 1943 “fluctuation experiments,” Luria and Delbruck showed that even if mutational markers truly did evolve neutrally, estimates of mutation rates based on such markers would be extremely noisy (Luria and Delbruck 1943). Luria and Delbruck were attempting to understand the following phenomenon. When a pure bacterial culture is exposed to a bacteriophage, the culture will disappear because of destruction of cells sensitive to the virus. After further incubation, the culture will often become turbid again because of growth of a variant that is resistant to the phage. Once the variant is isolated, it often remains resistant even if it is cultured for many generations in the absence of any phage. At the time Luria and Delbruck were considering this problem, very little was known about the molecular mechanisms of mutation. Yet, they already understood that such a phenomenon could either occur because of resistance mutations occurring before the viral challenge, or because a certain proportion of sensitive cells somehow acquire resistance once they are exposed to phage (Luria and Delbruck 1943).

Luria and Delbruck modeled the variance expected in the number of resistant mutants under both these scenarios (Luria and Delbruck 1943). Their models showed that a much higher variance would be expected if the emergence of resistance were caused by mutations occurring before exposure to viruses. If mutation is a Poisson process and if mutations occur after and in response to viral exposure, one would expect the number of resistant mutants following exposure to be distributed around a certain mean, with the variance equal to the mean (a known characteristic of the Poisson distribution). If, however, mutations occur before exposure, they can occur in any generation of growth. Mutations occurring in earlier generations will rise to higher frequencies by the end of an experiment, compared with mutations occurring in later generations. Therefore, the number of resistant mutants at the end of an experiment will depend not only on the number of mutations that have occurred, but also on when these mutations occurred. This should greatly enhance the variance in the numbers of resistant mutants observed between different experiments. Indeed, Luria and Delbruck then went on to show that in different experiments they saw a variance that was much higher than the mean number of resistant mutants. This provided the first ever demonstration that mutations occurred before selection for their outcome (Luria and Delbruck 1943).

In addition to showing for the first time that mutation precedes selection, the Luria and Delbruck study also shed light on the great variance in substitution rates one can expect to observe when considering phenotypic markers (Luria and Delbruck 1943). First, as mentioned above, they showed that the variance in marker substitution frequency was expected to be much higher than the mean marker substitution frequency. Second, Luria and Delbruck found that the mean substitution frequency they estimated by simply averaging substitution frequencies across different experiments was much higher than the substitution frequency estimated by assuming a Poisson distribution and considering the number of experiments in which no resistance substitutions were observed. This

R. Hershberg

exemplifies the strong effect mutations occurring early on in the experiment can have on calculated average mutation frequencies. One or a few experiments in which a relatively high number of mutations occurred early on, may greatly skew the estimated average frequency of mutations upward. Thus, estimates of mutation frequencies and rates obtained by using marker substitutions can often be very noisy (Luria and Delbruck 1943). Fortunately, we can now, in many cases, move away from using markers and rather use whole-genome sequencing to study mutation.

METHODS FOR ELIMINATING THE EFFECTS OF NATURAL SELECTION WHEN STUDYING MUTATION

To be able to study different parameters of the mutational process, we must be able to disentangle mutation from the effects of natural selection. The easiest way of accomplishing this is by focusing on scenarios in which selection is expected to have less of an effect on patterns of substitution (Fig. 1). A number of studies have used pseudogenes to study mutational biases (e.g., see Andersson and Andersson 1999; Nachman and Crowell 2000). Such studies assume that sequence variation within pseudogenes is unaffected by selection, because pseudogenes are no longer under selection to maintain function. Therefore, it is assumed that patterns of sequence variation within pseudogenes will be determined solely by mutation. Although useful, this approach has limitations. For one, although pseudogenes should not be under selection stemming from protein function, they may be under selection owing to genome-wide factors. For example, if there is selection to maintain a certain genomic nucleotide content (Hershberg and Petrov 2010; Hildebrand et al. 2010), it might affect pseudogenes as strongly as it does other sequences. Second, for most microbial genomes, we can only identify a very small number of pseudogenes, because bacterial pseudogenes tend to be lost very quickly (Kuo and Ochman 2010).

A second approach is to focus on evolutionary scenarios in which the efficiency of selection

is reduced across the entire genome (Fig. 1). Such genome-wide relaxations of selection can be the result of either close relatedness (Akashi 1995; Messer 2009) and/or small effective population sizes (N_e) (Lynch 2007). Bacterial lineages exist for which genetic variation between members of the lineage has naturally been only weakly affected by selection, probably caused by a combination of close relatedness and small N_e (Hershberg et al. 2008; Holt et al. 2008; Hershberg and Petrov 2010; Lieberman et al. 2011). Large quantities of genomic data from many members of several such lineages are publicly available. Patterns of sequence variation between members of bacterial lineages evolving under relaxed selection can be used to characterize mutational patterns (Fig. 1).

The efficiency of selection can also be artificially reduced in the laboratory through repeated single-cell bottlenecks of growing bacterial populations, which severely reduces N_e . Such experiments are called mutation accumulation (MA) experiments (Elena and Lenski 2003; Lind and Andersson 2008; Brockhurst et al. 2010). It is now possible to follow up MA experiments with whole-genome sequencing of the ancestor strain and its resulting progeny, thus allowing for the genome-wide identification of the MA mutations. The number of generations a bacterial population underwent during an MA experiment can be easily estimated. MA experiments therefore make it possible to estimate not only the relative rates with which different classes of mutations occur, but also the overall, absolute mutation rates. This is a clear advantage of MA experiments over approaches that rely on sequencing data from naturally evolving bacteria, which cannot be used to estimate absolute mutation rates. At the same time, MA experiments are much more labor intensive. It is also important to note that the mutation rates and patterns estimated through MA experiments may be influenced by the conditions under which these experiments are performed. This is a particular concern if mutation rates and patterns change under different growth conditions. For example, the stress-induced mutagenesis theory suggests that mutation rates could be much higher during

stationary phase (reviewed in Galhardo et al. 2007, and discussed in depth later in this review).

