1-1-2016

Hypermutability in Asexuals: Investigating the Effects of Deleterious Mutations

Tanya Singh
University of Pennsylvania, tanyasingh23@gmail.com

Hypermutability in Asexuals: Investigating the Effects of Deleterious Mutations

Abstract
Mutation is the ultimate source of the genetic variation—including genetic variation for mutation rate itself—that fuels evolution. Selection to increase the genomic mutation rate, driven by selective sweeps of beneficial mutations, can be strong and rapid where genetic linkage is present, as evidenced by numerous observations in experimental microbial populations. Selection to decrease the mutation rate, in contrast, is expected to depend on avoidance of mutational load and act over a longer time scale. In keeping with this latter expectation, there have been relatively few experimental observations of the evolution of reduced genomic mutation rates. Here, I report the rapid evolution of reduced mutation rates in hypermutable E. coli populations propagated at extremely small effective size—a circumstance under which selection is generally minimized. I hypothesize that high deleterious mutation pressure can strengthen indirect selection favoring lower mutation rates in these populations, and find both phenotypic and genotypic evidence to support this hypothesis. Additionally, I use simulations to analyze the effect of high deleterious mutation pressure on nascent neutral lineages that arise in an expanding asexual population and find that the spread of these lineages can be impaired. I discuss these results in the light of fates of novel mutations and point to future work that will involve studying the fates of adaptive mutation under high deleterious mutation pressure.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Biology

First Advisor
Paul D. Sniegowski

Keywords
Deleterious Mutation, Evolutionary Genetics, Mutation, Mutation Accumulation, Mutation Rate Evolution, Neutral Mutation

Subject Categories
Biology | Microbiology

HYPERMUTABILITY IN ASEXUALS: INVESTIGATING THE EFFECTS OF DELETERIOUS MUTATIONS

Tanya Singh

A DISSERTATION

in

Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

_____

Dr. Paul D Sniegowski

Professor of Biology

Graduate Group Chairperson

_____

Dr. Michael Lampson

Associate Professor of Biology

Dissertation Committee

Dr. Timothy Linksvayer, Assistant Professor of Biology

Dr. Paul Schmidt, Associate Professor of Biology

Dr. Joshua Plotkin, Professor of Biology, Mathematics, and Computer and Information Science

Dr. Dustin Brisson, Associate Professor of Biology

Dr Paul Sniegowski, Professor of Biology

HYPERMUTABILITY IN ASEXUALS: INVESTIGATING THE EFFECTS OF DELETERIOUS MUTATIONS

ABTRACT

HYPERMUTABILITY IN ASEXUALS: INVESTIGATING THE EFFECTS OF DELETERIOUS MUTATIONS

Tanya Singh

Paul Sniegowski

Mutation is the ultimate source of the genetic variation—including genetic variation for mutation rate itself—that fuels evolution. Selection to increase the genomic mutation rate, driven by selective sweeps of beneficial mutations, can be strong and rapid where genetic linkage is present, as evidenced by numerous observations in experimental microbial populations. Selection to decrease the mutation rate, in contrast, is expected to depend on avoidance of mutational load and act over a longer time scale. In keeping with this latter expectation, there have been relatively few experimental observations of the evolution of reduced genomic mutation rates. Here, I report the rapid evolution of reduced mutation rates in hypermutable *E. coli* populations propagated at extremely small effective size—a circumstance under which selection is generally minimized. I hypothesize that high deleterious mutation pressure can strengthen indirect selection favoring lower mutation rates in these populations, and find both phenotypic and genotypic evidence to support this hypothesis. Additionally, I use simulations to analyze the effect of high deleterious mutation pressure on nascent neutral lineages that arise in an expanding asexual population and find that the spread of these lineages can be impaired. I discuss these results in the light of fates of novel mutations and point to future work that will involve studying the fates of adaptive mutation under high deleterious mutation pressure.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1

Introduction

The Nature of Mutations

Genetic mutations occur and spread in a population when haphazard genomic changes are carried over to the next generation. Mutations include, but are not limited to the substitution of an incorrect nucleotide, the insertion or deletion of nucleotides at a site, the movement of mobile genetic elements such as transposable elements (McClintock 1938), and even the deletion or duplication of entire genes or genomic regions. Although there are many ways in which mutations occur at the molecular level, from a phenotypic standpoint, mutations can be divided into three categories: neutral, deleterious and beneficial. Of these, mutations that are detrimental to fitness are likely to be much more common than mutations that enhance fitness (Muller 1928).

Deleterious mutations are more common than beneficial mutations because a random alteration or disruption of a functional genomic sequence is more likely to be detrimental than advantageous. It is helpful to consider an analogy presented by RA Fisher who compares the state of adaptation of an organism, which is the product of many eons of evolution by natural selection, to a microscope that is tuned at a good degree of focus, but not a perfect degree of focus (Fisher 1930:

40). Any additional tuning of moderate or large effect is more likely to make the focus worse, and only very few fine changes can be thought to improve the focus. From this simple analogy, it becomes quite clear that most mutations are expected to be detrimental rather than beneficial.

It also follows from the previous discussion that beneficial mutations are expected to be rare. The ‘focus’ of the microscope can be translated as the fitness or the degree to which a certain individual is adapted to its environment, but it also represents the developmental plan of an organism which might preclude acquiring certain phenotypes, rendering some mutations that might otherwise be beneficial to be deleterious (Maynard Smith et al. 1985). Hence, in order to be beneficial, mutations need to fit in with the existing phenotypic and developmental constraints of the organism. In addition, whether a mutation is beneficial is highly dependent on the environment, which conspires to increase the rarity of beneficial mutations. The effect of deleterious mutations, in contrast, is less environmentally determined, and hence most deleterious mutations are expected to be detrimental regardless of the environment.

Evolution of the Genomic Mutation Rate

The genomic mutation rate is known to be an evolutionarily pliable trait (Michaels, Cruz, and Miller 1990; Hong et al. 2005; Sniegowski et al. 2012; Thompson, Desai, and Murray 2006), and multiple loci affecting mutation rates

have been implicated by research in a variety of organisms (Cox, Degnen, and Scheppe 1972; Michaels, Cruz, and Miller 1990; Painter 1975; Mansky and Cunningham 2000). Mutation rate modifiers that elevate the genomic mutation rate are referred to as mutators, and modifiers that reduce mutation rate (by increasing, for example DNA polymerase fidelity) are called antimutators. With the realization that a large number of loci are implicated in maintaining the mutation rate because of the need to repair DNA damage and maintain DNA replication fidelity, the prospect of selection and other evolutionary forces acting to change the frequencies of alleles at such loci arises. Indeed, changes in frequencies of mutation rate modifier alleles have been inferred or directly observed both in natural (LeClerc et al. 1996; Giraud et al. 2001; Hermisson et al. 2002; Richardson et al. 2002) and experimental populations (Chao and Cox 1983; Cox, Cox et al 1972; Mao et al. 1997; Sniegowski, Gerrish, and Lenski 1997a; Giraud et al. 2001; Notley-McRobb and Ferenci, 2000 ; Shaver et al. 2002; Thompson, Desai, and Murray 2006; Wielgoss et al. 2012; McDonald et al. 2012).

If mutation rate modifiers do not themselves affect fitness, then selection can only act indirectly to change their frequencies via linkage with fitness-affecting mutations (Gentile et al. 2011; Raynes and Sniegowski 2014). In asexual populations, for example, mutator alleles can go to fixation via the process of genetic hitchhiking, i.e. by virtue of being linked to a single or multiple

advantageous mutations (Sniegowski, Gerrish, and Lenski 1997b; Shaver et al. 2002). Such hitchhiking occurs because a mutator allele is more likely to become associated with beneficial mutations, and hence increase in frequency in the population. In the absence of genetic recombination, it is almost impossible to dissociate the mutator allele and the beneficial mutation, thereby ensuring that the mutator allele goes to fixation.

Recent theoretical studies have predicted that the evolution of mutation rates in asexual populations that are undergoing adaptive evolution will be upwardly biased because of recurrent hitchhiking of different mutator alleles with beneficial mutations (Gentile et al. 2011; Andre and Godelle 2006). One study even predicts that this bias toward the evolution of a higher genomic mutation rate should culminate in extinction of an asexual population as fitness ultimately crashes under the influence of deleterious mutations (Gerrish et al. 2007). Investigation of the dynamics of neutral, beneficial, and deleterious mutations--and of potential mutation rate evolution itself--under such hypermutable circumstances is a fundamental aspect of my dissertation work.

Effects of Deleterious Mutations

Deleterious mutations are purged from populations via purifying selection; copies of other mutations--good, bad, or neutral--that are linked to a particular copy of a deleterious mutation are also removed from the population when the deleterious

mutation copy is purged. Numerous experimental and theoretical studies have been conducted on the influence of deleterious mutations on adaptive evolution in asexual populations (Johnson 2000; Andersson and Hughes 1996; Howe and Denver 2008; Balick et al. 2012) and the influence of deleterious mutations on linked genomic regions in sexual populations (Fisher 1930; Hill and Robertson 1966; Peck 1994; Charlesworth, Morgan, and Charlesworth 1993). One strong focus has been on the influence of deleterious mutations on the fate of linked beneficial mutations. Haldane (Haldane 1927), showed that the fixation probability of a beneficial mutation that arises as a single copy is approximately twice the selection coefficient, *s*, in its favor (for small *s*). Haldane's approach, however, assumed that no other mutations--beneficial or deleterious--affect the fate of the beneficial mutation. A variety of studies have shown through analytical or simulation approaches that the fixation probability of a beneficial mutation can be considerably reduced below Haldane's expectation by the influence of linked deleterious mutations (Bachtrog and Gordo 2004; Peck 1994). Intuitively, one would predict that this effect should be all the stronger in a hypermutable population.

The general threat that deleterious mutations pose to finite (real) asexual populations was pointed out by HJ Muller (Muller 1964). Muller described how the most-fit class of individuals in an asexual population is lost by the irreversible accumulation of deleterious mutations, in the absence of reverse mutations. This

process, described as Muller's ratchet by Felsenstein (Felsenstein 1974) is accelerated at a higher mutation rate, and in theory, can lead to the decline of fitness of an asexual population, consequently a decrease in population size, which may ultimately lead to extinction of the population. Although Muller's model is simplistic, since it does not include any beneficial mutations, it has been shown that if deleterious mutations do indeed hitchhike to fixation with adaptive mutations, they lower the overall benefit of the adaptive mutation since they are harmful to the organism (Johnson and Barton 2002; Peck 1994; Jiang et al. 2011). Indeed, theoretical studies have predicted that excessive mutation pressure can overwhelm adaptive evolution (Bull, Sanjuán, and Wilke 2007; Gerrish et al. 2007; Lynch et al. 1993). These predictions suggest that a high mutation rate population is likely to suffer a decrease in population size and may face the risk of extinction. In theory, such a population, if it substituted a lower mutation rate, could delay its extinction and step back from the brink, albeit temporarily.

Mutation rates can be decreased by the substitution of antimutator mutations, but these mutations, which presumably involve gains of function, are expected on genetic grounds to be rare (Drake 1993). Moreover, they are typically not expected to be present at appreciable frequencies in asexual populations for two reasons. First, an antimutator mutation may not be neutral with respect to its direct effect on fitness since it may increase the biochemical cost of replication by

increasing replication fidelity (unlike a mutator, which is more likely to be a loss of function mutation). Second, an antimutator may have a negative indirect effect on fitness, by lowering the chances of acquiring beneficial mutations, thereby reducing the competitive ability of individuals to their wild type counterparts. Despite these considerations, it is possible that in hypermutable populations that accumulate deleterious load at a high rate, an antimutator mutation may rise in frequency because it helps alleviate the fitness cost of accumulation of deleterious mutations (Wielgoss et al. 2012).

Mutation accumulation (MA) experiments, which have a rich experimental history (Bateman 1959; Lynch et al. 1999; Denver et al. 2009), show that when a population is propagated at very small size, its fitness declines, undoubtedly due to the accumulation of deleterious mutations and the scarceness of beneficial mutations. MA experiments are a method to estimate the deleterious mutation rate and have been employed to that end (reviewed in Foster 2006). However, these experimental setups may also be a great tool to study the evolution of mutation rates under a low rate of supply of beneficial mutations. In the first part of my dissertation, I employ the MA paradigm as a platform to ask questions about the evolution of mutation rate and fitness at low effective population size and very high mutation rate. Chapter 2 presents and discusses the implications of results from an MA experiment carried out on a hypermutable *Escherichia coli* strain in two contrasting growth media. One was a stressful growth medium

(minimal glucose, MG), and one was a rich growth medium (lysogeny broth, LB). Interestingly, I observed a reduction in mutation rates in populations propagated in both media that survived to the conclusion of the experiment, but an increase in mutation rates in populations that went extinct during the experiment. This evolution of mutation rates in such short timescales is novel and might suggest that these populations were under strong selection pressure to avoid deleterious mutation load. These arguments and their implications are discussed in greater detail in Chapter 2.

Chapter 3 presents a genomic analysis that I carried out in order to investigate the genetic basis of the evolutionary reduction in mutation rates documented and discussed in Chapter 2. I discuss various methodologies used to discover potential genomic changes affecting mutation rates in these populations. I also discuss the implications of these genetic changes, and point to future work that is requisite to confirm the phenotypic effects of these genetic candidates.

Chapter 4 discusses the effect that deleterious mutations may have on the distribution of neutral mutants in an expanding population. I use computer simulations to explore the effects of a high influx of deleterious mutations on the distribution of neutral mutations, and show that these distributions can be significantly altered when the deleterious mutation rate is very high. I discuss the implications of this finding and suggest future experimental work related to it.

Chapter 5 summarizes the findings of my dissertation work and discusses possible future directions for research in the general area of my dissertation.

References

Andersson, D I, and D Hughes. 1996. “Muller’s Ratchet Decreases Fitness of a DNA-Based Microbe.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (2): 906–7. doi:10.1073/pnas.93.2.906.

Andre, Jean Baptiste, and Bernard Godelle. 2006. “The Evolution of Mutation Rate in Finite Asexual Populations.” *Genetics* 172 (1): 611–26. doi:10.1534/genetics.105.046680.

Bachtrog, Doris, and Isabel Gordo. 2004. “Adaptive Evolution of Asexual Populations under Muller’s Ratchet.” *Evolution; International Journal of Organic Evolution* 58 (7): 1403–13. doi:10.1554/03-595.

