NATURAL SELECTION AND THE EMERGENCE OF A MUTATION PHENOTYPE: An Update of the Evolutionary Synthesis Considering Mechanisms that Affect Genome Variation

Lynn Helena Caporale

One Sherman Square, New York, NY 10023; email: caporale@usa.net

Key Words implicit genome, Darwin, explicit genome, codon degeneracy, reversible mutations, contingency genes

■ **Abstract** Most descriptions of evolution assume that all mutations are completely random with respect to their potential effects on survival. However, much like other phenotypic variations that affect the survival of the descendants, intrinsic variations in the probability, type, and location of genetic change can feel the pressure of natural selection. From site-specific recombination to changes in polymerase fidelity and repair of DNA damage, an organism's gene products affect what genetic changes occur in its genome. Through the action of natural selection on these gene products, potentially favorable mutations can become more probable than random. With examples from variation in bacterial surface proteins to the vertebrate immune response, it is clear that a great deal of genetic change is better than "random" with respect to its potential effect on survival. Indeed, some potentially useful mutations are so probable that they can be viewed as being encoded implicitly in the genome. An updated evolutionary theory includes emergence, under selective pressure, of genomic information that affects the probability of different classes of mutation, with consequences for genome survival.

CONTENTS

OVERVIEW

Charles Darwin and Alfred Russel Wallace (37b) provided a robust framework for studying evolution: From among "the amount of individual variation that my experience as a collector had shown me to exist (123)," the "most fitted" survive; but they lacked the tools to investigate the source of that variation. As genes and mutation were incorporated into evolutionary theory (37a), evolution began to be described in terms of "random mutation" followed by natural selection (49). That mutation of DNA is completely "random" was not, of course, Darwin's idea.

It has been argued that mutations must be random because natural selection cannot "assist the process of evolutionary change," since "selection lacks foresight, and no one has described a plausible way to provide it" (34). If the challenges that confronted genomes were unprecedented and completely random, it would be hard to disagree with the statement that selection "lacks foresight." However, to the extent that classes of challenges and opportunities tend to recur, a response that is better than random can be favored by selection (17, 18, 88). Examples of recurring challenges include host/pathogen battles, access to valuable information encoded by other genomes, and the evolution of new members of gene families.

Due in part to the assumption that mutation is random, most discussions of evolution have focused on selection rather than on the biochemical mechanisms related to the generation of the variation upon which natural selection acts. Yet intrinsic variations in the physical-chemical properties of the DNA sequence context, and its interactions with polymerase, proofreading, repair, and recombination machinery, alter the probability of distinct types of mutation along a DNA sequence. Darwin "called [the] principle, by which each slight variation, if useful, is preserved, by the term Natural Selection" and asked "why should we doubt that variations in any way useful to beings . . . would be preserved, accumulated, and inherited?" (26). Variations in the probability of mutation along a genome can be "in any way useful to beings" and thus "preserved, accumulated, and inherited."

It is appropriate that a proposal to update evolutionary theory appear in the *Annual Reviews of Microbiology*, because insights from microbiology (and immunology), when considered together, deepen our understanding of the reach of natural selection and thus represent a breakthrough in our understanding of evolution. The research reviewed here leads to the conclusion that, under the pressure of natural selection, a "mutation phenotype" evolves in which, first, certain classes of mutation are more probable than others and, second, some of the more probable classes of mutation can have an increased probability of being useful, or at least not harmful, compared with completely random mutation (15, 18). (This is not the same as suggesting that a genome "knows" that it if it replaces a particular A with a G, it will be able to digest a specific sugar.) This review begins with the familiar

subject of site-specific recombination, refers to the multiple biochemical activities that, when integrated, result in the probability of distinct mutations along a genome's sequence, and then discusses mechanisms that are known to focus indels and point mutations along a nucleic acid sequence in a manner that can provide a selective advantage.

SECOND-ORDER SELECTION AND SITE-SPECIFIC FRAMEWORKS FOR EVOLUTION

Horizontal Transfer

Specific biochemical mechanisms have evolved that enable the horizontal transfer of blocks of DNA. This DNA often encodes pathogenicity, antibiotic resistance, or the ability to take up and utilize a new food source, such as lactose (39, 91). Integrons provide a framework for transfer of intact, expressible genes from organism to organism (59, 104). Phage also spread information. The integration of *CTXphi*, which encodes cholera toxin, specifically at *dif*-like sites in the *Vibrio cholerae* genome, is mediated by host-encoded *XerCD* recombinases, which are widely distributed among bacterial species. Thsese recombinases likely mediate the integration of other filamentous phage in various bacterial species, and indeed other mobile elements (66). A comparison of strains of *Streptococcus* A, including one isolated from a patient experiencing toxic shock syndrome, led to the conclusion that outbreaks of particularly virulent disease emerge from ongoing combinatorial assortment of virulence factors by phage-mediated recombination (9).

When the pathogenic *Escherichia coli* 0157:H7 was compared with the non-pathogenic laboratory strain *E. coli* K-12, there were 75,168 mostly "synonymous" individual base pair changes, but inserted clusters added up to 1.34 million bp unique to 0157:H7 (95). Indeed, horizontal transfer plays a role in evolution that dramatically rivals single nucleotide changes and can occur between species, such as *E. coli* and *Salmonella enterica* (11, 75). Type III protein secretion systems have spread widely and are found in a broad range of both pathogenic and symbiotic organisms, including those essential for nitrogen fixation (35, 67).

In contrast to genes selected for their ability, for example, to enable the digestion of a specific new sugar, DNA recognition sites and enzymes involved in horizontal transfer of genetic information emerge under what has been termed second-order selection (3): selection for their ability to access information that has evolved in other genomes, which in turn provides a potential selective advantage to generations of descendants of each genome that acquires and retains this ability (76).

Hidden Messages: The Degeneracy of the Genetic Code

FOCUSED HYPERMUTATON   The generation of antibody diversity begins with site-directed recombination of one out of the many V regions encoded by a genome into an expression site beside a J/C region. Following recombination, DNA encoding the antibody binding site exhibits an intrinsically increased mutation rate (i.e., increased even in the absence of selection for the ability to bind to an antigen).

Hypermutation in the variable region of a recombined immunoglobulin gene is reported to begin with regulated, targeted, enzymatic deamination of a C on either strand at the sequence RGYW (either puRine, G, either pYrimidine, A or T) (32), followed by the action of a "mutator" polymerase(s) (38). Blocking uracil deglycosylase increased the ratio of transitions to transversions. Changing the DNA sequence, including changes between "synonymous" codons, changes the location of the mutation hotspots (54). Although not every antibody-producing cell that the immune system generates will bind to an antigen and therefore be selected for expansion, the underlying diversity that generates the antibodies is not based on random nucleotide change.