ABSOLUTE/OVERALL RATES OF MUTATION

One of the key parameters of the mutational process is the absolute rate with which mutations happen, on average, across all types of mutations and along the entire genome. In 1991, based on data collected by using a combination of fluctuation and MA experiments, and quantifying mutation rates based on the frequency of marker substitutions, John Drake coined “Drake’s rule” (Drake 1991). According to this rule, per nucleotide point mutation rates inversely correlate with genome size in microbes. As a result, genome-wide mutation rates are an approximate constant of ~ 0.003 -point mutations per genome per generation (Drake 1991). These results were based on mutation rates of only seven microbes, but later results from many additional microbes provided further support for Drake’s rule, particularly in prokaryotes and in double-stranded DNA (dsDNA) viruses (Lynch 2010). Drake argued that such a fine-tuned mutation rate must be an evolved trait (Drake 1991).

It is generally accepted that natural selection favors the lowering of mutation rates, as mutations are mostly deleterious (Kimura 1967; Drake 1991; Dawson 1998; Lynch 2010). Drake and others postulated that reducing mutation rates comes at a certain physiological cost (Kimura 1967; Drake 1991; Dawson 1998). Drake suggested that mutation rates reached equilibrium when the benefit of further lowering mutation rates matched the physiological cost of so doing. In other words, according to Drake, natural selection drives both the reduction in mutation rates, as well as the ultimate tapering off of this reduction. In contrast, Michael Lynch suggested an alternative model under which the lower limit on mutation rates is not set by natural selection on physiological cost, but rather by genetic drift (Lynch 2010). As per-base mutation rates become lower, selection to further reduce mutation rates becomes weaker, until a point is reached in which selection is no

longer strong enough to counteract the action of genetic drift (Lynch 2010). Supporting this model, Lynch was able to show that per-base mutation rates inversely correlated with effective population sizes (N_e) in both prokaryotes and eukaryotes (Lynch 2010; Sung et al. 2012). Because N_e is inversely related to the power of drift, it can therefore be said that mutation rates become higher as the power of drift relative to selection becomes stronger, congruent with Lynch’s model.

Lynch later refined his “drift-barrier” model by showing that the regression of the mutation rates versus N_e is elevated for prokaryotes compared with eukaryotes (Sung et al. 2012). This finding suggested that, for a given N_e , selection is less effective at reducing mutation rates in prokaryotes. To explain this phenomenon, Lynch suggested that the magnitude of selection to reduce mutation rates is not just a function of the per-base mutation rate, but rather also of the genome-wide deleterious mutation potential of the genome (Sung et al. 2012). Prokaryotes that tend to have less coding sequences in total, provide a smaller target for the origin of deleterious mutations than eukaryotic genomes. Under this refined model, the strength of selection to reduce per nucleotide mutation rates will scale positively with what Lynch defined as the effective genome size, which he approximated as the sum of coding DNA within a genome. Fitting with this, Lynch observed that the effective genome-wide mutation rate, calculated as the per-site mutation rate multiplied by the effective genome size, inversely correlated with N_e , in a way that did not depend on whether an organism is a prokaryote or a eukaryote (Sung et al. 2012).

Under both Drake’s and Lynch’s models, the cost of deleterious mutations is what drives mutation rates down (Drake 1991; Lynch 2010; Sung et al. 2012). Therefore, under both models, an increase in the average cost of mutations would lead to a decrease in mutation rates. To examine this, Drake examined mutation rates of thermophiles and compared them to those of mesophiles (Drake 2009). The rationale was that many mutations that are tolerated at the standard growth temperature are highly harmful

R. Hershberg

when temperatures are higher. Thus, more mutations will have a fitness cost in thermophiles than in mesophiles, which should lead to lower mutation rates within thermophiles (Drake 2009). By again using data derived by use of marker substitutions, Drake then showed that mutation rates in two different thermophilic microbes were indeed much lower than in mesophilic microbes and phages (Drake 2009). This seems to support the model under which selection favors lowering of mutation rates, because of the deleterious effects of mutations.

Recently, many studies have been conducted in which MA lines from various microbes were fully sequenced to determine mutation rates (e.g., Lind and Andersson 2008; Lee et al. 2012; Sung et al. 2012). As discussed above, measures of mutation rates from whole-genome sequencing are expected to be more precise than those measured via the use of phenotypic markers. These recent studies have shown that although the Drake rule seems to generally apply in prokaryotes and dsDNA phages, the range of per genome mutation rates appears to be higher than originally postulated by Drake. For example, Lee et al. (2012) estimated mutation rates for a wild-type *Escherichia coli* laboratory strain, based on whole-genome sequencing of 59 MA lines. Based on these data, they estimated a mutation rate of ~ 0.001 mutations per genome per generation (lower than the 0.003 constant suggested by Drake) (Lee et al. 2012). Sung et al. (2012) sequenced MA lines of one of the smallest culturable bacteria, *Mesoplasma florum*, and found a genome-wide mutation rate of 0.008.

MUTATIONAL BIASES

Various types of mutations may occur at different rates. Such consistent variation in the rates of different categories of mutations means that the mutational process in itself, even in the absence of any natural selection, may introduce biases into patterns of genetic variation. Characterizing these biases is important for understanding which biases in patterns of genetic variation are selected and thus functionally important, and which may just be introduced by the mutational process.