Balick, D, Sidhartha Goyal, E Jerison, and R Neher. 2012. “Rare Beneficial Mutations Can Halt Muller’s Ratchet.” *Bull. Am. Phys. Soc.* 57: 1–8.

Bateman, A J. 1959. “The Viability of Near-Normal Irradiated Chromosomes.” *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine* 1 (2): 170–80.

Bull, J J, R Sanjuán, and C O Wilke. 2007. “Theory of Lethal Mutagenesis for Viruses.” *Journal of Virology* 81 (6): 2930–39. doi:10.1128/JVI.01624-06.

Chao, L, and E C Cox. 1983. “Competition Between High and Low Mutating Strains of *Escherichia coli."Evolution* 37 (1): 125–34.

Charlesworth, B, M T Morgan, and D Charlesworth. 1993. “The Effects of Deleterious Mutations on Neutral Molecular Varaition.” *Genetics* 134 (4): 1289–1303.

Cox, E C, G E Degnen, and M L Scheppe. 1972. “Mutator Gene Studies in *Escherichia coli - mutS* Gene.” *Genetics* 72 (4): 551–67.

Denver, Dee R, Peter C Dolan, Larry J Wilhelm, Way Sung, J Ignacio Lucas-Lledó, Dana K Howe, Samantha C Lewis, et al. 2009. “A Genome-Wide View of *Caenorhabditis elegans* Base-Substitution Mutation Processes.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (38): 16310–14. doi:10.1073/pnas.0904895106.

Drake, J. W. 1993. “General Antimutators Are Improbable. J. Mol. Biol. 229: 8–13.No Title.” *J. Mol Bio* 229: 8–13.

Felsenstein, J. 1974. “The Evolutionary Advantage of Recombination. Genetics.” *Genetics*.

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection.* Oxford: Oxford University Press.

Fontanari, J F, a Colato, and R S Howard. 2003. “Mutation Accumulation in

Growing Asexual Lineages." *Physical Review Letters* 91 (21): 218101. doi:10.1103/PhysRevLett.91.218101.
Foster, Patricia L. 2006. "Methods for Determining Spontaneous Mutation Rates." *Methods in Enzymology* 409: 195–213. doi:10.1016/S0076-6879(05)09012-9.
Gentile, Christopher F, Szi-Chieh Yu, Sebastian Akle Serrano, Philip J Gerrish, and Paul D Sniegowski. 2011. "Competition between High- and Higher-Mutating Strains of *Escherichia coli*." *Biology Letters* 7 (3): 422–24. doi:10.1098/rsbl.2010.1036.
Gerrish, Philip J, Alexandre Colato, Alan S Perelson, and Paul D Sniegowski. 2007. "Complete Genetic Linkage Can Subvert Natural Selection." *Proceedings of the National Academy of Sciences of the United States of America* 104 (15): 6266–71. doi:10.1073/pnas.0607280104.
Giraud, a, I Matic, O Tenaillon, a Clara, M Radman, M Fons, and F Taddei. 2001. "Costs and Benefits of High Mutation Rates: Adaptive Evolution of Bacteria in the Mouse Gut." *Science (New York, N.Y.)* 291 (5513): 2606–8. doi:10.1126/science.1056421.
Haldane, J.B.S. 1927. "The Mathematical Theory of Natural and Artificial Selection." *Proc. Camb. Philos. Soc.* 23: 838–44.
Hermisson, Joachim, Oliver Redner, Holger Wagner, and Ellen Baake. 2002. "Mutation-Selection Balance: Ancestry, Load, and Maximum Principle." *Theoretical Population Biology* 62 (1): 9–46. doi:10.1006/tpbi.2002.1582.
Hill, W. G., and A Robertson. 1966. "The Effect of Linkage on Limits to Artificial Selection." *Genetical Research* 8: 269–94.
Hong, Esther S., Annie Yeung, Pauline Funchain, Malgorzata M. Slupska, and Jeffrey H. Miller. 2005. "Mutants with Temperature-Sensitive Defects in the *Escherichia coli* Mismatch Repair System: Sensitivity to Mispairs Generated in Vivo." *Journal of Bacteriology* 187 (3): 840–46. doi:10.1128/JB.187.3.840-846.2005.
Howe, Dana K, and Dee R Denver. 2008. "Muller's Ratchet and Compensatory Mutation in Caenorhabditis Briggsae Mitochondrial Genome Evolution." *BMC Evolutionary Biology* 8: 62. doi:10.1186/1471-2148-8-62.
Jiang, Xiaoqian et al. "The Influence of Deleterious Mutations on Adaptation in Asexual Populations." Ed. Henry Harpending. *PLoS ONE* 6.11 (2011): e27757. *PMC*. Web. 27 May 2016.
Johnson, Toby. 2000. "Theoretical Studies of the Interaction Between Deleterious and Beneficial Mutations The University of Edinburgh."
Johnson, Toby, and Nick H. Barton. 2002. "The Effect of Deleterious Alleles on Adaptation in Asexual Populations." *Genetics* 162 (1): 395–411.
LeClerc, J E, B Li, W L Payne, and T a Cebula. 1996. "High Mutation Frequencies among Escherichia Coli and Salmonella Pathogens." *Science (New York, N.Y.)* 274 (5290): 1208–11. doi:10.1126/science.274.5290.1208.

Lynch, M, J Blanchard, D Houle, T Kibota, S Schultz, L Vassilieva, and J Willis. 1999. “Spontaneous Deleterious Mutation.” *Evolution* 53 (3): 645–63.
Lynch, M, R Bürger, D Butcher, and W Gabriel. 1993. “The Mutational Meltdown in Asexual Populations.” *The Journal of Heredity* 84 (5): 339–44.
Mansky, Louis M, and Kristopher S Cunningham. 2000. “Virus Mutators and Antimutators.” *Trends in Genetics* 16 (11): 512–17. doi:10.1016/S0168-9525(00)02125-9.
Mao, E F, L Lane, J Lee, J H Miller, Emily F Mao, Laura Lane, Jean Lee, and Jeffrey H Miller. 1997. “Proliferation of Mutators in A Cell Population . Proliferation of Mutators in A Cell Population” 179 (2): 417–22.
McClintock, B. 1938. “The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of the Aberrant Mitotic Behaviour of Ring Shaped Chromosomes, Genetics 21.” *Genetics* 21: 315.
McDonald, Michael J., Yu Ying Hsieh, Yen Hsin Yu, Shang Lin Chang, and Jun Yi Leu. 2012. “The Evolution of Low Mutation Rates in Experimental Mutator Populations of Saccharomyces Cerevisiae.” *Current Biology* 22 (13): 1235–40. doi:10.1016/j.cub.2012.04.056.
Michaels, M L, C Cruz, and J H Miller. 1990. “mutA and mutC: Two Mutator Loci in Escherichia Coli That Stimulate Transversions.” *Proceedings of the National Academy of Sciences of the United States of America* 87 (23): 9211–15. doi:10.1073/pnas.87.23.9211.
Muller, H.J. 1964. “Muller, H.J., 1964. The Relation of Recombination to Mutational Advance. Mutat.” *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* 1: 2–9.
Notley-McRobb, L, and Ferenci T. 2000. “Experimental Analysis of Molecular Events during Mutational Periodic Selections\rin Bacterial Evolution.” *Genetics* 156 (4): 1493–1501.
Painter, P. R. 1975. “Mutator Genes and Selection for the Mutation Rate in Bacteria.” *Genetics* 79 (4): 649–60.
Peck, Joel R. 1994. “A Ruby I N the Rubbish: Beneficial Mutations, Deleterious Mutations.”
Raynes, Y, and P D Sniegowski. 2014. “Experimental Evolution and the Dynamics of Genomic Mutation Rate Modifiers.” *Heredity*, no. April: 1–6. doi:10.1038/hdy.2014.49.
Richardson, Anthony R, Zhong Yu, Tanja Popovic, and Igor Stojiljkovic. 2002. “Mutator Clones of Neisseria Meningitidis in Epidemic Serogroup A Disease.” *Proceedings of the National Academy of Sciences of the United States of America* 99 (9): 6103–7. doi:10.1073/pnas.092568699.
Shaver, Aaron C., Peter G. Dombrowski, Joseph Y. Sweeney, Tania Treis, Renata M. Zappala, and Paul D. Sniegowski. 2002. “Fitness Evolution and the Rise of Mutator Alleles in Experimental Escherichia Coli Populations.” *Genetics* 162 (2): 557–66.

Smith, J Maynard, R Burian, S Kauffman, P Alberch, J Campbell, B Goodwin, R Lande, D Raup, and L Wolpert. 1985. “Developmental Constraints and Evolution: A Perspective from the Mountain Lake Conference on Development and Evolution.” *The Quarterly Review of Biology* 60 (3). University of Chicago Press: 265–87.

Sniegowski, Paul D, Philip J Gerrish, Toby Johnson, and Aaron Shaver. 2012. “The Evolution of Mutation Rates: Separating Causesfrom Consequences.” *BioEssays*, 1–10.

Sniegowski, Paul D, Philip J Gerrish, and Richard E Lenski. 1997a. “Evolution of High Mutation Rates in Experimental Populations of E . Coli.” *Nature* 606: 703–5.

———. 1997b. “Evolution of High Mutation Rates in Experimental Populations of E . Coli.” *Nature* 387: 703–5. doi:10.1128/AEM.02595-09.

Thompson, Dawn a., Michael M. Desai, and Andrew W. Murray. 2006. “Ploidy Controls the Success of Mutators and Nature of Mutations during Budding Yeast Evolution.” *Current Biology* 16 (16): 1581–90. doi:10.1016/j.cub.2006.06.070.

Wielgoss, Sébastien, Jeffrey E Barrick, Olivier Tenaillon, Michael J Wiser, W James Dittmar, Stéphane Cruveiller, Béatrice Chane-Woon-Ming, Claudine Médigue, Richard E. Lenski, and Dominique Schneidera. 2012. “Mutation Rate Dynamics in a Bacterial Population Re Fl Ect Tension between Adaptation and Genetic Load.” *Proc Natl Acad Sci (USA)* 110 (1): 222–27. doi:10.5061/dryad.hb3b5.

CHAPTER 2

Evolution of reduced genomic mutation rates in hypermutable populations of *Escherichia coli* propagated at extremely small effective population size

(Adapted from a manuscript that will be submitted to the journal *Biology Letters*)

Introduction

Mutation is the ultimate source of the genetic variation—including genetic variation for mutation rate itself—that fuels evolution. Selection to increase the genomic mutation rate, driven by selective sweeps of beneficial mutations, can be strong and rapid where genetic linkage is present, as evidenced by numerous observations in experimental microbial populations. Selection to decrease the mutation rate, in contrast, is expected to depend on avoidance of mutational load and act over a longer time scale. In keeping with this latter expectation, there have been relatively few experimental observations of the evolution of reduced genomic mutation rates. Here, I report the rapid evolution of reduced mutation rates in hypermutable *E. coli* populations propagated at extremely small effective size—a circumstance under which selection is generally minimized. I suggest that a combination of two factors accounts for my observations: 1) the strength and immediacy of selection against accumulated deleterious mutations at a very high mutation rate, and 2) the ineffectiveness of selection on beneficial mutations

at small effective population size. I discuss the relevance of my results to current ideas about the evolution, maintenance, and consequences of high mutation rates in asexual populations.

Because the genomes of all organisms harbor loci that affect the genome-wide mutation rate, mutation rates can evolve through the effects of natural selection and other evolutionary forces. If mutation-rate-modifying alleles have negligible direct effects on individual fitness, then natural selection can only alter mutation rates indirectly, via linkage disequilibrium between modifiers and fitness-affecting mutations (Sniegowski et al. 2000; Lynch 2010; Baer 2008). Indirect selection to increase mutation rate is driven by hitchhiking of up-modifiers of mutation (mutators) with sweeping beneficial alleles and has been documented numerous times in experimental and natural microbial populations (Sniegowski, Gerrish, and Lenski 1997; LeClerc et al. 1996; Giraud et al. 2001; reviewed in Raynes and Sniegowski 2014). In contrast, indirect selection to decrease mutation rate depends on avoidance of mutational load, is expected to be relatively slow and weak, and has seldom been observed (McDonald et al. 2012; Wielgoss et al. 2012). Existing theory and observations thus suggest that where selection is minimized, the systematic evolution of reduced mutation rate is unlikely.

Selection is minimized intentionally in mutation accumulation (hereafter, MA) experiments, in which replicate populations founded from a single ancestral genome are propagated at extremely small effective size (N_e) for many generations (Bateman 1959; Lynch et al. 1999). Because genetic drift governs the fate of mutations when their selective effect is less than approximately the reciprocal of effective population size (Wright 1931; Ohta and Kimura, 1971), deleterious mutations that would otherwise be suppressed by purifying selection are free to accumulate along with truly neutral mutations in MA experiments, allowing estimation of their rate of occurrence (Bateman 1959; Lynch et al. 1999).

I have carried out an MA experiment with replicate populations derived from a hypermutable *E. coli* strain. I report that several of these populations evolved significantly lower mutation rates than that of their common ancestor and that at the conclusion of the experiment, population fitness was negatively correlated with mutation rate. I discuss the relevance of my results to current ideas about the evolution, maintenance, and consequences of high mutation rates in asexual populations.

Materials and Methods

1. Mutation accumulation experiment

Forty independent MA populations were established using random isolates from *E. coli* strain PS2534, which is resistant to the antibiotic tetracycline, harbors defects in mismatch repair (*mutL13*) and proofreading (*dnaQ905*), and exhibits a genomic mutation rate ~4500 fold higher than that of wild type *E. coli* (Gentile et al. 2011). Twenty populations were propagated on minimal glucose (MG) agar plates (Lenski 1988) and the remaining 20 populations were propagated on lysogeny broth (LB) agar plates (Miller 1972). All plates were supplemented with tetracycline (15µg/ml) to avoid contamination. The effect of tetracycline on mutation rates of these populations was not significant (see Appendix). Populations were incubated at 37° C. Every 24 h, each population was bottlenecked to a size of one by streaking a random, isolated colony derived from a single cell to a fresh agar plate in order to isolate a new such colony. Intermediate stages of the experiment were archived at -80° C in 15% glycerol every 5 days. If colony growth was not visible after 24 hours for a given population during its propagation, I incubated the population for another 24 hours; if no growth was visible at 48 hours, the population was provisionally considered to have gone extinct and was restarted from its previously frozen time point. After three consecutive failed restarts, no further attempt was made to propagate the population and it was considered to have gone extinct.