Thus, information that modulates the rate and type of genetic change can evolve within the protein-coding region of genes much as the protein-coding sequence itself evolves (15). The infrastructure that creates each functional antibody gene focuses variation at locations well matched to the functional requirements of the gene product, i.e., within the variable-region binding site. Thus, the probability of variation can become aligned with the potential biological effect of a mutation at that site. Of course, the new genes must face selection, but the event that created them is not completely random.

PATHOGEN VARIATION   Biochemical mechanisms that generate coat diversity enable arthropod-borne infectious agents to avoid removal by immune surveillance while remaining accessible in blood for transfer to a new host (4). *Borrelia burgdorferi* (19) exchanges patches in its coat protein through site-directed recombination (128). Second-order selection operates on the ability of *B. burgdorferi* to generate diversity, for when cultured in the laboratory, outside the selective pressure of a host immune system, *B. burgdorferi* tends to lose plasmids and infectivity (109). Although there are nearly 200 ways to encode the five amino acids at the borders of the varied coat patch, the plasmid-borne information encoding this EGAIK repeat is embedded in a completely conserved 17-bp repeat.

Conserved recognition sites for DNA invertase of long tail-fiber genes of the double-stranded tailed phage that infect enteric bacteria (105) enable a combinatorial assortment of host-specificity regions to be exchanged at the appropriate position of the tail fiber gene, either by inversion of a DNA segment that encodes two options in opposite orientations or by recombination (106), extending the potential host range.

REGION-SPECIFIC VARIATION   Although homologous and site-directed recombination are familiar mechanisms, the immunoglobulin class switch is an example of region-specific recombination (63). The breaks in DNA that initiate a class switch are not always between the same two base pairs, but they are always within the "switch regions." The "environment" of the B cell regulates the site of the DNA cut; for example, interleukin (IL)-4 directs the construction of a gene encoding IgE by inducing the expression of an appropriate "germline transcript" (80), which targets the double-strand break to the switch region that is upstream of DNA encoding the constant region of the epsilon heavy chain.

Endonuclease cuts in meiosis also appear to be region specific. Local DNA sequence and chromatin structure (10) and the presence of binding sites for certain proteins (122) affect the accessibility of the region to the endonuclease Spo11, which makes the double-strand break that initiates meiotic recombination (68). In *Schizosaccharomyces pombe*, about half the recombination events occurred within 50–200 bases of the hotspot sequence ATGACGT (27). In a 216-kb segment of the class II region of the human major histocompatibility complex, hotspots of crossover in sperm correspond to areas where linkage disequilibrium breaks down (70). Because mutations within Spo11 can alter the location at which the DNA will be cut and near which variation will be generated (33), they will feel the pressure of natural selection, as will the accessibility of each region of DNA and cleavage-prone sequences within the accessible region.

GENE DUPLICATION   Some locations in the genome are more likely than others to participate in gene duplication and amplification events. A genome-wide survey of changes during adaptation to thermal stress in *E. coli B* revealed repeated duplications of the same region of the chromosome and suggested that this was facilitated by repeats (102). In *Saccharomyces cerevisiae*, repeated, independent, but nevertheless similar, chromosomal rearrangements, including identical breakpoints at transposon-related sequences, emerged under the sustained strong selective pressure of growth in glucose-limited chemostats (36). Frequent duplication under selection, followed by rapid loss of the duplicates when the selective pressure is removed, has been described as a "reversible" form of mutation, observed, for example, under starvation conditions in which growth essentially is limited to organisms in which duplication of a region of the chromosome enables increased transport of the limiting carbon source (114).

The high intrinsic rate of genome variation in mammalian histocompatibility antigens, and at focused places in the immunoglobulins, points to the possibility of a genomic framework that facilitates evolution of gene families. Each time a gene is duplicated, as a gene family expands, the same challenge recurs: the need to avoid mutations that would destroy the common function of the gene family while changing other amino acids that underlie the new gene family member's target specificity (15, 17). If information that facilitates adaptation can evolve at sites of high variation of immunoglobulins, histocompatibility regions, contingency genes, and pathogen coats (5, 72, 129), it is likely to be found in other locations of which we currently are unaware and which should be the subject of future research (16).

A GENOME'S PATTERN OF MUTATION EVOLVES THROUGH A BALANCE OF MULTIPLE ACTIVITIES

Overview

The probability of distinct genetic changes varies in a sequence-context-dependent manner, affected by the K_m and k_{cat} of enzymes that polymerize and repair DNA,

and by the relative pool sizes of the nucleotides (53, 130). Changes in pool sizes, such as through changes in nucleotide diphosphate kinase activity, change the rate of distinct types of mutation through effects on both polymerase fidelity and mismatch repair (86).

A mutation in DNA that encodes a DNA polymerase affects the future probability of specific types of mutations at myriad places throughout the genome. For example, in *Haemophilus influenzae*, tetranucleotide repeats change in length more quickly when the activity of polI is decreased (7). For generation after generation, such mutations affect the viability, and thus the total number, of progeny that inherit an altered polymerase, along with its unique classes of more and less probable mutations.

Distinct mutation spectra result from changes in the activity of distinct components of mismatch repair and the polymerases (42, 56). Analysis of 164 spontaneous lacI⁻ mutations recovered from a uracil-DNA glycosylase-deficient strain of *E. coli* indicated that DNA context and different levels of gene expression and DNA repair all affect the classes and frequencies of "spontaneous" mutation (40).

DNA Sequence Context

The effect of the physical properties of each DNA sequence context on its own likelihood of mutation often is overlooked in discussions of evolution. DNA sequences can have profound effects on DNA structure (94, 97), the fidelity of DNA polymerases (69), and mismatch repair (81). Because sequence context affects the access and activity of distinct polymerase and repair proteins, sequence context affects local genome composition.

Certain mutations could be called "predictable" because they occur with orders-of-magnitude-higher probability than other mutations do; therefore, given a routinely achieved combination of time and population size, they essentially certainly will occur. "Correction" of quasipalindromes to perfect inverted repeats (99) occurs relatively frequently and preferentially during replication of the leading strand, whereas deletions between direct repeats, at sites where misalignment can be stabilized by sequence-context-dependent DNA secondary structure, are observed frequently and preferentially on the lagging strand (112). The leading and lagging strands also can have different probabilities that, for example, an A will mutate to a G, resulting in different base compositions on the two strands (101).