Adenine-Thymine (AT) Bias of Mutation and Bacterial Nucleotide Content Variation

Bacterial nucleotide content is extremely variable. Some bacteria have guanine-cytosine (GC) content $< 25\%$, whereas the GC content of other bacteria can reach 75%. This variation was for a very long time considered to be entirely neutral, and the result of extreme variation in mutational biases between different bacteria (Sueoka 1962; Muto and Osawa 1987). It was thought that GC-rich bacteria were simply ones in which AT to GC mutations occurred more frequently than GC to AT mutations. The opposite pattern of mutation was thought to occur in AT-rich bacteria (Sueoka 1962; Muto and Osawa 1987). However, it was more recently shown, using data from bacteria evolving under varying degrees of relaxed selection, that mutation is universally AT biased across both AT-rich and GC-rich bacteria (Balbi et al. 2009; Hershberg and Petrov 2010; Hildebrand et al. 2010). Given that mutation is always AT biased, some other force must be driving elevated GC content in bacteria with intermediate to high GC content. The most obvious culprit is natural selection, favoring such higher GC content, but other non-selective mechanisms could also be involved.

One nonselective mechanism that may be driving GC content up in bacteria with intermediate to high GC content, is biased gene conversion (BGC) (reviewed in Duret and Galtier 2009). It has been shown that gene conversion is GC biased in many eukaryotes, including humans and other mammals. In other words, the probability of a GC allele to be passed on to the next generation through gene conversion is higher in these eukaryotes than that of an AT allele. As a result of such BGC, in these eukaryotes, regions with lower recombination rates tend to be more AT rich, whereas regions undergoing more recombination will tend to be more GC rich (Fullerton et al. 2001). A relationship between levels of recombination and GC content was also demonstrated for many bacteria, suggesting that BGC, or a mechanism similar to BGC, may affect nucleotide content in bacteria in a similar manner (Touchon et al. 2009; Lasalle et al. 2015).

A second nonselective mechanism that may be increasing GC content in bacteria relates to mismatch-repair (MMR) systems. Lee et al. (2012) conducted MA experiments on both wild-type *E. coli* and mutants deficient in MMR. They found that although mutation was indeed AT biased in wild-type *E. coli*, it was GC biased in the absence of MMR. This suggests that the nucleotide content of genomes might be influenced by how well their MMR systems function (Lee et al. 2012). Nucleotide content is a slowly evolving trait, because many substitutions need to occur for genome-wide patterns of nucleotide content to substantially change. Therefore, the nucleotide content of a bacterium may not be influenced solely by its current MMR functionality. Rather, MMR function during the evolution of the lineage to which the bacterium belongs may influence its current GC content. Fitting with this, it has been shown that the relationship between the nucleotide content of a bacterium and the current presence of MMR genes within its genome is not a straightforward one (Garcia-Gonzalez et al. 2012).

When it comes to selection affecting nucleotide content, the first big question that arises concerns the nature of selection. If indeed natural selection favors higher GC content in some bacteria, why? What is the advantage conferred on these bacteria by having higher genome-wide GC content? The currently available answers to this question are far from complete.

A study that examined metagenomic samples collected from aquatic and soil environments, was successful in demonstrating that soil bacteria are substantially more GC rich than aquatic bacteria, even when differences in phylogeny are accounted for (Foerstner et al. 2005). These results suggest that environmental selection plays a role in determining nucleotide content. The study in question was performed in 2005 when metagenomic data were only starting to become available, and used samples from only four different environments (Foerstner et al. 2005). A more recent study used a much larger collection of 183 metagenomic data sets, extracted from 14 environment types, to investigate the effects of environment on nucleotide composition (Reichenberger et al. 2015). This

study supported the results of the smaller scale metagenomic analysis and demonstrated that environment affects microbial nucleotide content in a manner that cannot be entirely explained by differences in phylogenetic composition. Intriguingly, the data used in the more recent study made it possible to show that environmental factors drive changes in nucleotide content, not only between highly diverged environment types (e.g., soil vs. aquatic), but also between samples extracted from the guts of different human subjects (Reichenberger et al. 2015). These results imply that the environmental factors that select for certain nucleotide compositions may be quite subtle.

The most obvious reason selection would favor high GC content in some bacteria is that higher GC content may provide better genome stability when temperatures are elevated. Many studies have attempted to investigate the correlation between GC content and optimal growth temperatures, with mixed results (Galtier and Lobry 1997; Lobry 1997; Hurst and Merchant 2001; Marashi and Ghalanbor 2004; Musto et al. 2004, 2006; Wang et al. 2006). In the end, it is very possible that growth temperature does affect nucleotide content. However, high growth temperatures are likely not the only environmental factors affecting nucleotide content, and they likely do not explain why so many bacteria have high or intermediate GC content in the face of universally AT-biased mutation.

Recently, Raghavan et al. (2012) have suggested an alternative force selecting for elevated GC content related to gene expression. Raghavan et al. inserted a plasmid containing the green fluorescent protein (GFP) gene into strains of *E. coli*. They generated their GFP genes to differ in the GC content of their synonymous sites. This allowed them to show that strains harboring a more GC-rich GFP gene grew faster than strains harboring a more AT-rich version of the gene, in a manner that depended on the construct being expressed, at both the mRNA and protein levels (Raghavan et al. 2012). They then showed that this effect was not limited to the GFP gene but also occurred when other genes were so inserted into *E. coli* (Raghavan et al. 2012). This finding fits the observation

R. Hershberg

that bacteria with intermediate to high GC content tend use GC-rich optimal codons (Hershberg and Petrov 2009)—a trend that results in a much higher GC content of protein-coding synonymous sites, compared with noncoding intergenic sequences, within GC-intermediate and GC-rich genomes (Hershberg and Petrov 2009, 2012; Raghavan et al. 2012). If indeed GC-rich coding sequences are expressed more efficiently and/or accurately, selection may indeed drive GC content up in coding sequences. However, this suggested mechanism does not explain why intergenic, noncoding regions also have higher GC content than expected at mutational equilibrium in genomes with intermediate to high GC content (Hershberg and Petrov 2010). Thus, although selection for genes to be more GC rich may contribute to elevated GC content, it cannot explain them in their entirety.