Populations that did not go extinct were propagated for a total of 50 daily transfers, corresponding to approximately 1250 generations of binary fission based on daily growth from 1 to ~3 x 10^7 cells.

2. Fitness measurements

I estimated population fitnesses in liquid media every ten transfers by measuring growth rates and maximal (24 h) absorbance values at 600 nm. Frozen intermediate time points were inoculated into and grown in flat-bottomed 96-well microplates containing 15µg/ml of tetracycline in 250 µL of either Davis minimal medium (DM) supplemented with 1g/L of glucose (Carlton and Brown 1981), or LB; absorbance values during culture growth were measured on an automated plate reader (Thermo Fisher MultiSkan GO) every 2 min. Viable CFU (colony forming units) counts were obtained as an additional measure of fitness by dilution and plating of 24 h liquid cultures to appropriate media. Plates were incubated at 37° C with shaking over a 24-hour period in an automated plate reader (Thermo-Fisher Multiskan-GO); every 2 min during this incubation, absorbance at 600 nm was estimated by the plate reader as a proxy for cell density. These results are shown in Figure 2.1. As an additional assay of fitness at the end of the MA experiment, I carried out dilution and plating of 24 h liquid cultures of the ancestral strain and the endpoint MA populations (grown in LB or DM broth) in order to estimate viable cell densities via counts of CFUs.

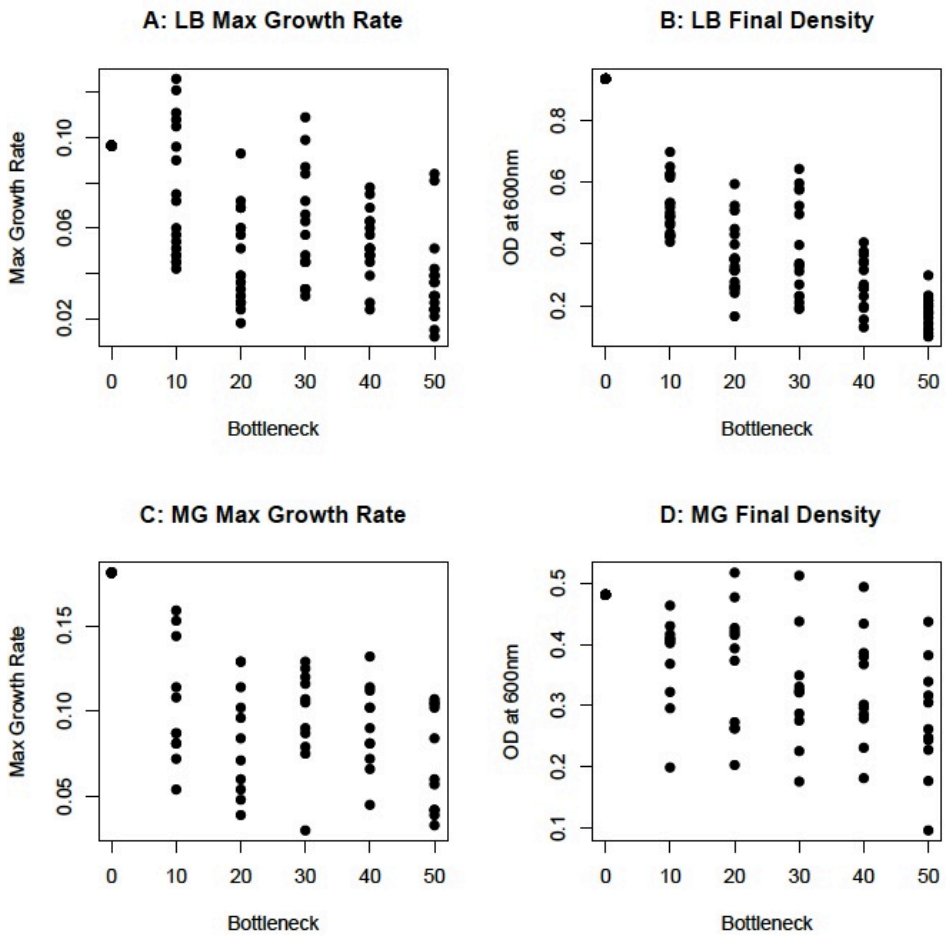


Figure 2.1 Measures of fitness in the evolving MA populations. Maximum growth rates (change in OD at 600nm/hr.) and final absorbance values (600 nm) of LB (A-B) and minimal glucose (C-D) MA populations propagated for 24 h on liquid media as described in the Methods. The decline in maximum growth in LB and MG and final density in LB and MG are statistically significant (p<0.05).

3. Estimation of mutation rates

Mutation rates to nalidixic acid resistance and streptomycin resistance (which arise at different genetic loci) were estimated in the ancestral strain and the

evolved MA populations using a modified version (Jones 1994; Gerrish 2008) of the Luria-Delbrück fluctuation assay (Luria and Delbrück 1943). The fluctuation assay was done as follows: A small number of large independent cultures of the strain was grown in the appropriate media (MG or LB liquid media) and a fixed fraction of each culture is plated to selective medium to enumerate mutants. In the modified version of the fluctuation assay employed here, a 30 ml culture is grown in triplicate for each clone for which a mutation rate estimate is desired. The cultures are inoculated with a small number of cells from an overnight growth of the clone from a frozen stock. Usually the overnight culture is diluted about a million fold before inoculation into the large cultures. The large cultures are then incubated for 48 hours and then a fraction of the culture (usually 100 µl from a 30 ml culture) is plated on a selective plate in order to enumerate mutants. To accurately estimate the mutation rate, it is also required to know the final population size of the large cultures. This is estimated by diluting the large culture appropriately and plating on a permissive medium so that all cells may grow. After obtaining the number of mutants that grew on the selective medium and the population size of each replicate, mutation rates were estimated using a maximum likelihood approach. Maximum likelihood mutation rates and 95% confidence intervals from these assays were calculated with software kindly provided by Dr. Philip Gerrish.

4. Computer Simulations

In order to test the hypothesis that the antimutator alleles may have an indirect effect on fitness by reducing the rate at which deleterious mutations accumulate, and that such an effect may be strong enough to drive these alleles to fixation, I performed computer simulations that would mimic the process of mutation accumulation, i.e. exponential growth of a population starting from a single individual and then drawing a random individual from the population, and repeating the process over again. I incorporated mutation rates into my model, such that every individual carried a hypermutable mutation rate at the start of the simulation. There was, however, some probability associated with acquiring a lower mutation rate; this was fixed at $1x10^{-6}$, which is based on the assumption that antimutator mutations are gain of function mutations and therefore tend to be rare (Drake 1993). Individuals could also acquire a higher mutation rate, with a certain constant probability (fixed at $1x10^{-4}$), following the notion that most mutator mutations are likely to be loss of function mutations (Miller et al. 2002; Siegel and Bryson 1967; Michaels, Cruz, and Miller 1990; Shaver and Sniegowski 2003). Both the mutator and the antimutator mutations have the same effect on mutation rate in the simulations, i.e. they increase or decrease the mutation rate by the same factor, in this case 100.

In addition to the mutator and antimutator mutations, I also included beneficial mutations and deleterious mutations in my model. Deleterious mutations

occurred with a probability of 0.9, which is based on the mutation rate of the hypermutable strain that was employed for the mutation accumulation experiments which has been previously described. That particular strain was constructed by transducing a non-functional *dnaQ* allele that contains 4 point mutations, in a strain that is already defective for a mismatch repair gene, *mutL*. The resultant strain, PS2534, has a mutation rate approximately 4500-fold higher than wild type. Wild type *E. coli* possesses a deleterious mutation rate of 0.0002, per generation per genome (Lynch et al. 1999), and therefore the hypermutable strain PS2534 has a mutation rate that translates to approximately 0.9 deleterious mutations per generation per genome. The deleterious mutations in my simulations were drawn from a gamma distribution, with shape parameter 0.3 and scale parameter 0.1, which were selected to produce an overall mean effect of deleterious mutations close to 0.03, to be consistent with previous simulation studies, and the general consensus of the effect size of deleterious mutations in the literature (Piganeau and Eyre-Walker 2003; Gerrish et al. 2007; Keightley 2012). In addition, beneficial mutations were also included in my simulation, with a mean effect size of 0.03 as well, and drawn from an exponential distribution, implying that very few beneficial mutations of very large effect existed.

The computer simulations were carried out in C++ (code available upon request). They are individual based, with exponential growth starting from 1 individual (at every bottleneck) and increasing in size to 4×10^7 (~23 generations) before being

bottlenecked again. Both deleterious and beneficial mutations are included in the simulations, and the mutation rate to deleterious mutations is ~0.9 per genome per generation, akin to the hypermutable strain I employed for the MA experiment.

Results

1. Mutation Accumulation Experiment

Measures of fitness declined significantly in the MA populations during their propagation (Figure 2.1). Indeed, 2 of the LB populations and 9 of the MG populations went extinct. Among-population variance in fitness, however, showed little or no evidence of increase over the course of propagation. Mutation rates to nalidixic acid resistance were significantly lower than that of the ancestor in 6 of the 9 surviving MG populations for which I was able to estimate mutation rates and 12 of the 18 surviving LB populations, with some populations exhibiting evolved reductions in mutation rate of over tenfold compared with the ancestor (Figure 2.2). Mutation rates to nalidixic acid resistance from two independent assays were significantly correlated ($p < 0.05$), supporting the overall reliability of the fluctuation assay (see Appendix). Mutation rates to streptomycin resistance were lower than that of the ancestor in all populations for which nalidixic acid mutation rate had decreased with the exception of population 14 from the MG subset, although there was not a significant correlation between nalidixic acid

and streptomycin resistance mutation rates. Moreover, all of the populations that went extinct during the experiment exhibited streptomycin resistance mutation rates that were significantly higher than the ancestral mutation rate, and nalidixic acid resistance mutation rates that were also significantly higher than the ancestor, with the exception of populations LB4 and LB11 (Figure 2.2 and Figure 2.3). Finally, there was a significant negative correlation between mutation rate to nalidixic acid resistance and endpoint population fitness relative to the ancestor (as measured by CFUs in 24 h cultures), in the surviving MA lines, (see figure 2.4), for the LB populations. Although there was a negative correlation between the endpoint fitness and mutation rates to nalidixic acid in the MG populations, it was not statistically significant, perhaps due to a very low sample size, since I was able to obtain precise mutation rates for only 9 out of the 11 populations that survived the 50 bottlenecks in the MG medium.

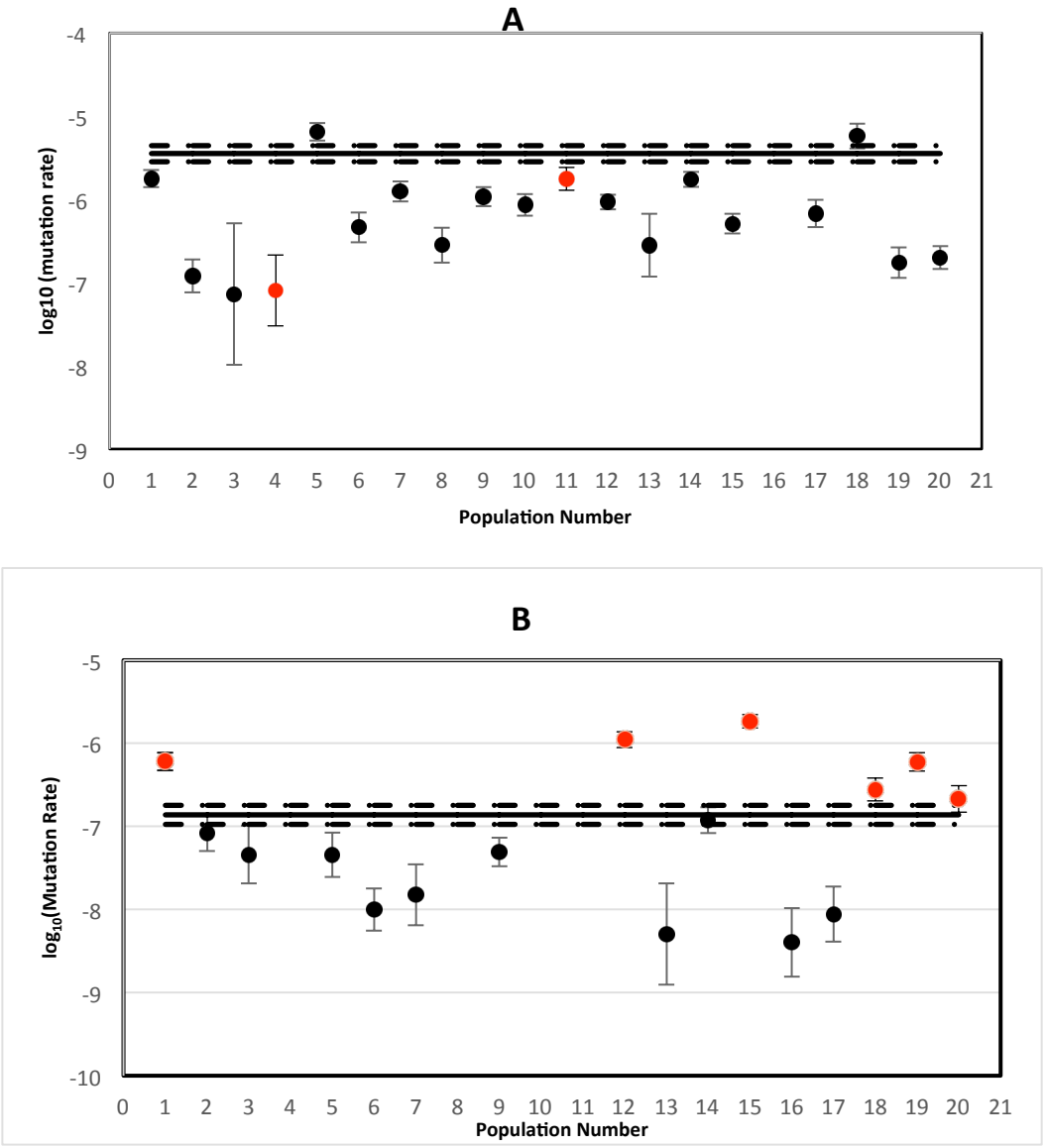


Figure 2.2. Maximum-likelihood mutation rates to nalidixic acid resistance and associated 95% confidence limits for the LB (A) and MG (B) populations. Black markers represent the populations that survived all 50 bottlenecks and red markers represent the populations that went extinct during the experiment. Solid horizontal line gives the estimated mutation rate in the ancestral strain, PS2534; upper and lower dotted lines represent the 95% confidence interval of the estimate.