Emergence of a Mutation Phenotype

Through integration of a wide range of cellular activities, including the level of and balance between distinct repair, polymerase (12, 52), and proofreading activities encoded and expressed by that genome, and their interaction with different sites in the genome, an overall mutation rates emerges.

Although it is no surprise that mutation increases if exonucleolytic proofreading is decreased (108), at first glance it is surprising that mutation can increase if the activity of certain repair proteins is increased (46, 90). For example, in *E. coli*, increased expression of 3-methyladenine DNA glycosylase II (which excises damaged, but to some extent undamaged, bases from DNA) increases the mutation rate, as measured by increased "spontaneous" mutation to rifampicin resistance (in a manner that is sensitive to the local sequence context) (8). One of the strongest reported mutators in *S. cerevisiae* resulted from high levels of expression of 3-methyladenine DNA glycosylase relative to expression of the apurinic/apyrimidinic endonuclease, creating an imbalance between the first two enzymes involved in DNA base excision repair; this high mutation rate is not observed in the absence of the *Rev1/Rev3/Rev7*-catalyzed lesion bypass system (51). Yeast strains with a mutation that interferes with the exonucleolytic proofreading activity of polymerase delta or epsilon have an "antimutator" phenotype with respect to frameshift errors if they also lack MSH2 mismatch repair activity (58).

When expression of the SOS-inducible polymerase dinB was increased in *E. coli*, both frameshift and base substitution mutations increased, although not to the same extent (73). The ability to survive and to respond to genome damage depends upon expression and/or activation of proteins involved in the SOS response (79, 126), polymerases with unusual specificity (45), and repair proteins (23, 120). Induction of the SOS response increases the efficiency of global nucleotide excision repair of cyclobutane pyrimidine dimers (25), and alkylation sensitivity varies depending upon different mutations in repair pathways (84).

Some genomes, such as the radiation-resistant *Deinococcus radiodurans* (124), survive under conditions that seem to be inescapably mutagenic. Thermophiles would risk multiple mutations per gene per generation without mechanisms that repair and protect DNA (57). In fact, the mutation rate of 37°C genomes also would be high without mechanisms that repair "spontaneous" damage (78). From repair of apurinic sites to the removal of mismatches resulting from C tautomerization during replication and from deamination of C to U (43) [the rate of which is further increased opposite O6-alkylated guanines (41)], different levels of repair result in changes in base composition. For example, in certain mollicutes, lack of uracil deglycosylase is correlated with increased AT content (125).

Slow repair of deaminated Cs under stress might enable a "toe in the water" test of the effect of replacing Cs with Ts (13), because mRNA synthesized prior to repair incorporates an A rather than a G opposite the deaminated C. If this point mutation has a survival advantage, then organisms with that "damage" might divide prior to repair and thus incorporate an A into the newly synthesized strand of DNA.

The initiation and focusing of hypermutation of immunoglobulin variable regions by targeted deamination of Cs (32) suggests that repair and protection mechanisms may be captured, regulated, and focused to specific regions of the genome.

REVERSIBLE MUTATIONS AND A GENOME'S IMPLICIT RANGE

We describe one nucleotide sequence as an organism's genome and expect that progeny of this organism will inherit the same nucleotide sequence, except when mutation intervenes. Yet it is predictable, within population sizes of only thousands of bacteria, that certain mutations will occur. For example, tetranucleotide repeats increase and decrease in length as the new and old strands of DNA misalign during synthesis and/or repair (107, 117). Such mutations are not only predictable, but also reversible: Because these repeats continue to change in length, a parental type will reappear among the population of descendants (100). Therefore, rather than view a genome as encoding a specific repeat length that can mutate, we can view that genome as encoding a specific repeat length explicitly, but a range of repeat lengths implicitly, and consider that the range of lengths is an inherited phenotype of the genome.

Because changes in the lengths of repeats can change the strength of promoters or shift the reading frame of genes, each of the genes associated with these repeats will have a range of activities within a population descended from essentially any individual with any one combination of repeat lengths (6, 118). For example, in *Neisseria meningitidis*, individuals with spacers of 11, 10, or 9 Gs between the -35 and -10 consensus motifs in its promoter have high, medium, or no detectable levels of expression of *porA* (119). Changes in the length of a repeat also may change how sensitive a gene is to being regulated by specific molecules in the environment. For example, in *E. coli*, as a tract of Ts that begins eight nucleotides from the promoter -10 region is shortened from seven to three, pyrimidine-mediated regulation of uracil phosphoribosyltransferase expression is reduced and then becomes undetectable (21).

Because their tendency to change in length has quantitative effects, tandem repeats have been described as "tuning knobs" (71, 116), generating diversity that facilitates adaptation at multiple loci within a comparatively few generations. In *H. influenzae* and *N. meningitidis*, genes associated with tetranucleotide repeats, termed "contingency loci," are involved in LPS biosynthesis, adhesion, iron acquisition, restriction-modification systems and the evasion of host immunity (6).

The infection process is a dynamic one (89), during which the implicitly encoded variation facilitates adaptation to variations in the environment and access to different tissue sites (29, 87). For example, unencapsulated *N. meningitidis* invade epithelial cells (61), but encapsulated organisms are resistant to serum complement, facilitating systemic spread (121). Although most meningococci carried asymptomatically in the upper respiratory tract are unencapsulated, capsular forms predominated during an outbreak of meningococcal disease (61). In this study, the presence or absence of capsule, and thus the virulence, correlated with insertion or deletion of a C in a run of Cs within the coding region of the polysialyltransferase gene, causing premature termination of translation and then restoration of the reading frame from generation to generation. Thus, both encapsulated and

unencapsulated individuals generate a mixture of encapsulated and unencapsulated progeny.

The amount of combinatorial diversity available to some species through changes in the length of repeats is impressive; a survey of three strains of *Neisseria* spp. suggests nearly 100 candidate phase-variable genes (113). Other more complex "reversible" mutations include the "flip-flop" system, in which gene expression is turned on and off by inversion of a segment of DNA. [In addition to the "predictable" variants, this system also generates "out of the box" diversity through lower-probability recombination at diverse sites (3).] Other examples of inversion-mediated reversible phenotypic change include alterations in the sequence of a defined region of pilin, an "on/off" switch for fimbriae expression involving inversion of the promotor (60, 62) and a change in bacteriophage host specificity through in-frame inversions in the coding region of the tail fiber gene (106).

The range of genomes encoded implicitly through the many potential combinations of repeat lengths extends the range of environmental niches accessible to a population of descendants without committing all descendants to a sequence path that may be favored only by the circumstances of the moment. Progeny inherit multiple sequences, one explicitly and others implicitly.