A second question that arises when considering natural selection acting on nucleotide composition is the question of how such selection would work. A problem arises because each individual base mutation only minutely alters overall nucleotide content, and an enormous number of mutations are needed to have any significant effect on overall nucleotide content. If so, how can selection on GC content affect each individual mutation? Additionally, if selection were to affect each mutation, the associated genetic load would be staggering. A possible solution to this conundrum is that natural selection may not act on individual mutations. Rather if there is selection in favor of elevated GC content and there is a nonselective mechanism, such as BGC, that elevates GC content (Duret and Galtier 2009), it is possible that strong selection will exist on that mechanism. For example, if indeed BGC affects nucleotide content in some bacteria, as has been shown for eukaryotes (Duret and Galtier 2009), bacteria that lose the ability to carry out BGC may gradually become more AT rich. Once their GC content becomes low enough to be disfavored by natural selection, these bacteria will be removed from the population. In this example, it is not each GC to AT mutation that is affected by selection, but rather the mutational event that

leads to the loss of BGC. This is currently just an idea, and much further theoretical and experimental work needs to be performed to examine its validity.

Variation in Mutation Rates along the Chromosome

Mutation may also bias patterns of genetic variation if certain regions of the genome are more prone to mutation than other regions. In a recent study, Foster et al. (2013) sequenced 24 MA lines of MMR defective *E. coli*. They found a striking pattern by which mutations are not randomly distributed along the chromosome. Rather, mutations fall in a wave-like pattern that is repeated in an almost exact mirror image in the two separately replicated halves (replisomes) of the *E. coli* chromosome (Foster et al. 2013). They further showed that mutation density was higher in regions of the *E. coli* chromosome where gene expression is regulated by nucleotide-associated proteins. These results were interpreted by Foster et al. (2013) to imply that mutation rates are affected by chromosome structure.

In a recent study, Martincorena et al. (2012) claimed to show that mutation rates are significantly lower in highly expressed genes and genes undergoing stronger selection. They postulated that by lowering mutation rates, particularly in genes that are more highly expressed and more important, *E. coli* was using an evolutionary risk-management strategy. These results were obtained by analyzing patterns of synonymous substitution between 34 *E. coli* strains, and relied on an assumption that these patterns of substitution evolved under relaxed selection, because of close relatedness of these strains (Martincorena et al. 2012). It is important to note, however, that different *E. coli* strains are highly diverged and that patterns of substitution between strains of *E. coli* are, in fact, subject to extremely strong selection (Hershberg et al. 2007). It is therefore quite possible that the differences in the frequency of *E. coli* synonymous substitutions between highly expressed and less highly expressed genes are because of selection, rather than mutation. Indeed, it was very re-

cently shown that the theory of adaptive risk management via lowering of mutation rates in highly expressed genes is theoretically untenable (Chen and Zhang 2013). Furthermore, the negative correlation suggested by Martincorena et al. (2012) between mutation rates and levels of expression was not supported by MA studies in *E. coli*, *Salmonella*, and yeast (Lind and Andersson 2008; Lee et al. 2012; Park et al. 2012; Chen and Zhang 2013; Foster et al. 2013). To the contrary, in some MA studies, a significant positive correlation is observed between levels of expression and mutation frequencies (Lind and Andersson 2008; Park et al. 2012; Chen and Zhang 2013).

CONSTITUTIVE MUTATORS AND STRESS-INDUCED MUTAGENESIS

As mentioned above, natural selection is thought to favor the lowering of mutation rates, because many mutations are deleterious (Kimura 1967; Drake 1991; Dawson 1998; Lynch 2010). In sharp contrast to this expectation, it was observed that ~1% of all natural bacterial isolates are mutators that have high mutation rates, compared with the remainder of the population (Gross and Siegel 1981; LeClerc et al. 1996). If indeed selection disfavors high mutation rates, why would hypermutating bacteria be present at such high frequencies? The best explanation currently available is that mutators accelerate adaptation in asexual clonal populations (Sniegowski et al. 1997; Taddei et al. 1997; Giraud et al. 2001; Notley-McRobb et al. 2002). Mutator alleles may thus be linked to adaptive alleles that arise as a result of hypermutation. It is therefore thought that when bacteria are exposed to strong pressure to adapt quickly (e.g., when they are faced with new challenges), mutator alleles may become beneficial, which increases their frequencies (Sniegowski et al. 1997; Taddei et al. 1997; Giraud et al. 2001; Notley-McRobb et al. 2002).

The mutators discussed above are constitutive mutators—bacteria that are defective in their repair mechanisms and that constitutively mutate at higher frequencies (LeClerc et al. 1996). However, it has also been postulated

that bacteria may be able to selectively increase mutation rates when they are exposed to certain “stressful” or growth-limiting conditions (reviewed in Foster 2007; Galhardo et al. 2007). Modeling has shown that such stress-induced mutagenesis (SIM) should be highly beneficial (Ram and Hadany 2012), as it could allow bacteria to transiently increase mutagenesis particularly when they are most pressured to adapt. Yet, the study of SIM has been plagued by fierce debate (e.g., Slechta et al. 2002, 2003; Roth et al. 2003, 2006; Wrande et al. 2008; Katz and Hershberg 2013). In this review, I do not have sufficient space to delve into the full debate, but will only introduce some points of contention.