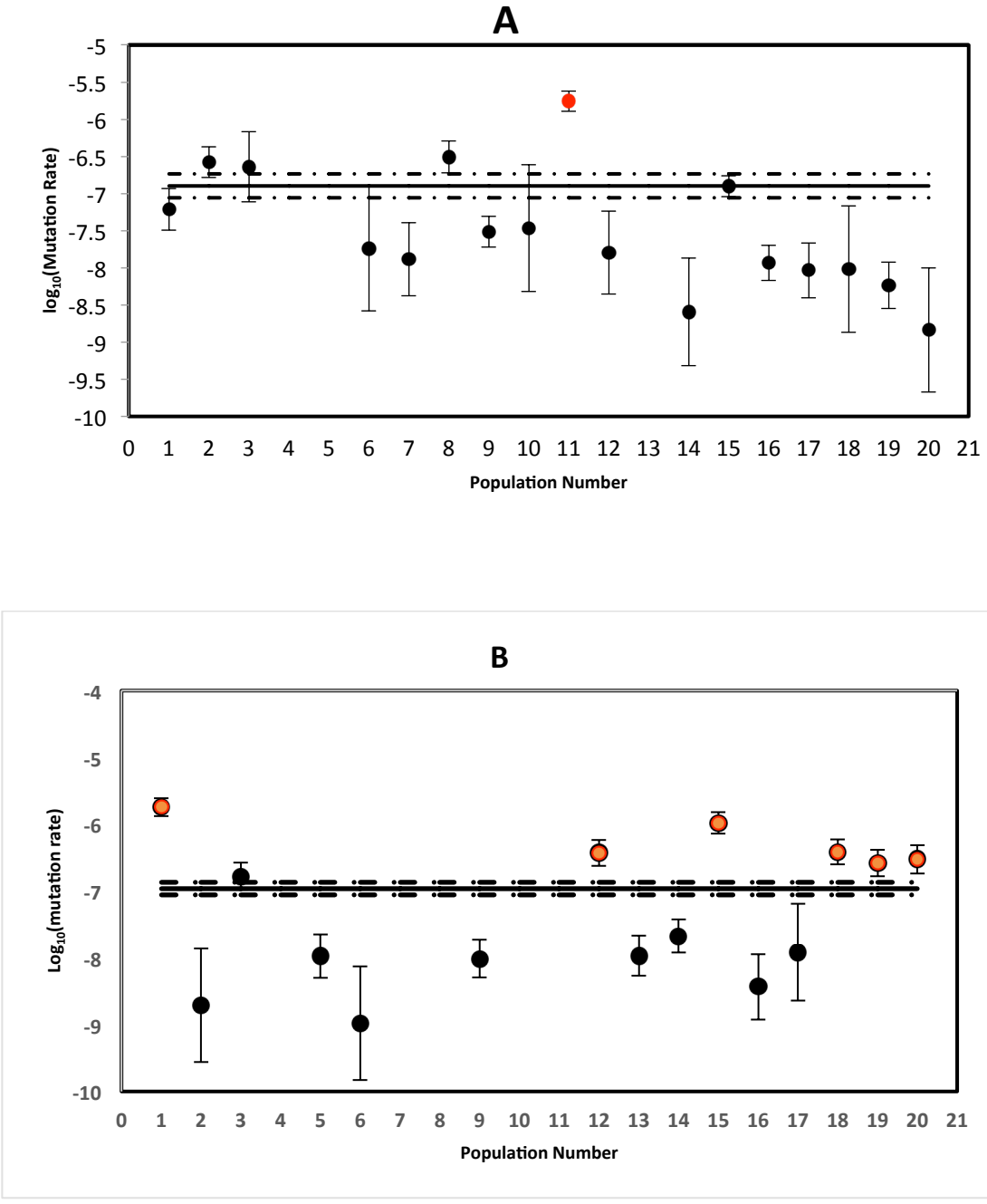


Figure 2.3. Maximum-likelihood mutation rates to streptomycin resistance and associated 95% confidence limits for the LB (A) and MG (B) populations that survived all 50 transfers of the MA experiment. The black markers represent the surviving populations and the red markers represent the extinct populations. Solid

horizontal line gives the estimated mutation rate in the ancestral strain, PS2534; upper and lower dotted lines represent the 95% confidence interval of the estimate.

2. Computer Simulations

As mentioned previously in the methods section, for my simulations I assumed that antimutator mutations are rarer than mutator mutations (100 fold rarer). I found that in simulations for 30 replicates, the antimutator fixed 10 times and the mutator fixed only one time. Only 30 replicates were carried out since these simulations tend to be very memory intensive and hence take a long time to run. In order to test whether the fixation could have resulted from drift alone, I encoded a marker in the simulations that has no effect on fitness and tested its probability of fixation over the 50 bottlenecks in 30 replicates. I did not observe a single fixation event for this neutral marker, indicating that the increased likelihood of fixation of antimutators in my simulations could not have resulted from drift alone. Also, to test whether the probability of fixation is indeed higher when the population size is extremely reduced, i.e. when beneficial mutations are rare enough that they cannot offset the cost of deleterious mutation accumulation, I carried out some simulations at higher effective population sizes. This was accomplished via bottlenecking the population down to 1024 individuals at every bottleneck, followed by another twelve generations of growth. In this case, the effective population size is ~12000 individuals and hence the probability of obtaining a beneficial mutation was non-negligible. In this case, I did not observe fixation of either antimutator alleles or mutator alleles over 50

bottlenecks for 30 replicates, suggesting that it is unlikely that an antimutator would fix at a higher effective population size.

Discussion

MA populations are expected to decline in average fitness over time because most mutations that affect the phenotype decrease fitness; because mutations accumulate in a stochastic manner, fitness variance among MA populations is expected to increase over time (Bateman 1959). Based on the number of generations between bottlenecks (Lenski et al. 1991), the effective size of my hypermutable MA populations was ~25 individuals. Thus, mutations of selective effect substantially smaller than 4% were free to accumulate in these populations. Indeed, these populations showed substantial and significant fitness declines and some went extinct. There was little evidence, however, for increase in fitness variance among the populations, perhaps owing to their extraordinarily high mutation rate. Because my *E. coli* ancestor strain is expected to have a deleterious mutation rate of at least 0.9 per generation per genome (Gentile et al. 2011), it may well be that substantial among-population variance in fitness was to be expected immediately in my MA experiment and that further increases in variance would be negligible.

What was unexpected in my experiment was the observation of reduced genomic mutation rates in some populations. Preliminary genome sequencing of these

populations confirms that this result is not simply a consequence of external contamination (see Chapter 3), and thus an evolutionary explanation is required. In general, natural selection based on the avoidance of mutational load is predicted to act only slowly and weakly to decrease mutation rates (Lynch 2010). Two factors, however, may favor the rapid evolution of reduced mutation rates in hypermutable MA populations: First, selective pressure to avoid mutational load may well be quite strong, especially if fitness approaches a minimum viable value as deleterious mutations accumulate. Under these circumstances, a modifier that reduces genomic mutation rate by several-fold (an "antimutator") could spread because it increases the average relative fitness of an individual's descendants by more than the selective threshold imposed by the daily bottleneck regime.

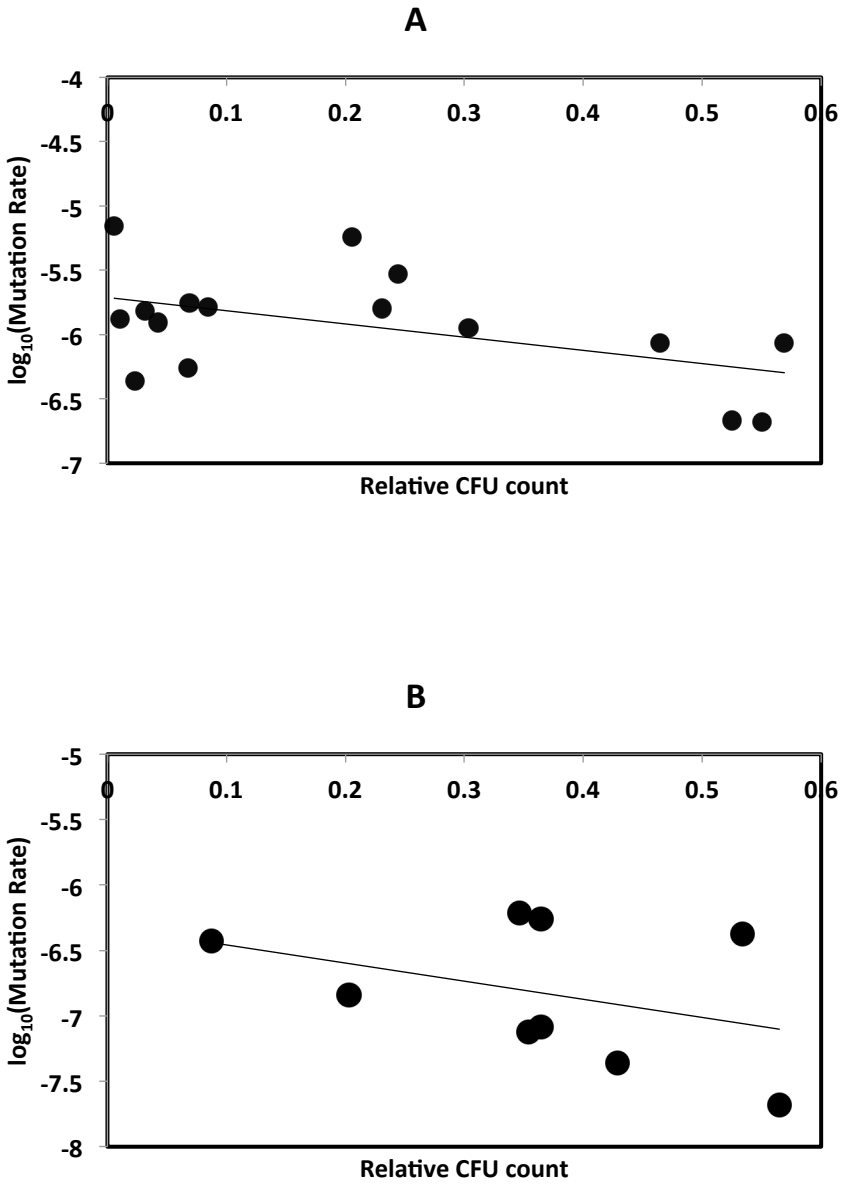


Figure 2.4 Fitnesses of the surviving of MA lines estimated as relative number of CFUs (colony forming units) with respect to the ancestor, PS2534, in both LB (A) and MG (B) media plotted against $\log_{10}$(Mutation Rate) to nalidixic acid resistance. The negative correlation between relative CFU counts and mutation rates is statistically significant ($p<0.05$) for the LB population, but not for the MG populations.

This interpretation is consistent with my finding that the mutation rates to streptomycin and nalidixic acid resistance were in general significantly higher in those populations that went extinct than in the populations that persisted to the end of the MA experiment (Figure 2). The finding that the number of viable cells at 24 h is negatively correlated ($p<0.05$) with mutation rates to nalidixic acid resistance in the surviving LB populations lends further support to this idea. Second, because selection in favor of mildly to moderately beneficial mutations is neutralized by small effective population size, any advantage that a high-mutation-rate lineage might have in its faster acquisition of beneficial mutations (de Visser 2002) could be substantially diminished in MA populations. Consistent with both of the foregoing ideas, individual-based computer simulations in which mutations affecting both mutation rates and fitness can occur strongly suggest that hypermutable MA populations are more likely to substitute antimutators than mutators (see Results). In sum, the cost of a high mutation rate (deleterious mutations) seems likely to persist or even increase while its benefit (faster acquisition of beneficial mutations) diminishes in very small hypermutable populations. This phenomenon of reduced mutation rates has been previously observed via simulations when beneficial mutations are absent, and deleterious mutations are abundant (Gerrish et al. 2007).

Selection is not the only evolutionary force likely to be operating in hypermutable MA populations: mutation pressure and genetic drift could play significant roles

as well. Indeed, the observation that, in general, the extinct populations (and one of the surviving populations) had mutation rates *higher* than that of the ancestor suggests a role for mutational pressure, and perhaps drift, in further compromising the ability of some populations to prevent mutations (Lynch 2008; Gerrish et al. 2007).

Historically, MA experiments have been used as an effective means to estimate the deleterious mutation rate, with the underlying premise that the deleterious mutation rate remains constant throughout the experiment. My results indicating that mutation rate may be liable to evolve during an MA experiment have implications for estimation of mutation rates via MA experiments, especially since many MA studies have been carried out with mutator strains in recent years because the high mutability of mutator strains requires fewer replicate populations to be propagated (Maharjan et al. 2013; Heilbron et al. 2014).

Finally, my results have some implications for recent ideas concerning mutation rate evolution and the fate of asexual populations. Theoretical and experimental work (Gerrish et al. 2007; Andre and Godelle 2006; Gentile et al. 2011) predicts that recurrent mutator hitchhiking can cause mutation rate evolution to be upwardly biased in adapting asexual populations, perhaps even culminating in a mutation rate that causes extinction (Gerrish et al. 2007). The extremely small size and high mutation rate of my MA populations may well mimic the terminal

circumstances envisioned by these recent studies. The results of my work suggest the interesting possibility that, as population size declines and deleterious mutations accumulate under the influence of a very high mutation rate, some populations may pull back from the brink of extinction—if only temporarily—by evolving reduced mutation rates.

References

Andre, Jean Baptiste, and Bernard Godelle. 2006. “The Evolution of Mutation Rate in Finite Asexual Populations.” *Genetics* 172 (1): 611–26. doi:10.1534/genetics.105.046680.

Baer, Charles F. 2008. “Does Mutation Rate Depend on Itself?” *PLoS Biology* 6 (2): 0233–35. doi:10.1371/journal.pbio.0060052.

Bateman, A J. 1959. “The Viability of Near-Normal Irradiated Chromosomes.” *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine* 1 (2): 170–80.

Carlton, B C, and B J Brown. 1981. “Gene Mutation. In: Manual of Methods for General Bacteriology.” In , 222–42.

de Visser, J. Arjan G M. 2002. “The Fate of Microbial Mutators.” *Microbiology* 148 (5): 1247–52. doi:10.1099/00221287-148-5-1247.