ACCELERATED EXPLORATION

An "optimal" overall mutation rate is high enough to access the variation needed for long-term survival, but low enough to avoid being selected against by damaging mutations (31, 37). In one study, ∼1000 generations after loss of *MutS*, most lineages had reduced colony size, 4% had died out, and 55% had auxotrophic requirements, yet only 3% of the wild-type lineages had detectable mutations of any type (48). Under stress, however, bacteria that are less efficient at mismatch repair may be favored. Patients chronically infected with *Pseudomonas* have a high proportion of "mutator" bacteria (93), which can evolve resistance to subsequent antibiotic treatment at an increased rate (50).

To the extent that it is possible to increase, selectively, the rate of sampling of alternative implicit genomes under stress, this would expedite exploration while protecting "housekeeping" functions. In *N. meningitidis*, errors in mismatch repair are reported to increase the rate of mutation at contingency loci by an order of magnitude more than at other loci (98), and are more likely to be observed in invasive strains isolated during pandemics than in strains isolated from patients in years when there were no epidemics.

To the extent that the implicit genome encodes variations in levels of mismatch repair (MMR), a population of bacteria descendant from any one individual is composed of individuals with mismatch repair genes of varying levels of alacrity in repairing DNA. Repeats in MMR genes that facilitate the loss and regain of mutator activity, through recombination at a rate that is higher than the genome's background rate, could provide a selective advantage. Certain strains of *Pseudomonas* obtained from people with cystic fibrosis had a deletion between two repeats in

MutS and decreased mismatch repair activity (93). In *E. coli*, there is a relatively high number of neighboring repeats in genes such as *MutS* and *MutL* compared with random DNA sequences and with other *E. coli* genes (102). As would be expected if mismatch repair genes recombine at a comparatively high rate, they exhibit high sequence mosaicism derived from diverse phylogenetic lineages (14, 30). Although mismatch repair activity can be eliminated and then regained by recombination, it also might be "tuned" (suggested by E.P. Rocha, personal communication).

Laboratory-constructed deletions of *MutS* and *MutL* decrease the virulence of *Listeria monocytogenes* (85). In reviewing experimental data, it will become increasingly important to distinguish laboratory-constructed mutants from genes that are operating in the context in which they evolved. For example, it would be interesting to compare genes inactivated by frameshifts and deletions that are reversible through the use of any "intrinsic genome"–related mechanisms of the organism under investigation (for example, in which the repeats enable the hyper-recombination phenotype of MMR-deficient mutators to revert to a nonmutator phenotype through regain of MMR) with laboratory-constructed mutations that are not so readily reversible.

Decreased activity of the mismatch repair proteins MLH1 or MSH6 increases the rate of gene amplification in eukaryotic cells (20). [The extent of duplication and amplification of genes is affected by the action of a number of proteins; tandem duplications of the histidine operon on *S. typhimurium* are reduced by more than three orders of magnitude in $recA^-$ strains (2a).] Decreased MMR activity and/or induction of the SOS system allows increased recombination between divergent sequences (82). When the SOS response is induced in mismatch repair-deficient cells, *E. coli* accepts DNA from *S. enterica* (83). Following two rounds of selection for recombinants in an interspecies mating between *S. enterica* and *E. coli*, MMR^- cells represented 95% of the population (as defined by spontaneous mutation to rifampicin resistance and backed up by mapping a subset of mutations to *MutS* or *MutL*) (47).

Contingency loci also can affect the activity of restriction/modification systems (28), thus opening the door to "out of the box" diversity acquired from other organisms; the diversity of restriction/modification specificities in a population of bacteria enables varying levels of acceptance of distinct sequences of DNA. Bacteria may increase uptake (110) and release (115) of DNA under the influence of quorum sensing signals, and have evolved recognition signals within their DNA to facilitate uptake by conspecifics. Thus the extent of DNA uptake can be affected by the DNA sequence itself and by biochemical activities in both the donor and the recipient cell, all of which can fall under regulation and natural selection.

FINDING INFORMATION

Sequence alignment revealed the conserved nucleotides recognized by *B. burgdor-feri*'s coat-changing integrase (129). Statistical analysis was used to detect RGYW at sites that experience hypermutation in the immunoglobulin V regions (54).

However, as we examine whole genomes, encoded information will not always be obvious from examination of the sequence alone.

Identification of the recognition pattern for integron "59 base" elements has proven to be challenging. "59 base" elements are recognized by the site-specific recombinases that are responsible for inserting additional cassettes into the integron, and cassettes from the integron into genomic DNA. Recognition appears to involve a relationship between neighboring sequences in that mutations within one side of an imperfect inverted repeat can overcome mutations within the other side [see preliminary data cited in (22)].

We focus on DNA sequences as we write them, as strings of the letters A,T,G, and C; however, completely different sequences of bases can create a three-dimensional pattern that proteins recognize as similar and that encodes novel information. A simple example is provided by AT and GC pairs, both of which "fit" across the same width of double helix. Polymerase, editing, and mismatch repair proteins accept either an AT or a GC pair, but not other combinations.

Examples that have begun to look beyond linear base sequence analysis include the recognition of a palindromic major groove H-bond donor acceptor pattern, which appears to define favored sites of integration of the P transposon in *Drosophila melanogaster* (77), and the use of physical properties of DNA sequences in a probabilistic model used to recognize promoters (92). In the future we should be able to "read" DNA more as proteins do; to calculate and represent, in a comprehensible way, the breathing, tilt, and propeller-like twist of the base pairs along each DNA sequence; and to assess the extent to which the unique physical and chemical properties presented by different sequences affect the rate, nature, and location of genetic change.

DIRECTED MUTATION?

Distinct from the question as to whether mutations that tend to be favorable can become more probable is assessing whether mutation might be directed to specific sites that are particularly relevant to an environmental challenge faced at that moment (96, 103).

That the mutation rate in contingency and mismatch repair genes can be selectively increased sails close to stating that mutation can be targeted to the specific needs of the organism. Similarly, recombination that turns on *E. coli* fimbriae synthesis becomes more likely at the host's body temperature (60). However, these examples represent the cloud of "implicit" genomes and are not examples of mechanisms that would create an unprecedented novel base change that is targeted to overcome a specific stress.

Much as the control of virulence system in group A *Streptococcus* enables rapid changes in expression of gene products in diverse functional categories that interact with the host (55), a genetic "engineer" can learn how to use different stress sensors to target diversity generators, such as C deamination, to specific genes under appropriate environmental cues. But to what extent has evolution done this?