The strongest support of SIM, and the most detailed understanding of its mechanisms has come from the use of a particular assay suggested originally by Cairns and Foster (1991). In this assay, a special *E. coli* strain, deleted for its chromosomal *lac* operon, and carrying a *lacI-lacZ* fusion gene with a frameshift mutation in *lacI* on an F' conjugative plasmid, is plated onto lactose plates. On such plates, only cells that become *lac* positive can form colonies, and so the frequency of reversion mutants can be monitored. Original proponents of SIM assumed that growth could only be achieved on the plates by reversion mutants that corrected the frameshift mutation in *lacI*. Colonies forming from mutants that arose before plating were expected to emerge within 2 days of plating, and any subsequent colonies were assumed to result from mutations occurring on the plates, in non-growing bacteria. Any such mutations were assumed to be the result of SIM.

Studies utilizing the Cairns and Foster Lac assay suggested that the occurrence of stress-induced frameshift mutations depended on double-strand breaks (DSBs), repair of these DSBs by an error-prone polymerase, *dinB*, and also depended on the bacterial stress response, mediated by the stationary phase σ factor, *rpoS* (reviewed in Galhardo et al. 2007). Although these results suggested a mechanism by which SIM could occur, use of the Lac assay was severely debated. First, it was argued that increased reversion could be caused by amplifica-

R. Hershberg

tion of the inactive lac gene, slow growth of cells carrying this amplification, consequent frameshift reversion mutations, and selection for these mutants that could now grow freely on the lactose plates (reviewed in Roth et al. 2006). Thus, it was suggested that frameshift reversions were not necessarily because of SIM. Second, it was argued that the particular F' conjugative plasmid used was problematic, as it contained an extra copy of the *dinB* gene that was shown to be important for increased frequency of reversion (Roth et al. 2006). It was therefore argued that the results obtained using the Lac assay were not general, but rather particular to the assay used.

To address these concerns, Shee et al. (2011) more recently developed an alternative chromosomal assay for studying SIM. In this assay, the frequency of frameshift reversions to an artificially introduced tetracycline resistance cassette containing a deactivating frameshift mutation is quantified. DSBs are induced artificially by placing the tetracycline cassette 8.5 kb from an I-sceI double-strand endonuclease cut-site. The cells are engineered to contain an SceI gene, controlled by a P_{BAD} promoter, which is repressed when glucose is available, but derepressed once glucose becomes depleted and cells begin to starve. Using this assay, Shee et al. (2011) could show that there was an increase in the frequency of tetracycline-resistant reversion mutants in response to starvation. This increase was shown to be dependent on DSBs, *dinB*, and *rpoS* (Shee et al. 2011). Shee et al. interpreted their results as demonstrating that results obtained using the Lac assay are not specific to that assay, and that SIM indeed occurs in *E. coli*, and depends on the stress response being induced and on error-prone repair of DSBs.

So far, I have discussed SIM as it has been studied in artificial laboratory models, but has SIM been shown to occur within natural bacterial populations? Until very recently, the best, most well cited evidence for the natural occurrence of SIM came from experiments conducted by Bjedov et al. in 2003 (Bjedov et al. 2003). In these experiments, ~800 natural isolates of *E. coli*, extracted from a large variety of host-associated and non-host-associated environ-

ments were tested for the frequency with which they accumulate resistance to rifampicin in young and aging colonies. It was observed that, to varying extents in different isolates, the frequency of mutants resistant to rifampicin increases in aging colonies compared with young colonies. This increase in the frequency of resistant mutants was a priori assumed to result from increased mutagenesis, resulting from the starvation stress incurred via growth in aging colonies. Indeed the resulting paper was titled “stress-induced mutagenesis in bacteria” (Bjedov et al. 2003). A subsequent study, published in 2008, showed that increased frequency of resistance to rifampicin could also be explained by natural selection, as it showed that many rifampicin-resistant mutants carried a growth advantage in aging colonies (Wrande et al. 2008). Yet the Bjedov et al. study continued to be very widely cited as conclusive evidence for the occurrence of SIM within natural bacterial populations (e.g., Bogumil and Dagan 2012; Buerger et al. 2012; Feher et al. 2012; Obolski and Hadany 2012; Rosenberg et al. 2012; Ryall et al. 2012; Sanchez-Alberola et al. 2012; Maclean et al. 2013; Martincorena and Luscombe 2013).

We have recently repeated the Bjedov et al. experiments on a single laboratory strain of *E. coli*. Consistent with their results, we were able to show a substantial increase in the frequency of resistance to rifampicin in aging colonies compared with young colonies. We also observed a sharp increase in the frequency of resistance to a second antibiotic, nalidixic acid (Katz and Hershberg 2013). We then used whole-genome sequencing to show conclusively that increased mutagenesis could not explain the increased frequency of resistance observed to either of the two antibiotics (Katz and Hershberg 2013). Therefore, SIM cannot explain the Bjedov et al. results, and these results cannot be seen as evidence of SIM occurring in natural bacterial populations.

We further showed that, as was previously shown for rifampicin resistance mutations (Wrande et al. 2008), nalidixic acid resistance mutations can also confer a growth advantage in aging colonies (Katz and Hershberg 2013).