Gentile, Christopher F, Szi-Chieh Yu, Sebastian Akle Serrano, Philip J Gerrish, and Paul D Sniegowski. 2011. “Competition between High- and Higher-Mutating Strains of Escherichia Coli.” *Biology Letters* 7 (3): 422–24. doi:10.1098/rsbl.2010.1036.

Gerrish, Philip. 2008. “A Simple Formula for Obtaining Markedly Improved Mutation Rate Estimates.” *Genetics* 180 (3): 1773–78. doi:10.1534/genetics.108.091777.

Gerrish, Philip J, Alexandre Colato, Alan S Perelson, and Paul D Sniegowski. 2007. “Complete Genetic Linkage Can Subvert Natural Selection.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (15): 6266–71. doi:10.1073/pnas.0607280104.

Giraud, a, I Matic, O Tenaillon, a Clara, M Radman, M Fons, and F Taddei. 2001. “Costs and Benefits of High Mutation Rates: Adaptive Evolution of Bacteria in the Mouse Gut.” *Science (New York, N.Y.)* 291 (5513): 2606–8. doi:10.1126/science.1056421.

Heilbron, Karl, Macarena Toll-Riera, Mila Kojadinovic, and R. Craig MacLean. 2014. “Fitness Is Strongly Influenced by Rare Mutations of Large Effect in a

Microbial Mutation Accumulation Experiment." *Genetics* 197 (3): 981–90. doi:10.1534/genetics.114.163147.
Jones, Michael E. 1994. "LB Fluctuation Experiements; Accounting Simultaneously for Plating Efficiency and Differential Growth Rate." *J. Theor. Biol* 166: 355–63.
Keightley, Peter D. 2012. "Rates and Fitness Consequences of New Mutations in Humans." *Genetics* 190 (2): 295–304. doi:10.1534/genetics.111.134668.
LeClerc, J E, B Li, W L Payne, and T a Cebula. 1996. "High Mutation Frequencies among Escherichia Coli and Salmonella Pathogens." *Science (New York, N.Y.)* 274 (5290): 1208–11. doi:10.1126/science.274.5290.1208.
Lenski, R. E. 1988. "Experimental Studies of Pleiotropy and Epistasis in Escherichia Coli. I. Variation in Competitive Fitness among Mutants Resistant to Virus T4." *Evolution* 42: 425–32.
Lenski, Richard E., Michael R. Rose, Suzanne C. Simpson, and Scott C. Tadler. 1991. "Long-Term Experimental Evolution in Escherichia Coli. I. Adaptation and Divergence During 2,000 Generations." *The American Naturalist* 138 (6): 1315. doi:10.1086/285289.
Luria, S E, and M Delbrück. 1943. "Mutations of Bacteria from Virus Sensitivity to Virus Resistance." *Genetics* 28 (6): 491–511. doi:10.1038/nature10260.
Lynch, M, J Blanchard, D Houle, T Kibota, S Schultz, L Vassilieva, and J Willis. 1999. "Spontaneous Deleterious Mutation." *Evolution* 53 (3): 645–63.
Lynch, Michael. 2008. "The Cellular, Developmental and Population-Genetic Determinants of Mutation-Rate Evolution." *Genetics* 180 (2): 933–43. doi:10.1534/genetics.108.090456.
———. 2010. "Evolution of the Mutation Rate." *Trends in Genetics* 26 (8). Elsevier Ltd: 345–52. doi:10.1016/j.tig.2010.05.003.
Maharjan, Ram P, Bin Liu, Yang Li, Peter R Reeves, Lei Wang, and Thomas Ferenci. 2013. "Mutation Accumulation and Fitness in Mutator Subpopulations of Escherichia Coli." *Biology Letters* 9 (1): 20120961. doi:10.1098/rsbl.2012.0961.
McDonald, Michael J., Yu Ying Hsieh, Yen Hsin Yu, Shang Lin Chang, and Jun Yi Leu. 2012. "The Evolution of Low Mutation Rates in Experimental Mutator Populations of Saccharomyces Cerevisiae." *Current Biology* 22 (13): 1235–40. doi:10.1016/j.cub.2012.04.056.
Michaels, M L, C Cruz, and J H Miller. 1990. "mutA and mutC: Two Mutator Loci in Escherichia Coli That Stimulate Transversions." *Proceedings of the National Academy of Sciences of the United States of America* 87 (23): 9211–15. doi:10.1073/pnas.87.23.9211.
Miller, J. H. 1972. *Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.*
Miller, Jeffrey H., Pauline Funchain, Wendy Clendenin, Tiffany Huang, Anh Nguyen, Erika Wolff, Annie Yeung, 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 (1): 5–13.

Ohta, Tomoko, and M Kimura. 1971. "On the Constancy of the Evolutionary Rate of Cistrons " 25 (813).

Piganeau, Gwenaël, and Adam Eyre-Walker. 2003. "Estimating the Distribution of Fitness Effects from DNA Sequence Data: Implications for the Molecular Clock." *Pnas* 100 (18): 10335–40. doi:10.1073/pnas.1833064100.

Raynes, Y, and P D Sniegowski. 2014. "Experimental Evolution and the Dynamics of Genomic Mutation Rate Modifiers." *Heredity*, no. April: 1–6. doi:10.1038/hdy.2014.49.

Shaver, a C, and P D Sniegowski. 2003. "Spontaneously Arising mutL Mutators in Evolving Escherichia Coli Populations Are the Results of Changes in Repeat Length." *J. Bacteriol.* 185 (20): 6076–82. doi:10.1128/JB.185.20.6076.

Siegel, Eli C., and Vernon Bryson. 1967. "Mutator Gene of Escherichia Coli B1." *Journal of Bacteriology* 94 (1): 38–47.

Sniegowski, Paul D, Philip J Gerrish, Toby Johnson, and Aaron Shaver. 2000. "The Evolution of Mutation Rates: Separating Causesfrom Consequences." *BioEssays*, 1–10.

Sniegowski, Paul D, Philip J Gerrish, and Richard E Lenski. 1997. "Evolution of High Mutation Rates in Experimental Populations of E . Coli." *Nature* 606: 703–5.

Springman, R., T. Keller, I. J. Molineux, and J. J. Bull. 2010. "Evolution at a High Imposed Mutation Rate: Adaptation Obscures the Load in Phage T7." *Genetics* 184 (1): 221–32. doi:10.1534/genetics.109.108803.

Wielgoss, Sébastien, Jeffrey E Barrick, Olivier Tenaillon, Michael J Wiser, W James Dittmar, Stéphane Cruveiller, Béatrice Chane-Woon-Ming, Claudine Médigue, Richard E. Lenski, and Dominique Schneidera. 2012. "Mutation Rate Dynamics in a Bacterial Population Re Fl Ect Tension between Adaptation and Genetic Load." *Proc Natl Acad Sci (USA)* 110 (1): 222–27. doi:10.5061/dryad.hb3b5.

Wright, Sewall. 1931. "Evolution in Mendelian Populations." *Genetics* 16: 97–159.

CHAPTER 3

Genomic Analysis of Hypermutable *E. coli* Mutation Accumulation Populations

Introduction

In the preceding chapter, I reported the reduction of mutation rates in hypermutable *E. coli* populations that were propagated at extremely low population sizes and presented hypotheses to explain this unexpected observation. I predicted that the observed reductions in mutation rate may not have been the consequence of a direct effect on fitness, but may have resulted instead from a reduction in the deleterious mutation load going forward in time, thus increasing the likelihood for survival relative to other individuals that do not reduce mutation rates. This hypothesis was supported by my finding that there was a negative correlation between the evolved mutation rate and fitness at the end of the experiment. In addition, estimation of mutation rates of the populations that went extinct during the experiment revealed that they had evolved higher mutation rates, strengthening the argument that lower mutation rates may have had some indirect selective advantage in my experiment. In this chapter, I report findings from analysis of genomic data of the populations that were mentioned in the preceding chapter.

Genomic sequencing is an extremely powerful tool to analyze long-term evolution experiments. Recently, many genomic studies on long term evolution experiments have been undertaken and have brought interesting results to light (Barrick and Lenski 2009; Cooper, Rozen, and Lenski 2003; Kinnersley et al. 2014; Kao and Sherlock 2008). In addition, recent advances in genomic sequencing technology have made the process less expensive and more user-friendly (Buermans and den Dunnen 2014). The presence of multiple open-source pipelines for the analysis of genome sequence data also makes the process of interpretation less time- and labor-intensive and more accessible to the uninitiated (Deatherage and Barrick 2014).

I obtained next-generation genomic sequences for all of the surviving hypermutable populations that were propagated in the MA experiment described in Chapter 2. There were 18 surviving LB populations and 11 surviving MG populations. Genome sequencing was done in collaboration with Dr. Vaughn Cooper at the University of Pittsburgh, and the subsequent analysis was done using the open source genome sequence analysis pipeline breseq (Deatherage and Barrick 2014) developed by Dr. Jeffrey Barrick. I obtained approximately 200X coverage for each of my samples.

The sequence analysis was undertaken to understand the mutational dynamics of these high mutation rate strains when propagated at very low effective

population sizes. Deleterious mutations are expected to accumulate under propagation at low effective population sizes since selection is ineffective at purging any deleterious mutations that have an effect size smaller than the reciprocal of the effective population size (Wright 1931; Ohta and Kimura, 1971).

I was primarily interested in analyzing the genomic basis of the reduced mutation rates in the surviving populations of the MA experiment. Since many loci are implicated in mutation rate evolution, this required an extensive literature survey, to curate a list of mutations that are known to have beneficial effects in these media.

In the past, mutator alleles that have arisen in laboratory evolution experiments have been identified by sequencing approaches (Shaver and Sniegowski 2003; Wielgoss et al. 2012). However, the incidence of lower mutation rates is rarer and thus the genomic basis of lower mutation rates is largely unexplored (but see Wielgoss et al. 2012). Here, I was able to identify some candidate mutations that may have resulted in a lower mutation rate in some of the surviving populations. These mutations are listed and discussed in the sections that follow. In addition, the genome sequencing confirmed that all of the surviving populations had the original mutator alleles (*dnaQ905, mutL13*), confirming the absence of any external contamination in the experiment.

Methods

Genomic DNA was extracted from the surviving populations using the QIAGEN blood and tissue DNA extraction kit, which was specially optimized for bacterial species by heating the Elution Buffer and increasing the incubation period in the elution column before elution (Dr. Kathleen Sprouffske, personal communication). This optimization was necessary because some of the surviving MA lines had an extremely low growth rate and would not produce very dense cultures. In addition, RNAse was added to the mix before the enzymatic extraction of DNA was done in order to eliminate RNA contamination in the samples.

Genomic DNA concentration was determined using a NanoDropTM 8000 Spectrophotometer. Any sample with a concentration less than 10ng/µl was rejected and the extraction was repeated till the concentration was higher than that threshold. In addition, the genomic DNA was also analyzed by carrying out gel electrophoresis to determine if there was any contamination by RNA. RNA usually produces a fainter band, which sometimes appears as a smear on the gel. After ensuring that there was no RNA contamination in my samples, the samples were shipped on dry ice to Dr. Vaughn Cooper's laboratory, where they were prepared for next-generation sequencing on an Illumina Hi-Seq platform. The technician in Dr. Cooper's laboratory performed the library preparation and

ligation of adapters. After the sequencing was done, Dr. Cooper's laboratory shared the output files (one forward and reverse file for every population) after they had performed a quality check on the data.

I used the breseq pipeline to analyze the genomic sequences. I used Bowtie2 (Langmead and Salzberg 2012) to build the index files for the reference sequences that would be used by breseq to analyze the data. I then ran breseq on the individual forward and reverse files of each population and obtained candidate SNP (Single Nucleotide Polymorphism) information, along with information regarding insertions and deletions. The SNPs were called using a standard *E. coli* K12 reference sequence, which differed slightly from my own ancestral sequence, and so I used a custom script I wrote in the statistical data analysis software R to change the reference sequence to more closely resemble the common ancestor of my MA experiment. This sequence was subsequently used to call SNPs. After these SNPS were obtained, I used the Missing Coverage information that breseq provides to remove any SNPs that were called that occurred in the regions where the coverage was missing or low. After these processes were all carried out, the SNP data were analyzed for candidate changes that could have influenced the mutation rates of these populations. I used a script I wrote in the statistical data analysis program R to filter the list of all SNPs based on certain criteria (for example, non-synonymous).

The first step in identifying candidate mutator/antimutator SNPs was curating a list of loci that have historically been known to be associated with mutation rate changes. After an extensive survey of the microbial mutation rate literature and consulting with the EcoCyc database (Keseler et al. 2011; Keseler et al. 2013), I compiled a list of 55 loci that have been documented to be associated with changes in mutation rates. Below, I discuss the most promising candidates among these, as evidenced by my SNP data, as well as some results from comparing the mutational spectrum under the two different environments.

Results

1. Ratio of non-synonymous to synonymous SNPs and transitions to transversions

It is well documented that bacteria have a transition bias with more AT -> GC transitions. Also, because the hypermutable strain used as a common ancestor for my MA experiment is a mismatch repair deficient strain, it confers an even higher transition to transversion rate. This was confirmed by the SNP data (see table 3.1), in both the overall genome and the mutation loci of interest, which will be discussed in greater detail subsequently. In addition, I also measured the correlation between endpoint mutation rates and the total number of SNPs across all LB and MG populations (see Figure 3.1), and found that the number of SNPs and end point mutation rate are not significantly correlated. The lack of

significance could be perhaps due to the reduction in mutation rate occurring closer to the end of the MA experiment, and thus the number of SNPs would not be very reflective of the actual end point mutation rate. Additionally, in general, the LB populations have accumulated more SNPs than the MG populations, since they possess a higher mutation right from the start, i.e. the common ancestor for both LB and MG populations, although isogenic, has a perceptibly higher mutation rate in LB than that MG population (Chapter 2), which has been observed before (Ishizawa et al. 2015).