Acceptance of the concept that mutation can be targeted to specific biochemical pathways requires the demonstration of biochemical mechanisms that enable this to take place. With whole genomes before us, we can investigate the "wiring" to assess whether and to what extent the activity of specific polymerases, endonucleases, and repair proteins are altered under particular stresses and, to the extent we learn to recognize this, targeted to specific locations.

Discussions of potential targeting mechanisms have centered on expression of specific transcription factors that might block or allow access to specific gene regions and alter the location of mutation hotspots (44, 74), much as the germline transcript in the immune system directs the class switch. "Starvation-induced derepression" has been proposed to result in transcription-guided genome changes in bacteria (127). In yeast the nutritional status affects the pattern of transcription factor expression and alters sites of meiotic genetic variation (1). Transposon-carried regulatory DNA sequences may land next to physically separated genes that are expressed together, as the transposons "jump into" promoter regions with more open conformations (111). Given the new roles that we are just beginning to learn for RNA (24), a letter recently has been published (2) speculating that RNA could carry information back to the germline genome of eukaryotes.

It is interesting to explore the extent to which natural selection might have connected specific environmental stresses to certain classes of mutations. To facilitate careful discussion, it is important to note that this review has in general not examined the concept of focused mutation as used in that dynamic sense in which mutations would, for example, be targeted to revert auxotrophy. This review focuses on the ability of natural selection to alter the genome's mutation phenotype in ways that make classes of potentially adaptive (such as changes in the lengths of tetranucleotide repeats) or damaging mutations intrinsically more or less likely compared with random mutation. Even when mechanisms that can target genetic variation to a metabolically appropriate pathway are identified, that does not mean that the organism knows to change a specific A to a G to get the desired phenotypic effect. For V region hypermutation, focused mutation generates focused variation, upon which selection then acts.

SUMMARY: A NEW EVOLUTIONARY SYNTHESIS

The mutation phenotype of a genome represents an evolved balance between a myriad of biochemical activities, from nucleotide synthesis to the relative expression and selectivity of polymerases, proofreading, and repair in a sequence-context-dependent manner.

The number of distinct ways any genome might mutate is so vast that any fortuitous alignment of the tendency to make a certain type of mutation and the potential of that type of mutation to be biologically "useful" would be preserved through repeated cycles of pressure and survival.

Some genomes have evolved information encoding what I have termed here an implicit genome, which gives their progeny, taken as a group, predictable access to

a combinatorial assortment of variations in gene regions, such as host interaction surfaces, in which diversity is particularly important for survival. There is a selective value to the generation of diversity itself. Beyond the reach of the implicit genome, genomes can access additional, intact, information through recombination and horizontal transfer.

The concept of a mutation phenotype certainly does not imply that all mutations are targeted and helpful. However, it does suggest that when we observe different rates of mutation at different positions in a gene, we should consider the possibility this may be due to evolution of mechanisms that modulate the rate of variation, rather than selection for and against mutations one by one. Unless we look for such strategic information, we are unlikely to discover it, even if it is there.

Evolutionary theory has described variation as resulting from genetic changes that are forever random, with selection acting on the results of this random genetic variation. Because genomes do not inhabit a completely random world, genomes can evolve to be increasingly favored by repeated cycles of selection. The ability to handle predictable, repeated challenges is in fact a major challenge of evolution. Perhaps the most important factor in genome evolution will prove to be that the varied mechanisms that diversify and stabilize a genome themselves feel the pressure of natural selection.

ACKNOWLEDGMENTS

The *Annual Review of Microbiology* is online at http://micro.annualreviews.org

LITERATURE CITED

1. Abdullah MF, Borts RH. 2001. Meiotic recombination frequencies are affected by nutritional states in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 98: 14524–29

2. Adler MJ. 2001. A novel mechanism for evolution? *Science* 294:53–54

2a. Anderson RP, Roth JR. 1978. Tandem genetic duplications in *Salmonella typhimurium*: amplification of the histidine operon. *J. Mol. Biol.* 126:53–71

3. Arber W. 1999. Involvement of gene products in bacterial evolution. *Ann. NY Acad. Sci.* 870:36–44

4. Barbour AG, Restrepo BI. 2000. Antigenic variation in vector-borne pathogens. *Emerg. Infect. Dis.* 6:449–57

5. Barry JD, McCulloch R. 2001. Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Adv. Parasitol.* 49:1–70

6. Bayliss CD, Field D, Moxon ER. 2001. The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*. *J. Clin. Invest.* 107:657–62

7. Bayliss CD, van de Ven T, Moxon ER. 2002. Mutations in *polI* but not *mutSLH* destabilize *Haemophilus influenzae* tetranucleotide repeats. *EMBO J.* 21: 1465–76

8. Berdal KG, Johansen RF, Seeberg E. 1998. Release of normal bases from intact DNA by a native DNA repair enzyme. *EMBO J.* 17:363–67

9. Beres SB, Sylva GL, Barbian KD, Lei B, Hoff JS, et al. 2002. Genome sequence of a serotype M3 strain of group A *Streptococcus*: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc. Natl. Acad. Sci. USA* 99: 10078–83

10. Blumental-Perry A, Zenvirth D, Klein S, Onn I, Simchen G. 2000. DNA motif associated with meiotic double-strand break regions in *Saccharomyces cerevisiae*. *EMBO Rep.* 1:232–38

11. Boyd EF, Hartl DL. 1997. Recent horizontal transmission of plasmids between natural populations of *Escherichia coli* and *Salmonella enterica*. *J. Bacteriol.* 179: 1622–27

12. Bresson A, Fuchs RP. 2002. Lesion bypass in yeast cells: Pol eta participates in a multi-DNA polymerase process. *EMBO J.* 21:3881–87

13. Bridges BA. 1999. Dirty transcripts from clean DNA. *Science* 284:62–63

14. Brown EW, LeClerc JE, Li B, Payne WL, Cebula TA. 2001. Phylogenetic evidence for horizontal transfer of mutS alleles among naturally occurring *Escherichia coli* strains. *J. Bacteriol.* 183:1631–44

15. Caporale LH. 1984. Is there a higher level genetic code that directs evolution? *Mol. Cell Biochem.* 64:5–13

16. Caporale LH. 1999. *Molecular Strategies in Biological Evolution*. New York: NY Acad. Sci.

17. Caporale LH. 2000. Mutation is modulated: implications for evolution. *BioEssays* 22:388–95

18. Caporale LH. 2002. *Darwin in the Genome*. New York: McGraw-Hill

19. Casjens S, Palmer N, van Vugt R, Huang WM, Stevenson B, et al. 2000. A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. *Mol. Microbiol.* 35:490–516