An additional study showed that a mutation conferring resistance to streptomycin can also improve growth when bacteria are grown on poor carbon sources (Paulander et al. 2009). Combined, these results show that using antibiotic resistance as a marker for the study of mutation in general and SIM in particular may be highly problematic.

CONCLUDING REMARKS

Much remains to be understood about the rates and patterns of mutation and about how these vary between different bacterial isolates, within populations, as a factor of growth conditions, and along the chromosome. Mutation is difficult to study because it is a highly noisy process and because it affects variation in a manner that is highly entangled with the effects of natural selection. To characterize the effects of mutation, we need to acknowledge these complications and find creative ways to address them. Future studies will undoubtedly take advantage of our increasing ability to examine variation at the whole-genome level to reveal much more about mutation and how it acts as an engine of evolution in bacteria and beyond.

ACKNOWLEDGMENTS

I thank Sophia Katz, Wesley Field, and Talia Karasov for their helpful comments. R.H. is supported by a European Research Council (ERC) FP7 CIG Grant (No. 321780), by a BSF Grant (No. 2013463), by a Yigal Allon Fellowship awarded by the Israeli Council for Higher Education, and by the Robert J. Shillman Career Advancement Chair. Work by R.H. is performed in the Rachel & Menachem Mendelovitch Evolutionary Process of Mutation & Natural Selection Research Laboratory.

REFERENCES

Akashi H. 1995. Inferring weak selection from patterns of polymorphism and divergence at “silent” sites in *Drosophila* DNA. *Genetics* **139**: 1067–1076.

Andersson JO, Andersson SG. 1999. Insights into the evolutionary process of genome degradation. *Curr Opin Genet Dev* **9**: 664–671.

Bacterial Mutation as an Evolutionary Force

Balbi KJ, Rocha EP, Feil EJ. 2009. The temporal dynamics of slightly deleterious mutations in *Escherichia coli* and *Shigella* spp. *Mol Biol Evol* **26**: 345–355.

Bjedov I, Tenaillon O, Gerard B, Souza V, Denamur E, Radman M, Taddei F, Matic I. 2003. Stress-induced mutagenesis in bacteria. *Science* **300**: 1404–1409.

Bogumil D, Dagan T. 2012. Cumulative impact of chaperone-mediated folding on genome evolution. *Biochemistry* **51**: 9941–9953.

Brockhurst MA, Colegrave N, Rozen DE. 2010. Next-generation sequencing as a tool to study microbial evolution. *Mol Ecol* **20**: 972–980.

Buerger S, Spoering A, Gavriš E, Leslin C, Ling L, Epstein SS. 2012. Microbial scout hypothesis, stochastic exit from dormancy, and the nature of slow growers. *Appl Environ Microbiol* **78**: 3221–3228.

Cairns J, Foster PL. 1991. Adaptive reversion of a frameshift mutation in *Escherichia coli*. *Genetics* **128**: 695–701.

Chen X, Zhang J. 2013. No gene-specific optimization of mutation rate in *Escherichia coli*. *Mol Biol Evol* **30**: 1559–1562.

Dawson KJ. 1998. Evolutionarily stable mutation rates. *J Theor Biol* **194**: 143–157.

Drake JW. 1991. A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci* **88**: 7160–7164.

Drake JW. 2009. Avoiding dangerous missense: Thermophiles display especially low mutation rates. *PLoS Genet* **5**: e1000520.

Duret L, Galtier N. 2009. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu Rev Genomics Hum Genet* **10**: 285–311.

Elena SF, Lenski RE. 2003. Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nat Rev Genet* **4**: 457–469.

Feher T, Bogos B, Mehi O, Fekete G, Csorgo B, Kovacs K, Posfai G, Papp B, Hurst LD, Pal C. 2012. Competition between transposable elements and mutator genes in bacteria. *Mol Biol Evol* **29**: 3153–3159.

Foerster KU, von Mering C, Hooper SD, Bork P. 2005. Environments shape the nucleotide composition of genomes. *EMBO Rep* **6**: 1208–1213.

Foster PL. 2007. Stress-induced mutagenesis in bacteria. *Crit Rev Biochem Mol Biol* **42**: 373–397.

Foster PL, Hanson AJ, Lee H, Popodi EM, Tang H. 2013. On the mutational topology of the bacterial genome. *G3 (Bethesda)* **3**: 399–407.

Fullerton SM, Bernardo Carvalho A, Clark AG. 2001. Local rates of recombination are positively correlated with GC content in the human genome. *Mol Biol Evol* **18**: 1139–1142.

Galhardo RS, Hastings PJ, Rosenberg SM. 2007. Mutation as a stress response and the regulation of evolvability. *Crit Rev Biochem Mol Biol* **42**: 399–435.

Galtier N, Lobry JR. 1997. Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J Mol Evol* **44**: 632–636.

Garcia-Gonzalez A, Rivera-Rivera RJ, Massey SE. 2012. The presence of the DNA repair genes *mutM*, *mutY*, *mutL*, and