Environment	**Transitions (*Ti*)**	**Transversions (*Tv*)**	***Ti/Tv***
MG			
Whole Genome	3613	349	10.4
Mutation Rate Loci	74	5	14.8
LB			
Whole Genome	16243	956	16.99
Mutation Rate Loci	306	19	16.11

Table 3.1 Ratio of transitions to transversions in MG and LB populations

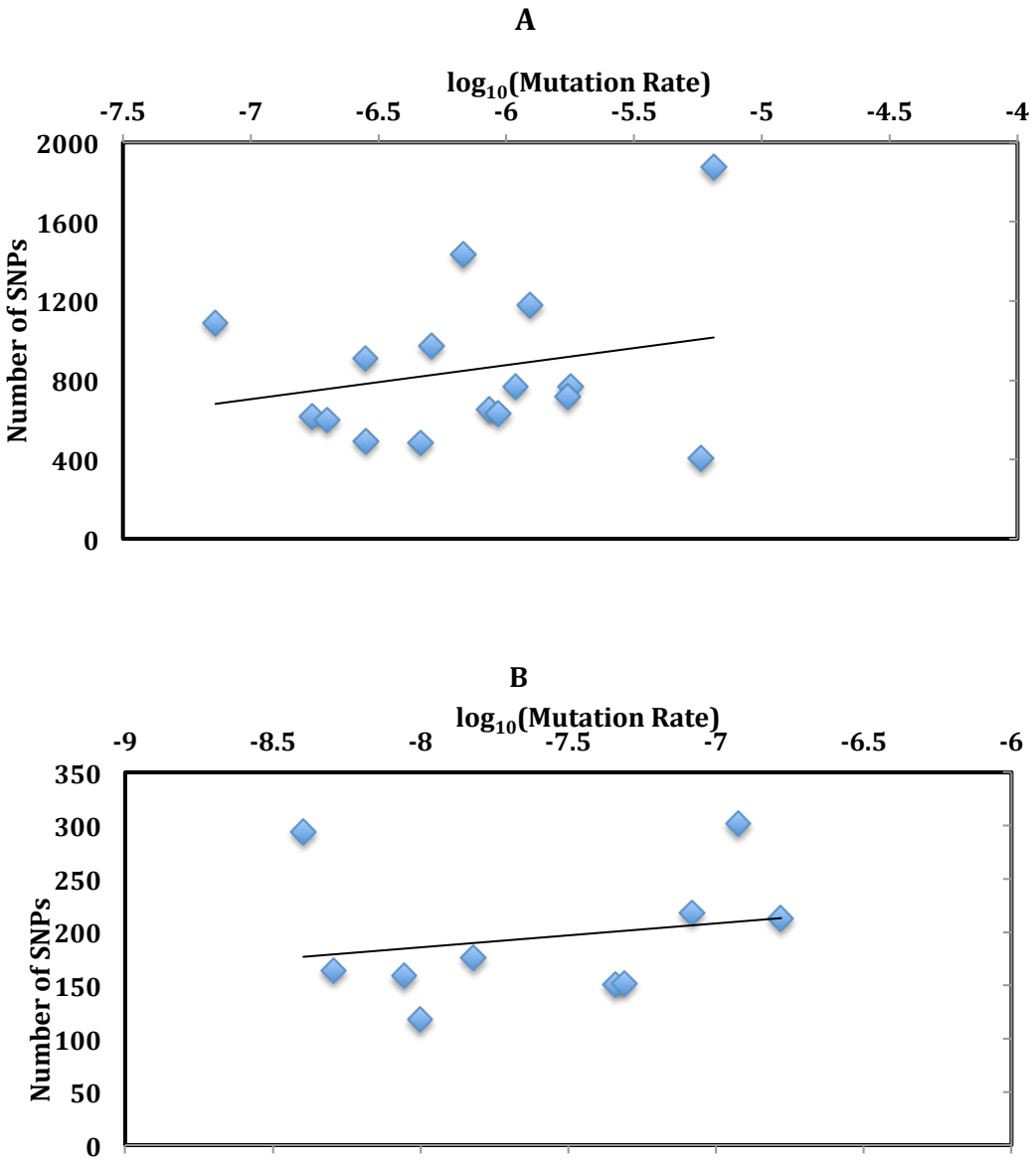
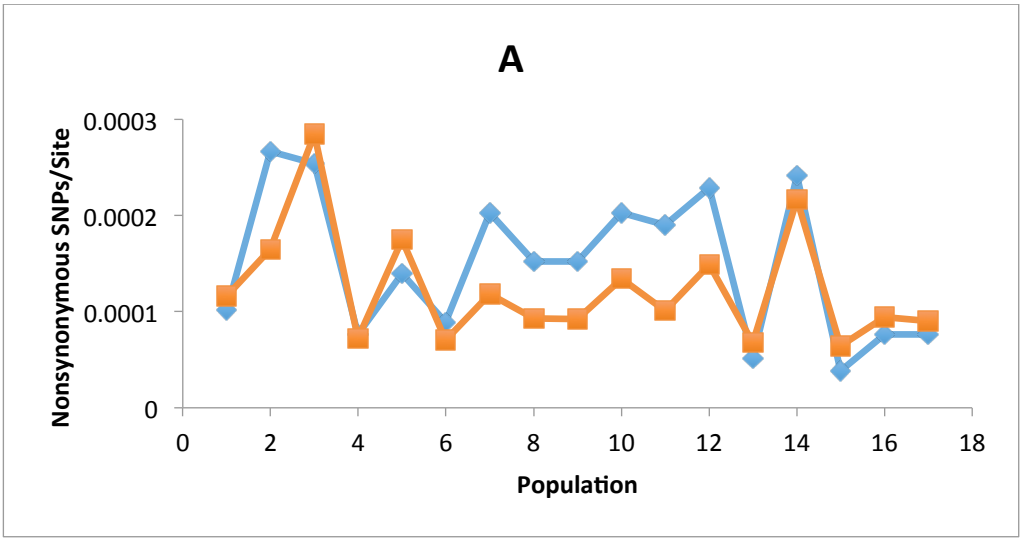


Figure 3.1 Total number of SNPs vs. end point mutation rates for both the LB (A) and MG (B) populations. The correlation between these variables is not significant for both LB and MG populations (p>0.05)

1. Number of non-synonymous substitutions per site



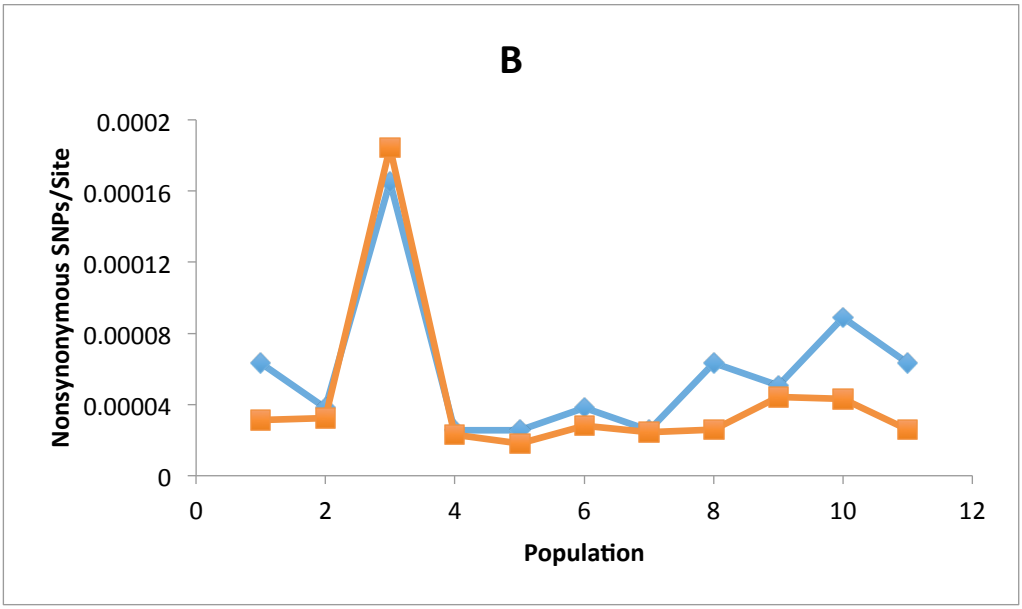


Figure 3.2 The number of non-synonymous substitutions per site for genes that are implicated in mutation rate evolution, compared with the rest of the genome in LB (A) and MG (B) populations. The BLUE lines denote the number of non-synonymous SNPs per site for the mutation rate associated loci and the AMBER lines denote the rate of non-synonymous SNPs for the whole genome. The estimates of non-synonymous substitutions are similar in most populations.

2. SNPs in Mutation Rate Loci

After running the breseq pipeline, many non-synonymous SNPs were discovered in the genes involved in influencing mutation rate. In order to confirm that these inflated numbers of SNPs were not just an artifact of a higher mutation rate, I sampled loci at random from the *E. coli* K12 genome and measured the number of non-synonymous SNPs to generate a null distribution that was then compared with the distribution of SNPs in mutation rate associated loci (see figure 3.4). These distributions are significantly different for the LB populations ($p<0.05$), but not the MG populations, when tested using a Wilcoxon Mann-Whitney U test. The significant difference between the background distribution of non-synonymous mutations and the mutation rate loci distribution of non-synonymous mutations may imply that these loci accumulate mutations at a different rate than the rest of the genome.

By probing the functional category DNA replication in EcoCyc (Keseler et al. 2013; Keseler et al. 2011), I found a list of 55 loci implicated in mutation rate evolution (see Appendix). Assuming that synonymous SNPs were unlikely to have an effect on protein evolution, I decided to explore the non-synonymous SNPs as candidates for reduction in mutation rates. There were 200 non-synonymous SNPs in mutation rate loci discovered in the surviving LB populations and about 50 non-synonymous SNPs in the MG populations that survived (see Appendix for complete list of SNPs). Of these, the ones that were

most likely to be responsible for a reduction in mutation rate, and may have undergone positive selection, are SNPs that occur independently in different populations, indicating convergent evolution. These SNPs are listed below along with their putative function.

Figure 3.2 shows the number of non-synonymous substitutions per site between the 55 mutation rate loci (blue lines) and the rest of the genome (amber lines). As discussed above, some populations have very high rates of non-synonymous substitutions, (for example MG4, the third point in the Fig 3.2B). This population has an extremely low fitness and for that reason I was unable to estimate the endpoint mutation rate of this population. The large number of non-synonymous substitutions, which are likely to be mostly deleterious, might explain the extremely low fitness of this population.

I generated a distribution of the number of non-synonymous mutations by the genes that they occurred in for the LB and MG populations. Out of the 55 loci that were screened for mutations in mutation rate loci, 54 of the loci had accumulated mutations in LB and 27 of the loci had accumulated mutations in MG. The distributions are presented in Figure 3.3. From observing the frequencies in these distributions, I picked the genes that were most commonly mutated from both the LB and the MG lines and decided to explore them further as candidates for reduction in mutation rates.

Candidate SNPs in LB Populations

1. *nrdE* (20 instances)

This gene encodes a subunit of the ribonucleoside-diphosphate reductase. Briefly, its function is to provide the precursors required for DNA synthesis. As part of this function, it catalyzes the formation of deoxyribonucleotides from the respective ribonucleotide. This protein contains binding sites for substrates as well as other molecules that are required for the reaction. Most of the SNPs (nine out of 20) observed in this gene seen in my experiments seem to have occurred at amino acid position 412, which was originally aspartate, an electrically charged amino acid. Most of the substitutions at this position have been replacements with asparagine, which is very similar in structure to aspartate, but also contains a polar side chain. It is interesting to note that the position 412 is extremely close to a hydrogen binding site at 415 in *nrdE*, which is a well documented active site and hence it may aid in binding to effector molecules and thereby increasing the efficiency of the reductase.

2. *dinB* (14 instances)

This gene encodes a stress-induced DNA polymerase (pol IV) that is devoid of any proofreading activity and hence prone to more errors during DNA replication. As such, it can be expected to be a target for mutations that might change the mutation rate, and it has already been implicated in lower mutation rates

(McKenzie et al. 2000). Most of the mutations that occurred in the *dinB* gene in my experiment were at amino acid positions 117-122 which are not in close proximity to the known active sites. However, a mutation that occurred in LB3, at position 103, which changed an Alanine to a Aspartate, has been previously documented as a knock out mutation (Wagner et al. 1999) and may be involved in lowering the mutation rate. The mutation rate of LB3 is significantly lower than that of the ancestor (see Chapter 2).

3. *dnaE* (10 instances)

This gene encodes the α subunit of DNA polymerase III, which is the most widely used DNA polymerase in *E coli.* It has an asymmetric dimeric structure that consists of 10 subunits. The core sub-units that are essential for DNA replication and proofreading are α, ε and θ. Mutations in this gene have previously been implicated in lower mutation rates in a study in which seven antimutator mutations were discovered (Fijalkowska and Schaaper 1993). Curiously, all the antimutator mutations that were discovered in their study were not concentrated in one region of the protein sequence but scattered all over the sequence. Incidentally, this seems true of the non-synonymous mutations that I observed in my populations as well, with mutations ranging from codons at position 114 to 926.

Candidate SNPs in MG populations

1. *nrdE* (5 instances)

Interestingly, this gene was also seen to have many non-synomynous SNPs across independent LB populations. It was mentioned that in the LB populations, most of the non-synonymous mutations occurred near an effector molecule binding site, which could potentially have some consequences for the efficacy of the reductase.

It is also extremely interesting to note that all of the 5 non-synonymous mutations that have occurred in this gene in the MG populations are identical. They occur at the same codon position (412) and involve an aspartate mutating to an asparagine, which was the dominant change noted in the LB populations as well. The convergent nature of this non-synonymous mutation not only across replicates under the same environment but also across different environments is suggestive; its role (if any) in reducing the genomic mutation rate could be tested in future work by carrying out allele replacement where a wild type copy of the gene *nrdE* would be inserted into the evolved strain and the effect on mutation rate would be investigated.