20. Chen S, Bigner SH, Modrich P. 2001. High rate of CAD gene amplification in human cells deficient in MLH1 or MSH6. *Proc. Natl. Acad. Sci. USA* 98:13802–7

21. Cheng Y, Dylla SM, Turnbough CL Jr. 2001. A long T.A tract in the *upp* initially transcribed region is required for regulation of *upp* expression by UTP-dependent reiterative transcription in *Escherichia coli*. *J. Bacteriol.* 183:221–28

22. Collis CM, Kim MJ, Partridge SR, Stokes HW, Hall RM. 2002. Characterization of the class 3 integron and the site-specific recombination system it determines. *J. Bacteriol.* 184:3017–26

23. Courcelle J, Hanawalt PC. 2001. Participation of recombination proteins in rescue of arrested replication forks in UV-irradiated *Escherichia coli* need not involve recombination. *Proc. Natl. Acad. Sci. USA* 98:8196–202

24. Couzin J. 2002. Breakthrough of the year. Small RNAs make big splash. *Science* 298:2296–97

25. Crowley DJ, Hanawalt PC. 1998. Induction of the SOS response increases the efficiency of global nucleotide excision repair of cyclobutane pyrimidine dimers, but not 6–4 photoproducts, in UV-irradiated *Escherichia coli*. *J. Bacteriol.* 180:3345–52

26. Darwin C. 1964 (1859). *On the Origin of Species*. Cambridge, MA: Harvard Univ. Press

27. Davis L, Smith GR. 2001. Meiotic recombination and chromosome segregation in *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci. USA* 98:8395–402

28. De Bolle X, Bayliss CD, Field D, van de Ven T, Saunders NJ, et al. 2000. The length of a tetranucleotide repeat tract in *Haemophilus influenzae* determines the phase variation rate of a gene with homology to type III DNA methyltransferases. *Mol. Microbiol.* 35:211–22

29. Deitsch KW, Moxon ER, Wellems TE. 1997. Shared themes of antigenic

variation and virulence in bacterial, pro-
tozoal, and fungal infections. *Microbiol.
Mol. Biol. Rev.* 61:281–93

30. Denamur E, Lecointre G, Darlu P, Tenail-
lon O, Acquaviva C, et al. 2000. Evolu-
tionary implications of the frequent hori-
zontal transfer of mismatch repair genes.
Cell 103:711–21

31. deVisser JAGM. 2002. The fate of micro-
bial mutators. *Microbiology* 148:1247–52

32. Di Noia J, Neuberger MS. 2002. Altering
the pathway of immunoglobulin hyper-
mutation by inhibiting uracil-DNA glyco-
sylase. *Nature* 419:43–48

33. Diaz RL, Alcid AD, Berger JM, Keeney
S. 2002. Identification of residues in yeast
Spo11p critical for meiotic DNA double-
strand break formation. *Mol. Cell Biol.*
22:1106–15

34. Dickinson WJ, Seger J. 1999. Cause and
effect in evolution. *Nature* 399:30

35. Downie JA, Young JP. 2001. Genome se-
quencing. The ABC of symbiosis. *Nature*
412:597–98

36. Dunham MJ, Badrane H, Ferea T, Adams
J, Brown PO, et al. 2002. Characteristic
genome rearrangements in experimental
evolution of *Saccharomyces cerevisiae.*
Proc. Natl. Acad. Sci. USA 99:16144–49

37. Eigen M. 2002. Error catastrophe and an-
tiviral strategy. *Proc. Natl. Acad. Sci. USA*
99:13374–76

37a. Evolution Since Darwin (PBS). http://
www.pbs.org/wgbh/evolution/library/06/
index.html

37b. Evolution Since Darwin (PBS). http://
www.wku.edu/%7Esmithch/index1.htm

38. Faili A, Aoufouchi S, Flatter E, Gueranger
Q, Reynaud CA, Weill JC. 2002. In-
duction of somatic hypermutation in im-
munoglobulin genes is dependent on DNA
polymerase iota. *Nature* 419:944–47

39. Finlay BB, Falkow S. 1997. Common
themes in microbial pathogenicity revis-
ited. *Microbiol. Mol. Biol. Rev.* 61:136–69

40. Fix DF, Glickman BW. 1987. Asym-
metric cytosine deamination revealed by
spontaneous mutational specificity in an

Ung-strain of *Escherichia coli. Mol. Gen.
Genet.* 209:78–82

41. Fix DF, Koehler DR, Glickman BW. 1990.
Uracil-DNA glycosylase activity affects
the mutagenicity of ethyl methanesul-
fonate: evidence for an alternative path-
way of alkylation mutagenesis. *Mutat.
Res.* 244:115–21

42. Flores-Rozas H, Kolodner RD. 1998. The
Saccharomyces cerevisiae MLH3 gene
functions in MSH3-dependent suppres-
sion of frameshift mutations. *Proc. Natl.
Acad. Sci. USA* 95:12404–9

43. Fondufe-Mittendorf YN, Harer C,
Kramer W, Fritz HJ. 2002. Two amino
acid replacements change the substrate
preference of DNA mismatch glycosy-
lase Mig.MthI from T/G to A/G. *Nucleic
Acids Res.* 30:614–21

44. Fox ME, Yamada T, Ohta K, Smith
GR. 2000. A family of cAMP-response-
element-related DNA sequences with
meiotic recombination hotspot activity
in *Schizosaccharomyces pombe. Genetics*
156:59–68

45. Friedberg EC, Wagner R, Radman M.
2002. Specialized DNA polymerases, cel-
lular survival, and the genesis of muta-
tions. *Science* 296:1627–30

46. Frosina G. 2000. Overexpression of en-
zymes that repair endogenous damage to
DNA. *Eur. J. Biochem.* 267:2135–49

47. Funchain P, Yeung A, Stewart J, Clen-
denin WM, Miller JH. 2001. Amplifica-
tion of mutator cells in a population as a
result of horizontal transfer. *J. Bacteriol.*
183:3737–41

48. Funchain P, Yeung A, Stewart JL, Lin R,
Slupska MM, Miller JH. 2000. The con-
sequences of growth of a mutator strain
of *Escherichia coli* as measured by loss
of function among multiple gene targets
and loss of fitness. *Genetics* 154:959–70

49. Futuyma DJ. 1998. *Evolutionary Biology.*
Sunderland, MA: Sinauer. 3rd ed.

50. Giraud A, Matic I, Radman M, Fons M,
Taddei F. 2002. Mutator bacteria as a risk
factor in treatment of infectious diseases.