R. Hershberg

- mutS* is related to proteome size in bacterial genomes. *Front Genet* **3**: 3.
- Gillespie J. 1998. *Population genetics, a concise guide*, 1st ed. The Johns Hopkins University Press, Baltimore.
- Giraud A, Radman M, Matic I, Taddei F. 2001. The rise and fall of mutator bacteria. *Curr Opin Microbiol* **4**: 582–585.
- Graur D, Li W. 2000. *Fundamentals of molecular evolution*. Sinauer Associates, Sunderland, MA.
- Gross MD, Siegel EC. 1981. Incidence of mutator strains in *Escherichia coli* and coliforms in nature. *Mutat Res* **91**: 107–110.
- Hershberg R, Petrov DA. 2009. General rules for optimal codon choice. *PLoS Genet* **5**: e1000556.
- Hershberg R, Petrov DA. 2010. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet* **6**: e1001115.
- Hershberg R, Petrov DA. 2012. On the limitations of using ribosomal genes as references for the study of codon usage: A rebuttal. *PLoS ONE* **7**: e49060.
- Hershberg R, Tang H, Petrov DA. 2007. Reduced selection leads to accelerated gene loss in *Shigella*. *Genome Biol* **8**: R164.
- Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, Roach JC, Kremer K, Petrov DA, Feldman MW, et al. 2008. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol* **6**: e311.
- Hildebrand F, Meyer A, Eyre-Walker A. 2010. Evidence of selection upon genomic GC-content in bacteria. *PLoS Genet* **6**: e1001107.
- Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill FX, Goodhead I, Rance R, Baker S, Maskell DJ, Wain J, et al. 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nat Genet* **40**: 987–993.
- Hurst LD, Merchant AR. 2001. High guanine-cytosine content is not an adaptation to high temperature: A comparative analysis amongst prokaryotes. *Proc Biol Sci* **268**: 493–497.
- Katz S, Hershberg R. 2013. Elevated mutagenesis does not explain the increased frequency of antibiotic resistant mutants in starved aging colonies. *PLoS Genet* **9**: e1003968.
- Kimura M. 1967. On the evolutionary adjustment of spontaneous mutation rates. *Genet Res* **9**: 23–34.
- Kuo CH, Ochman H. 2010. The extinction dynamics of bacterial pseudogenes. *PLoS Genet* **6**: e1001050.
- Lassalle F, Périán S, Bataillon T, Nesme X, Duret L, Daubin V. 2015. GC-content evolution in bacterial genomes: The biased gene conversion hypothesis expands. *PLoS Genet* **11**: e1004941.
- LeClerc JE, Li B, Payne WL, Cebula TA. 1996. High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. *Science* **274**: 1208–1211.
- Lee H, Popodi E, Tang H, Foster PL. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci* **109**: E2774–E2783.
- Lieber MR. 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79**: 181–211.
- Lieberman TD, Michel JB, Aingaran M, Potter-Bynoe G, Roux D, Davis MR Jr, Skurnik D, Leiby N, LiPuma JJ, Goldberg JB, et al. 2011. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat Genet* **43**: 1275–1280.
- Lind PA, Andersson DI. 2008. Whole-genome mutational biases in bacteria. *Proc Natl Acad Sci* **105**: 17878–17883.
- Lobry JR. 1997. Influence of genomic G+C content on average amino-acid composition of proteins from 59 bacterial species. *Gene* **205**: 309–316.
- Luria SE, Delbruck M. 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**: 491–511.
- Lynch M. 2007. *The origins of genome architecture*. Sinauer Associates, Sunderland, MA.
- Lynch M. 2010. Evolution of the mutation rate. *Trends Genet* **26**: 345–352.
- Maclean RC, Torres-Barcelo C, Moxon R. 2013. Evaluating evolutionary models of stress-induced mutagenesis in bacteria. *Nat Rev Genet* **14**: 221–227.
- Marashi SA, Ghalanbor Z. 2004. Correlations between genomic GC levels and optimal growth temperatures are not “robust.” *Biochem Biophys Res Commun* **325**: 381–383.
- Martincorena I, Luscombe NM. 2013. Non-random mutation: The evolution of targeted hypermutation and hypomutation. *BioEssays* **35**: 123–130.
- Martincorena I, Seshasayee AS, Luscombe NM. 2012. Evidence of non-random mutation rates suggests an evolutionary risk management strategy. *Nature* **485**: 95–98.
- Messer PW. 2009. Measuring the rates of spontaneous mutation from deep and large-scale polymorphism data. *Genetics* **182**: 1219–1232.
- Modrich P. 1991. Mechanisms and biological effects of mismatch repair. *Annu Rev Genet* **25**: 229–253.
- Musto H, Naya H, Zavala A, Romero H, Alvarez-Valin F, Bernardi G. 2004. Correlations between genomic GC levels and optimal growth temperatures in prokaryotes. *FEBS Lett* **573**: 73–77.
- Musto H, Naya H, Zavala A, Romero H, Alvarez-Valin F, Bernardi G. 2006. Genomic GC level, optimal growth temperature, and genome size in prokaryotes. *Biochem Biophys Res Commun* **347**: 1–3.
- Muto A, Osawa S. 1987. The guanine and cytosine content of genomic DNA and bacterial evolution. *Proc Natl Acad Sci* **84**: 166–169.
- Nachman MW, Crowell SL. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics* **156**: 297–304.
- Notley-McRobb L, Seeto S, Ferenci T. 2002. Enrichment and elimination of *mutY* mutators in *Escherichia coli* populations. *Genetics* **162**: 1055–1062.
- Obolski U, Hadany L. 2012. Implications of stress-induced genetic variation for minimizing multidrug resistance in bacteria. *BMC Med* **10**: 89.
- Park C, Qian W, Zhang J. 2012. Genomic evidence for elevated mutation rates in highly expressed genes. *EMBO Rep* **13**: 1123–1129.