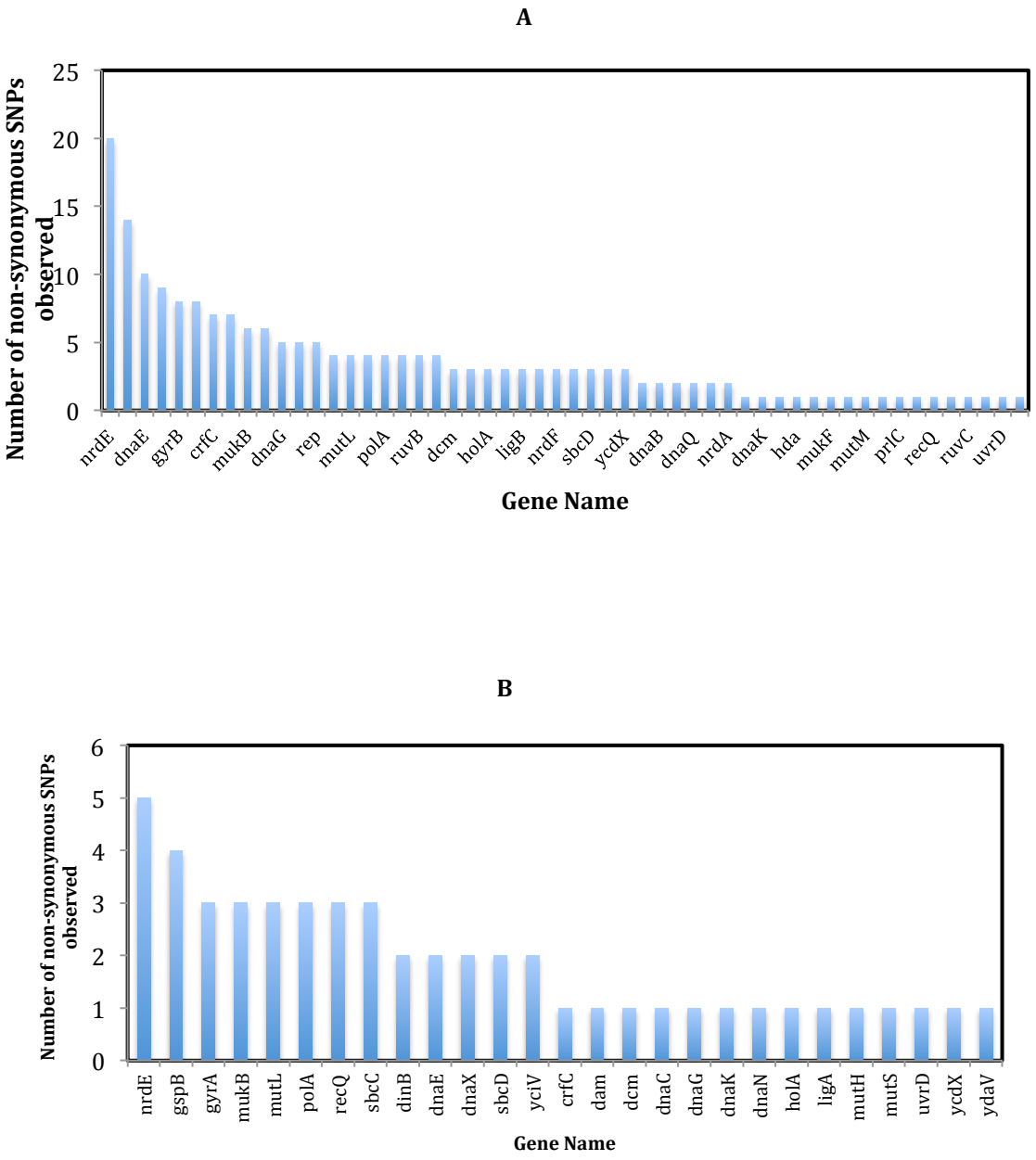


Figure 3.3 The distribution of non-synonymous SNPs that occur in mutation rate loci in the LB and MG MA populations, categorized by the genes in which they occur. In both instances, *nrdE* seems to be the gene with most non-synonymous substitutions.

2. *gspB* (4 instances)

This gene produces a calcium binding protein that is predicted to be involved with the initiation of DNA replication. It was first described by Guzman et al. in 1991 (Guzman, Pritchard, and Jimenez-Sanchez 1991) and has since then not received a lot of attention. It is a short protein of only 139 amino acids and it has a well described transmembrane domain that is encoded by the 24-48 position amino acids. However, very interestingly, all the populations that have accumulated non-synonymous mutations in this gene have substituted these mutations at the same codon position, i.e., position 97 from an aspartate to a glycine in all four cases. This is a change that may potentially have a strong effect since the aspartate is an electrically charged amino acid and may aid in binding with other molecules whereas the glycine is uncharged and the smallest amino acid.

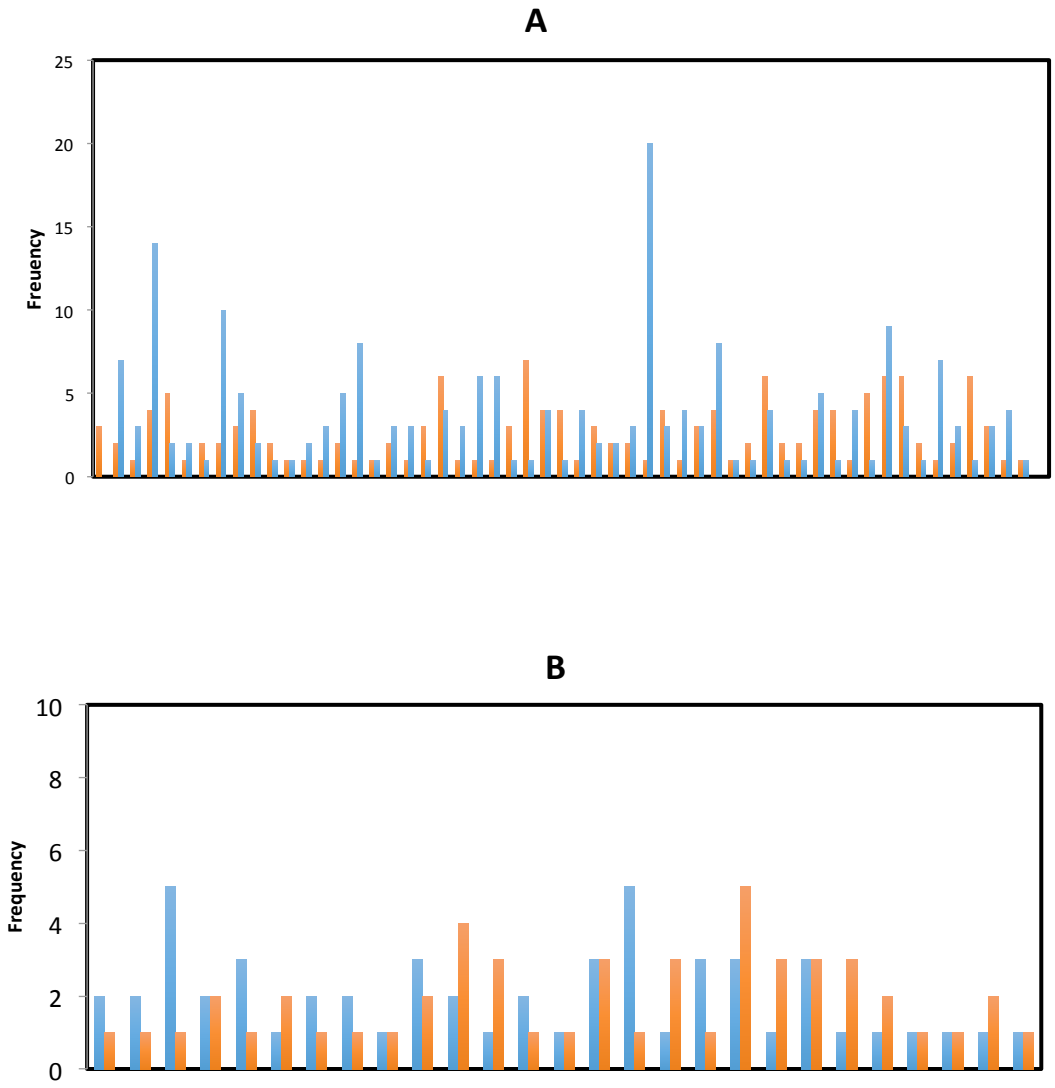


Figure 3.4 The distribution of non-synonymous SNPs that occur in the LB (A) and MG (B) MA populations, categorized by the whether they occur in mutation rate loci (BLUE) or randomly selected loci from anywhere in the genome (AMBER). The distributions in A are significantly different ($p<0.05$), but the distributions in B are not significantly different.

Discussion

In the previous chapter, I presented the phenotypic observation of reduced mutation rates in a short-term mutation accumulation experiment that was carried out with a hypermutable strain of *E. coli*. This observation seemed unexpected at first, since mutation rate is often observed to evolve to higher values in experimental populations and seldom seen to decline (reviewed in Raynes and Sniegowski 2014). Mutation rate is believed to evolve via indirect selection, implying that mutation rate modifier alleles do not have a direct effect on fitness, but may impact fitness indirectly by increasing or decreasing the likelihood of acquiring beneficial mutations or deleterious mutations. Even though most instances of mutation rate evolution have been evolution of higher mutation rates, there are a few instances of reduction in mutation rates (McDonald et al. 2012; Wielgoss et al. 2012). In the absence of beneficial mutations, the indirect advantage of a mutator allele might be severely diminished and hence mutators may be disfavored (Gerrish et al. 2007, simulation results). In fact, in the absence of beneficial mutations, most asexual lineages are expected to decline in fitness and eventually go extinct (Lynch et al. 1993), a process which is accelerated by the presence of a high mutation rate (Bull and Wilke 2008; Bull, Sanjuán, and Wilke 2007). The fact that I observed multiple extinctions in my MA experiment provides evidence to support these theories.

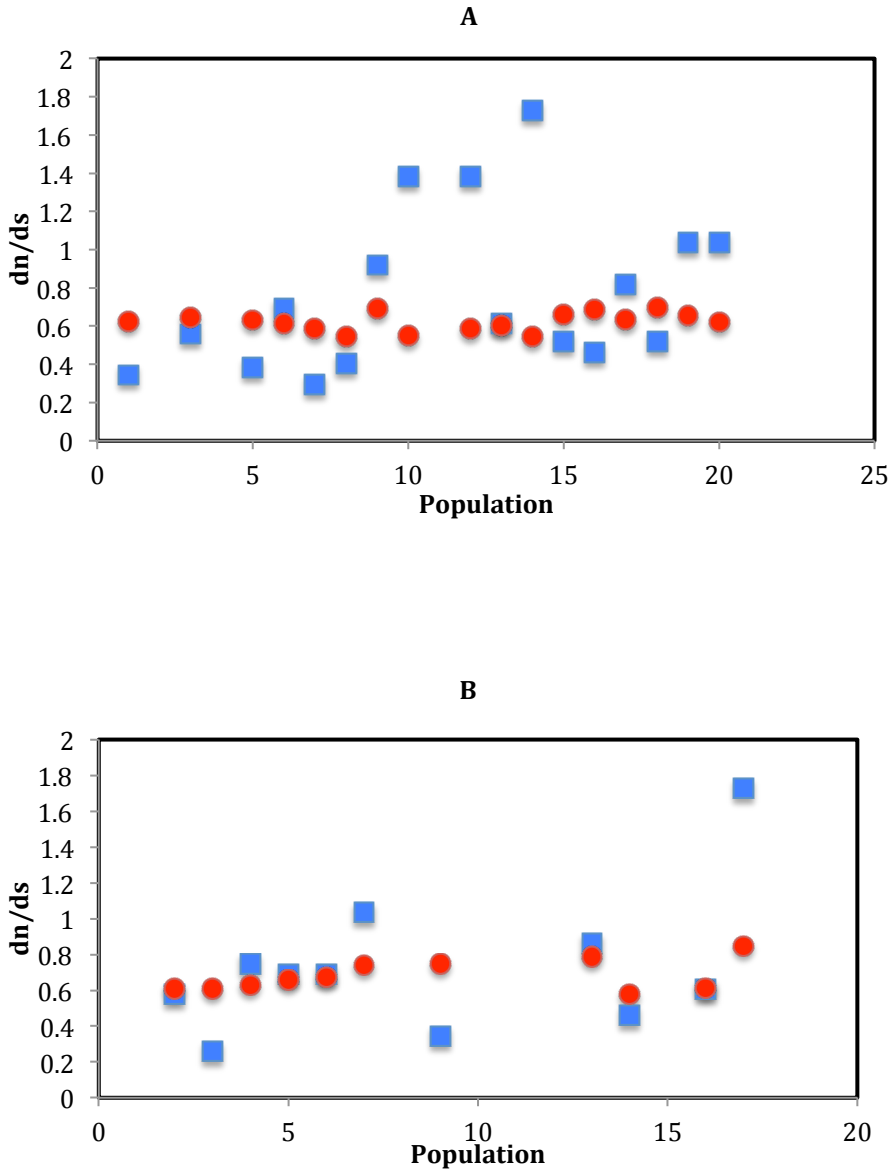


Figure 3.5 *dn/ds* ratios for LB (A) and MG (B) MA populations, categorized by mutation rate loci (BLUE) or the entire genome (RED).

In my populations, I did not find any significant correlation between the number of SNPs and the end point mutation rate (Figure 3.1), although there seems to be a positive relationship between these variables. The absence of a significant correlation may be because mutation rate evolved later in the experiment, and hence the number of SNPs is reflective of the mutation rate before it changed. This could be a reasonable explanation given our hypothesis that mutation rate may evolve to evade excess deleterious load under high deleterious mutation rates (Chapter 2).

Using the SNP data from my experiments, I have identified potential candidate SNPs for lower mutation rate from over 250 SNPs that occurred in the mutation rate loci from the *E. coli* genome, in my MG and LB populations, based on extensive parallelism that has been observed. Usually, the *dn/ds* ratio is measured in order to identify genes that may be under positive selection. However, in cases where the rate of non-synonymous mutation is either too low or too high, *dn/ds* may be not be informative, and may in fact lead to false positives (Barrick and Lenski 2013). Since the mutation rate is extraordinarily high in our populations, the genetic mutations that actually caused the phenotypic change might not be identifiable due to the high background rate of mutations. I have estimated the *dn/ds* ratios using the ratio of the total number of

non-synonymous changes that have occurred in conjuction with the total number of synonymous substitutions that have occurred, and taking into account the total number of possible non-synonymous and synonymous changes that can occur given the genomic sequence. These ratios are displayed in figure 3.5 and there is a general trend towards a higher *dn/ds* for the mutation rate loci, however, this needs to be further substantiated by estimating the confidence intervals of the genome wide *dn/ds* ratios.

In addition to identifying potential candidate mutations, I have described the specific biochemical process these substitutions might alter in order to understand how these substitutions may influence the mutation rate. The frequency of non-synonymous mutations in mutation rate loci observed in my data seems to be higher than the background frequency of non-synonymous mutations as evidenced by Figure 3.4. However, phenotypic studies are needed to test the hypothesis that these candidate mutations are responsible for the reduced mutation rates.