Antimicrob. Agents Chemother. 46:863–65

51. Glassner BJ, Rasmussen LJ, Najarian MT, Posnick LM, Samson LD. 1998. Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proc. Natl. Acad. Sci. USA* 95:9997–10002

52. Goodman MF. 2002. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* 71:17–50

53. Goodman MF, Fygenson KD. 1998. DNA polymerase fidelity: from genetics toward a biochemical understanding. *Genetics* 148:1475–82

54. Goyenechea B, Milstein C. 1996. Modifying the sequence of an immunoglobulin V-gene alters the resulting pattern of hypermutation. *Proc. Natl. Acad. Sci. USA* 93:13979–84

55. Graham MR, Smoot LM, Lux Migliaccio CA, Virtaneva K, Sturdevant DE, et al. 2002. Virulence control in group A *Streptococcus* by a two-component gene regulatory system: global expression profiling and in vivo infection modeling. *Proc. Natl. Acad. Sci. USA* 99:13855–60

56. Greene CN, Jinks-Robertson S. 1997. Frameshift intermediates in homopolymer runs are removed efficiently by yeast mismatch repair proteins. *Mol. Cell Biol.* 17:2844–50

57. Grogan DW. 1998. Hyperthermophiles and the problem of DNA instability. *Mol. Microbiol.* 28:1043–49

58. Hadjimarcou MI, Kokoska RJ, Petes TD, Reha-Krantz LJ. 2001. Identification of a mutant DNA polymerase delta in *Saccharomyces cerevisiae* with an antimutator phenotype for frameshift mutations. *Genetics* 158:177–86

59. Hall RM, Collis CM. 1995. Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. *Mol. Microbiol.* 15:593–600

60. Hallet B. 2001. Playing Dr Jekyll and Mr Hyde: combined mechanisms of phase variation in bacteria. *Curr. Opin. Microbiol.* 4:570–81

61. Hammerschmidt S, Muller A, Sillmann H, Muhlenhoff M, Borrow R, et al. 1996. Capsule phase variation in *Neisseria meningitidis* serogroup B by slipped-strand mispairing in the polysialyltransferase gene (*siaD*): correlation with bacterial invasion and the outbreak of meningococcal disease. *Mol. Microbiol.* 20:1211–20

62. Henderson IR, Owen P, Nataro JP. 1999. Molecular switches—the ON and OFF of bacterial phase variation. *Mol. Microbiol.* 33:919–32

63. Honjo T, Kinoshita K, Muramatsu M. 2002. Molecular mechanism of class switch recombination: linkage with somatic hypermutation. *Annu. Rev. Immunol.* 20:165–96

64. Deleted in proof

65. Deleted in proof

66. Huber KE, Waldor MK. 2002. Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* 417:656–59

67. Hueck CJ. 1998. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* 62:379–433

68. Hunter N, Kleckner N. 2001. The single-end invasion: an asymmetric intermediate at the double-strand break to double-Holliday junction transition of meiotic recombination. *Cell* 106:59–70

69. Jager J, Pata JD. 1999. Getting a grip: polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* 9:21–28

70. Jeffreys AJ, Kauppi L, Neumann R. 2001. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat. Genet.* 29:217–22

71. Kashi YD, King D, Soller M. 1997. Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.* 13:74–78

72. Kim DR, Park SJ, Oettinger MA. 2000. V(D)J recombination: site-specific cleavage and repair. *Mol. Cell* 10:367–74

73. Kim SR, Maenhaut-Michel G, Yamada M, Yamamoto Y, Matsui K, et al. 1997. Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: An overexpression of dinB/dinP results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc. Natl. Acad. Sci. USA* 94: 13792–97

74. Kirkpatrick DT, Fan Q, Petes TD. 1999. Maximal stimulation of meiotic recombination by a yeast transcription factor requires the transcription activation domain and a DNA-binding domain. *Genetics* 152:101–15

75. Lawrence JG, Ochman H. 1998. Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. USA* 95: 9413–17

76. Levin BR, Bergstrom CT. 2000. Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. *Proc. Natl. Acad. Sci. USA* 97:6981–85

77. Liao GC, Rehm EJ, Rubin GM. 2000. Insertion site preferences of the P transposable element in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 97:3347–51

78. Lindahl T, Wood RD. 1999. Quality control by DNA repair. *Science* 286:1897–905

79. Luo Y, Pfuetzner RA, Mosimann S, Paetzel M, Frey EA, et al. 2001. Crystal structure of LexA: a conformational switch for regulation of self-cleavage. *Cell* 106:585–94

80. Mao CS, Stavnezer J. 2001. Differential regulation of mouse germline Ig gamma 1 and epsilon promoters by IL-4 and CD40. *J. Immunol.* 167:1522–34

81. Marra G, Schar P. 1999. Recognition of DNA alterations by the mismatch repair system. *Biochem. J.* 338:1–13

82. Matic I, Rayssiguier C, Radman M. 1995. Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. *Cell* 80:507–15

83. Matic I, Taddei F, Radman M. 2000. No genetic barriers between *Salmonella enterica* serovar Typhimurium and *Escherichia coli* in SOS-induced mismatch repair-deficient cells. *J. Bacteriol.* 182:5922–24

84. Memisoglu A, Samson L. 2000. Contribution of base excision repair, nucleotide excision repair, and DNA recombination to alkylation resistance of the fission yeast *Schizosaccharomyces pombe*. *J. Bacteriol.* 182:2104–12

85. Merino D, Reglier-Poupet H, Berche P, Charbit A. 2002. A hypermutator phenotype attenuates the virulence of *Listeria monocytogenes* in a mouse model. *Mol. Microbiol.* 44:877–87

86. Miller JH, Funchain P, Clendenin W, Huang T, Nguyen A, et al. 2002. *Escherichia coli* strains (*ndk*) lacking nucleoside diphosphate kinase are powerful mutators for base substitutions and frameshifts in mismatch-repair-deficient strains. *Genetics* 162:5–13

87. Moxon ER, Rainey PB, Nowak MA, Lenski RE. 1994. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* 4:24–33

88. Moxon ER, Thaler DS. 1997. Microbial genetics. The tinkerer's evolving toolbox. *Nature* 387:659, 661–62

89. Nassif X, Pujol C, Morand P, Eugene E. 1999. Interactions of pathogenic *Neisseria* with host cells. Is it possible to assemble the puzzle? *Mol. Microbiol.* 32:1124–32

90. Nilsen H, Krokan HE. 2001. Base excision repair in a network of defence and tolerance. *Carcinogenesis* 22:987–98

91. Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299–304

92. Ohler U, Niemann H, Liao G, Rubin GM. 2001. Joint modeling of DNA sequence and physical properties to improve

eukaryotic promoter recognition. *Bioinformatics* 17:S199–206

93. Oliver A, Baquero F, Blazquez J. 2002. The mismatch repair system (mutS, mutL and uvrD genes) in *Pseudomonas aeruginosa*: molecular characterization of naturally occurring mutants. *Mol. Microbiol.* 43:1641–50