Bacterial Mutation as an Evolutionary Force

- Paulander W, Maisnier-Patin S, Andersson DI. 2009. The fitness cost of streptomycin resistance depends on *rpsL* mutation, carbon source and RpoS (σ^S). *Genetics* **183**: 539–546.
- Raghavan R, Kelkar YD, Ochman H. 2012. A selective force favoring increased G+C content in bacterial genes. *Proc Natl Acad Sci* **109**: 14504–14507.
- Ram Y, Hadany L. 2012. The evolution of stress-induced hypermutation in asexual populations. *Evolution* **66**: 2315–2328.
- Reichenberger ER, Rosen G, Hershberg U, Hershberg R. 2015. Prokaryotic nucleotide composition is shaped by both phylogeny and the environment Genome. *Biol Evol* **7**: 1380–1389.
- Rosenberg SM, Shee C, Frisch RL, Hastings PJ. 2012. Stress-induced mutation via DNA breaks in *Escherichia coli*: A molecular mechanism with implications for evolution and medicine. *BioEssays* **34**: 885–892.
- Roth JR, Kofoed E, Roth FR, Berg OG, Seger J, Andersson DI. 2003. Regulating general mutation rates: Examination of the hypermutable state model for Cairnsian adaptive mutation. *Genetics* **163**: 1483–1496.
- Roth JR, Kugelberg E, Reams AB, Kofoed E, Andersson DI. 2006. Origin of mutations under selection: The adaptive mutation controversy. *Annu Rev Microbiol* **60**: 477–501.
- Ryall B, Eydallin G, Ferenci T. 2012. Culture history and population heterogeneity as determinants of bacterial adaptation: The adaptomics of a single environmental transition. *Microbiol Mol Biol Rev* **76**: 597–625.
- Sanchez-Alberola N, Campoy S, Barbe J, Erill I. 2012. Analysis of the SOS response of *Vibrio* and other bacteria with multiple chromosomes. *BMC Genomics* **13**: 58.
- Shee C, Gibson JL, Darrow MC, Gonzalez C, Rosenberg SM. 2011. Impact of a stress-inducible switch to mutagenic repair of DNA breaks on mutation in *Escherichia coli*. *Proc Natl Acad Sci* **108**: 13659–13664.
- Slechta ES, Liu J, Andersson DI, Roth JR. 2002. Evidence that selected amplification of a bacterial *lac* frameshift allele stimulates Lac⁺ reversion (adaptive mutation) with or without general hypermutability. *Genetics* **161**: 945–956.
- Slechta ES, Bunny KL, Kugelberg E, Kofoed E, Andersson DI, Roth JR. 2003. Adaptive mutation: General mutagenesis is not a programmed response to stress but results from rare coamplification of *dinB* with *lac*. *Proc Natl Acad Sci* **100**: 12847–12852.
- Smith KC. 1992. Spontaneous mutagenesis: Experimental, genetic and other factors. *Mutat Res* **277**: 139–162.
- Sniegowski PD, Gerrish PJ, Lenski RE. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**: 703–705.
- Sueoka N. 1962. On the genetic basis of variation and heterogeneity of DNA base composition. *Proc Natl Acad Sci* **48**: 582–592.
- Sung W, Ackerman MS, Miller SE, Doak TG, Lynch M. 2012. Drift-barrier hypothesis and mutation-rate evolution. *Proc Natl Acad Sci* **109**: 18488–18492.
- Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. 1997. Role of mutator alleles in adaptive evolution. *Nature* **387**: 700–702.
- Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, et al. 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**: e1000344.
- Wang HC, Susko E, Roger AJ. 2006. On the correlation between genomic G+C content and optimal growth temperature in prokaryotes: Data quality and confounding factors. *Biochem Biophys Res Commun* **342**: 681–684.
- Wrande M, Roth JR, Hughes D. 2008. Accumulation of mutants in “aging” bacterial colonies is due to growth under selection, not stress-induced mutagenesis. *Proc Natl Acad Sci* **105**: 11863–11868.

Cold Spring Harbor Perspectives in Biology

Mutation—The Engine of Evolution: Studying Mutation and Its Role in the Evolution of Bacteria

Ruth Hershberg

Cold Spring Harb Perspect Biol 2015; doi: 10.1101/cshperspect.a018077

Subject Collection [Microbial Evolution](#)

Not So Simple After All: Bacteria, Their Population Genetics, and Recombination

William P. Hanage

Realizing Microbial Evolution

Howard Ochman

Thoughts Toward a Theory of Natural Selection: The Importance of Microbial Experimental Evolution

Daniel Dykhuizen

Coevolution of the Organization and Structure of Prokaryotic Genomes

Marie Touchon and Eduardo P.C. Rocha

Mutation—The Engine of Evolution: Studying Mutation and Its Role in the Evolution of Bacteria

Ruth Hershberg

The Origin of Mutants Under Selection: How Natural Selection Mimics Mutagenesis (Adaptive Mutation)

Sophie Maisnier-Patin and John R. Roth

Evolution of New Functions De Novo and from Preexisting Genes

Dan I. Andersson, Jon Jerlström-Hultqvist and Joakim Näsvall

Genome-Based Microbial Taxonomy Coming of Age

Philip Hugenholtz, Adam Skarshewski and Donovan H. Parks

Horizontal Gene Transfer and the History of Life

Vincent Daubin and Gergely J. Szöllosi

Early Microbial Evolution: The Age of Anaerobes

William F. Martin and Filipa L. Sousa

Microbial Speciation

B. Jesse Shapiro and Martin F. Polz

The Evolution of *Campylobacter jejuni* and *Campylobacter coli*

Samuel K. Sheppard and Martin C.J. Maiden

Paleobiological Perspectives on Early Microbial Evolution

Andrew H. Knoll

For additional articles in this collection, see <http://cshperspectives.cshlp.org/cgi/collection/>

**All Modifications and
Oligo Types Synthesized**
Long Oligos • Fluorescent • Chimeric • DNA • RNA • Antisense

Oligo Modifications?
Your wish is our command.