94. Packer MJ, Hunter CA. 2001. Sequence-structure relationships in DNA oligomers: a computational approach. *J. Am. Chem. Soc.* 123:7399–406

95. Perna NT, Plunkett G 3rd, Burland V, Mau B, Glasner JD, et al. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409:529–33

96. Poteete AR, Wang HR, Foster PL. 2002. Phage lambda red-mediated adaptive mutation. *J. Bacteriol.* 184:3753–55

97. Rich A. 1983. Right-handed and left-handed DNA: conformational information in genetic material. *Cold Spring Harbor Symp. Quant. Biol.* 47:1–12

98. Richardson AR, Yu Z, Popovic T, Stojiljkovic I. 2002. Mutator clones of *Neisseria meningitidis* in epidemic serogroup A disease. *Proc. Natl. Acad. Sci. USA* 99:6103–7

99. Ripley LS. 1999. Predictability of mutant sequences. Relationships between mutational mechanisms and mutant specificity. *Ann. NY Acad. Sci.* 870:159–72

100. Robertson BD, Meyer TF. 1992. Genetic variation in pathogenic bacteria. *Trends Genet.* 8:422–27

101. Rocha EP, Danchin A. 2001. Ongoing evolution of strand composition in bacterial genomes. *Mol. Biol. Evol.* 18:1789–99

102. Rocha EP, Matic I, Taddei F. 2002. Overrepresentation of repeats in stress response genes: a strategy to increase versatility under stressful conditions? *Nucleic Acids Res.* 30:1886–94

103. Rosenberg SM. 2001. Evolving responsively: adaptive mutation. *Nat. Rev. Genet.* 2:504–15

104. Rowe-Magnus DA, Mazel D. 2001. Integrons: natural tools for bacterial genome evolution. *Curr. Opin. Microbiol.* 4:565–69

105. Sandmeier H. 1994. Acquisition and rearrangement of sequence motifs in the evolution of bacteriophage tail fibres. *Mol. Microbiol.* 12:343–50

106. Sandmeier H, Iida S, Arber W. 1992. DNA inversion regions Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. *J. Bacteriol.* 174:3936–44

107. Saunders NJ, Jeffries AC, Peden JF, Hood DW, Tettelin H, et al. 2000. Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. *Mol. Microbiol.* 37:207–15

108. Scheuermann RH, Echols H. 1984. A separate editing exonuclease for DNA replication: the epsilon subunit of *Escherichia coli* DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. USA* 81:7747–51

109. Schwan TG, Burgdorfer W, Garon CF. 1988. Changes in infectivity and plasmid profile of the Lyme disease spirochete, *Borrelia burgdorferi*, as a result of in vitro cultivation. *Infect. Immun.* 56:1831–36

110. Shapiro JA. 1998. Thinking about bacterial populations as multicellular organisms. *Annu. Rev. Microbiol.* 52:81–104

111. Shapiro JA. 1999. Transposable elements as the key to a 21st century view of evolution. *Genetica* 107:171–79

112. Sinden RR, Hashem VI, Rosche WA. 1999. DNA-directed mutations. Leading and lagging strand specificity. *Ann. NY Acad. Sci.* 870:173–89

113. Snyder LA, Butcher SA, Saunders NJ. 2001. Comparative whole-genome analyses reveal over 100 putative phase-variable genes in the pathogenic *Neisseria* spp. *Microbiology* 147:2321–32

114. Sonti RV, Roth JR. 1989. Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. *Genetics* 123:19–28

115. Steinmoen H, Knutsen E, Havarstein LS. 2002. Induction of natural competence in Streptococcus pneumoniae triggers lysis and DNA release from a subfraction of the cell population. *Proc. Natl. Acad. Sci. USA* 99:7681–86

116. Trifonov E. 1999. Elucidating sequence codes: three codes for evolution. In *Molecular Strategies in Biological Evolution*, ed. LH Caporale, pp. 330–38. New York: Ann. NY Acad. Sci.

117. van Belkum A. 1999. Short sequence repeats in microbial pathogenesis and evolution. *Cell Mol. Life Sci.* 56:729–34

118. van Belkum A, Scherer S, van Alphen L, Verbrugh H. 1998. Short-sequence DNA repeats in prokaryotic genomes. *Microbiol. Mol. Biol. Rev.* 62:275–93

119. van der Ende A, Hopman CT, Zaat S, Essink BB, Berkhout B, Dankert J. 1995. Variable expression of class 1 outer membrane protein in *Neisseria meningitidis* is caused by variation in the spacing between the −10 and −35 regions of the promoter. *J. Bacteriol.* 177:2475–80

120. Van Houten B, Eisen JA, Hanawalt PC. 2002. A cut above: discovery of an alternative excision repair pathway in bacteria. *Proc. Natl. Acad. Sci. USA* 99:2581–83

121. Vogel U, Frosch M. 1999. Mechanisms of neisserial serum resistance. *Mol. Microbiol.* 32:1133–39

122. Wahls WP, Smith GR. 1994. A heteromeric protein that binds to a meiotic homologous recombination hot spot: correlation of binding and hot spot activity. *Genes Dev.* 8:1693–702

123. Wallace AR. 1905. *My Life, My Life: A Record of Events and Opinions, Vols. 1, 2.* New York: Dodd, Mead

124. White O, Eisen JA, Heidelberg JF, Hickey EK, Peterson JD, et al. 1999. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. *Science* 286:1571–77

125. Williams MV, Pollack JD. 1990. A mollicute (mycoplasma) DNA repair enzyme: purification and characterization of uracil-DNA glycosylase. *J. Bacteriol.* 172:2979–85

126. Witkin EM. 1976. Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol. Rev.* 40:869–907

127. Wright BE. 2000. A biochemical mechanism for nonrandom mutations and evolution. *J. Bacteriol.* 182:2993–3001

128. Zhang JR, Hardham JM, Barbour AG, Norris SJ. 1997. Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. *Cell* 89:275–85

129. Zhang JR, Norris SJ. 1998. Genetic variation of the *Borrelia burgdorferi* gene *vlsE* involves cassette-specific, segmental gene conversion. *Infect. Immun.* 66:3698–704

130. Zhang X, Mathews CK. 1995. Natural DNA precursor pool asymmetry and base sequence context as determinants of replication fidelity. *J. Biol. Chem.* 270:8401–4