References

Barrick, J E., and R. E. Lenski. 2013. *Genome Dynamics during Experimental Evolution*. Vol. 48. doi:10.1097/MPG.0b013e3181a15ae8.Screening.

Barrick, J. E., and R. E. Lenski. 2009. “Genome-Wide Mutational Diversity in an Evolving Population of Escherichia Coli.” *Cold Spring Harbor Symposia on Quantitative Biology* 74: 119–29. doi:10.1101/sqb.2009.74.018.

Buermans, H.P.J., and J.T. den Dunnen. 2014. “Next Generation Sequencing Technology: Advances and Applications.” *Biochimica et Biophysica Acta*

(BBA) - Molecular Basis of Disease 1842 (10). Elsevier B.V.: 1932–41. doi:10.1016/j.bbadis.2014.06.015.
Bull, J. J., Rafael Sanjuán, and Claus O. Wilke. 2007. “Theory of Lethal Mutagenesis for Viruses.” *Journal of Virology* 81 (6): 2930–39. doi:10.1128/JVI.01624-06.
Bull, James J., and Claus O. Wilke. 2008. “Lethal Mutagenesis of Bacteria.” *Genetics* 180 (2): 1061–70. doi:10.1534/genetics.108.091413.
Cooper, Tim F, Daniel E Rozen, and Richard E Lenski. 2003. “Parallel Changes in Gene Expression after 20,000 Generations of Evolution in Escherichiacoli.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (3): 1072–77. doi:10.1073/pnas.0334340100.
Deatherage, Daniel E, and Jeffrey E Barrick. 2014. “Identification of Mutations in Laboratory Evolved Microbes from next-Generation Sequencing Data Using Breseq.” *Methods in Molecular Biology (Clifton, N.J.)* 1151: 165–88. doi:10.1007/978-1-4939-0554-6_12.
Fijalkowska, IJ, and RM Schaaper. 1993. “Mutations in the Alpha Subunit of Escherichia Coli DNA Polymerase III: Identification of the Responsible Mutations and Alignment with Other DNA Polymerases.” *Genetics* 134: 1039–44.
Gerrish, Philip J, Alexandre Colato, Alan S Perelson, and Paul D Sniegowski. 2007. “Complete Genetic Linkage Can Subvert Natural Selection.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (15): 6266–71. doi:10.1073/pnas.0607280104.
Guzman, EC; Pritchard, R H, and Jiminez-Sanchez, A. 1991. “A Calcium-Binding Protein That May Be Required for the Initiation of Chromosome Replication in Escherichia Coli.” *Research in Microbiology* 142 (2-3): 137–40.
Ishizawa, Yuuka, Bei-Wen Ying, Saburo Tsuru, and Tetsuya Yomo. 2015. “Nutrient-Dependent Growth Defects and Mutability of Mutators In\n Escherichia Coli.” *Genes to Cells* 20 (1): 68–76. doi:10.1111/gtc.12199.
Kao, Katy C, and Gavin Sherlock. 2008. “Molecular Characterization of Clonal Interference during Adaptive Evolution in Asexual Populations of Saccharomyces Cerevisiae.” *Nature Genetics* 40 (12): 1499–1504. doi:10.1038/ng.280.
Keseler, Ingrid M., Julio Collado-Vides, Alberto Santos-Zavaleta, Martin Peralta-Gil, Socorro Gama-Castro, Luis Muniz-Rascado, César Bonavides-Martinez, et al. 2011. “EcoCyc: A Comprehensive Database of Escherichia Coli Biology.” *Nucleic Acids Research* 39 (SUPPL. 1): 583–90. doi:10.1093/nar/gkq1143.
Keseler, Ingrid M., Amanda Mackie, Martin Peralta-Gil, Alberto Santos-Zavaleta, Socorro Gama-Castro, César Bonavides-Martínez, Carol Fulcher, et al. 2013. “EcoCyc: Fusing Model Organism Databases with Systems Biology.”

Nucleic Acids Research 41 (D1): 605–12. doi:10.1093/nar/gks1027.
Kinnersley, Margie, Jared Wenger, Evgueny Kroll, Julian Adams, Gavin Sherlock, and Frank Rosenzweig. 2014. “Ex Uno Plures: Clonal Reinforcement Drives Evolution of a Simple Microbial Community.” *PLoS Genetics* 10 (6). doi:10.1371/journal.pgen.1004430.
Langmead, Ben, and Steven L Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nat Methods* 9 (4). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 357–59.
Lynch, M, R Bürger, D Butcher, and W Gabriel. 1993. “The Mutational Meltdown in Asexual Populations.” *The Journal of Heredity* 84 (5): 339–44.
McDonald, Michael J., Yu Ying Hsieh, Yen Hsin Yu, Shang Lin Chang, and Jun Yi Leu. 2012. “The Evolution of Low Mutation Rates in Experimental Mutator Populations of Saccharomyces Cerevisiae.” *Current Biology* 22 (13): 1235–40. doi:10.1016/j.cub.2012.04.056.
McKenzie, G J, R S Harris, P L Lee, and S M Rosenberg. 2000. “The SOS Response Regulates Adaptive Mutation.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (12): 6646–51. doi:10.1073/pnas.120161797.
Michaels, M L, C Cruz, and J H Miller. 1990. “mutA and mutC: Two Mutator Loci in Escherichia Coli That Stimulate Transversions.” *Proceedings of the National Academy of Sciences of the United States of America* 87 (23): 9211–15. doi:10.1073/pnas.87.23.9211.
Ohta, Tomoko, and M Kimura. 1971. “On the Constancy of the Evolutionary Rate of Cistrons 25 (813).
Raynes, Y, and P D Sniegowski. 2014. “Experimental Evolution and the Dynamics of Genomic Mutation Rate Modifiers.” *Heredity*, no. April: 1–6. doi:10.1038/hdy.2014.49.
Robertson, A. 1960. “A Theory of Limits in Artificial Selection.” *Proceedings of the Royal Society of London B: Biological Sciences* 153 (951): 234–49.
Shaver, a C, and P D Sniegowski. 2003. “Spontaneously Arising mutL Mutators in Evolving Escherichia Coli Populations Are the Results of Changes in Repeat Length.” *J. Bacteriol.* 185 (20): 6076–82. doi:10.1128/JB.185.20.6076.
Sniegowski, Paul D, Philip J Gerrish, and Richard E Lenski. 1997. “Evolution of High Mutation Rates in Experimental Populations of E . Coli.” *Nature* 387: 703–5. doi:10.1128/AEM.02595-09.
Wagner, Jérôme, Petr Gruz, Su Ryang Kim, Masami Yamada, Keiko Matsui, Robert P P Fuchs, and Takehiko Nohmi. 1999. “The dinB Gene Encodes a Novel E. Coli DNA Polymerase, DNA Pol IV, Involved in Mutagenesis.” *Molecular Cell* 4 (2): 281–86. doi:10.1016/S1097-2765(00)80376-7.
Wielgoss, Sébastien, Jeffrey E Barrick, Olivier Tenaillon, Michael J Wiser, W James Dittmar, Stéphane Cruveiller, Béatrice Chane-Woon-Ming, Claudine

Médigue, Richard E. Lenski, and Dominique Schneidera. 2012. “Mutation Rate Dynamics in a Bacterial Population Re Fl Ect Tension between Adaptation and Genetic Load.” *Proc Natl Acad Sci (USA)* 110 (1): 222–27. doi:10.5061/dryad.hb3b5.

CHAPTER 4

The effect of deleterious mutations on neutral mutations in hypermutable populations.

Introduction

Many mutations are neutral i.e., they do not affect the fitness of the organism. However, most mutations that do have an effect on fitness tend to be deleterious (Eyre-Walker and Keightley 1999). The deleterious mutation rate of a population is a function of the overall genomic mutation rate, which is an evolutionarily pliable trait. In addition, because effective population size determines whether selection will influence the fate of a mutation (Wright 1931; Ohta and Kimura 1971), the fraction of mutations that are deleterious is directly correlated with effective population size: a higher effective population size translates to a higher deleterious mutation rate. Although not enough is known about the evolutionary history of microbial populations to infer their effective population sizes in many cases (Mes 2008), most are likely to have large census sizes, especially compared to those of multicellular organisms (Lynch 2007). Thus, it is reasonable to assume that a large fraction of mutations in microbial populations are deleterious. Moreover, microbial populations--and, theoretically, asexual populations in general--tend to substitute mutator alleles by genetic hitchhiking,

which will further increase their deleterious mutation rate (Chapter 1). Although considerable attention has been paid to the role of deleterious mutations under genetic linkage (Charlesworth, Morgan, and Charlesworth 1993; Birky and Walsh 1988; Barton 2010; Wright and Andolfatto 2008; Keightley and Otto 2006), the effects of their accumulation on the spread of lineages bearing neutral and beneficial mutations have not been directly investigated and are the focus of the research described in this chapter.

In general, the effects of selection on linked loci were first observed and quantified by Hill and Robertson (Hill and Robertson 1966) who showed that a reduction in effective population size occurs at a locus that is linked to another locus under directional selection. In asexual populations, this effect can lead to substantial changes in allele frequency dynamics due to pervasive natural selection favoring beneficial mutations or purging deleterious mutations. Following from the approach of Hill and Robertson, most investigations of the influence of deleterious mutations in such processes have focused on their role in contaminating existing genetic backgrounds in an asexual population or linked genetic region (Hill and Robertson 1966; Charlesworth, Morgan, and Charlesworth 1993; Campos and Wahl 2010; Charlesworth 2012b; Charlesworth 2012a). For example, in the background selection model of Charlesworth et al. (Charlesworth, Morgan, and Charlesworth 1993), a population or genome region with complete genetic linkage essentially has its effective size reduced by a

factor f_0, where f_0 is the fraction of the population that does not carry any deleterious mutations. Charlesworth et al. reason that this is because only the descendants of the f_0 fraction will survive and contribute to future populations, whereas the others will be lost sooner or later. This reduction in effective population size reduces genetic variation and decreases the probability of fixation of beneficial mutations.

In the present study, I use simulations to examine the effect that the accumulation of deleterious mutations has on a lineage that initially starts with *zero* deleterious mutations. This effect has been tentatively called "lineage contamination" (Gerrish et al. 2016, in preparation). Lineage contamination is clearly distinct from background selection, although both effects are likely to have been occurring simultaneously in some previous simulation work (Johnson and Barton 2002; Peck 1994). My investigation of lineage contamination was stimulated by theoretical (Pénisson et al. 2013; Bull and Wilke 2008; Bull, Sanjuán, and Wilke 2007) and experimental (Gentile et al. 2011) work on the fate of fitness in populations with very high mutation rates--hypermutable populations. I hypothesize that the distribution of a neutral (and, by extension, beneficial) mutant in a hypermutable asexual population is distorted by deleterious mutations that accumulate differentially in the small subpopulation represented by the neutral mutation. I test this hypothesis using a computer simulation approach.

Methods

To investigate the effect of deleterious mutations on neutral mutant lineages, I designed computer simulations to simulate the exponential growth of an asexual population in a process similar to the classical fluctuation assay used to estimate mutation rates experimentally (Luria and Delbrück 1943). The fluctuation assay was first used by Luria and Delbrück to determine if the emergence of resistance to phage in bacteria was a product of random mutations or a result of induction by the phage. The random mutation hypothesis and the induced mutagenesis hypothesis led to different predictions regarding the shape of the final distribution of the desired resistance mutants. The distribution under random mutagenesis came to be known as the Luria-Delbrück distribution, and the fluctuation assay has become widely used for estimation of mutation rates. The assay itself is simple: a small number of clonal individuals is inoculated into replicate cultures of identical medium in order to undergo exponential growth. During this period, the mutation of interest (such as antibiotic resistance) will arise and ideally will be effectively neutral, because the selective agent is absent at the time of growth. At the end of the growth period, the number of such mutants is estimated in each culture by exposing the cultures to medium that is supplemented with the selective agent, and the distribution of the number of mutants across replicates is used to estimate the mutation rate.

In the current study, I was interested in investigating whether the accumulation of deleterious mutations linked to neutral mutants could affect the observed distribution of neutral mutants in the fluctuation test setting. In a real fluctuation test, of course, the experimenter cannot vary at will such parameters as the deleterious mutation rate and the distribution of effects of deleterious mutations; however, a simulation approach allows such manipulation. In the simulations employed for this study, I varied the neutral mutation rate independent of the genomic mutation rate, making it possible to study the effect of genomic mutation rate and more specifically deleterious mutations on the distribution of neutral mutations. I also varied the mean effect size of deleterious mutations in order to examine the effect of deleterious mutations on the fitness of individuals. The deleterious mutations in the simulations are drawn from a gamma distribution (Loewe et al. 2006; Piganeau and Eyre-Walker 2003), for which two different combinations for shape and scale parameters are used, in order to simulate a higher and lower mean effect size of deleterious mutations. The parameter values utilized for mutation rates are based on mutation accumulation studies in wild type *E. coli* (Lynch et al. 1999). I chose a particular deleterious mutation rate for the higher mutation rate simulations based on previous experimental studies in a hypermutable *E. coli* strain (Gentile et al. 2011). The two different values of mean effect size of deleterious mutations were selected in order to simulate two contrasting conditions under which there would be strong purifying selection

against deleterious mutations (s_d = 0.03) and lack thereof (s_d = 0.001). Deleterious mutations were drawn from a gamma distribution that best fits existing data on deleterious mutations (Gerrish et al. 2007; Piganeau and Eyre-Walker 2003; Eyre-Walker and Keightley 2007), although in some cases exponential distributions are used in the literature. Beneficial mutations are not included in these simulations, as it unlikely that beneficial mutations will arise and go to appreciable frequencies in such a short timescale. The mutation rate to the specific neutral mutation of interest is termed the focal mutation rate. The focal mutation rate value that I selected was based on the mutation rate to nalidixic acid resistance (1.5×10^{-6}) as measured by me in a hypermutable *E. coli* strain (Gentile et al. 2011).

The simulations start from a single individual, and final population size is approximately 4×10^6 individuals after 22 generations of growth. Fitness of individuals is initialized to 1 at the start of the simulation and decreases as mutations accumulate. The number of offspring produced by an individual is Poisson distributed with the mean of the Poisson distribution being equal to the relative fitness of the individual. The simulation reports the final number of neutral mutants from every replicate at the end of the exponential growth period. This number is recorded for all replicates and used to generate a distribution of neutral mutants. These results were used to compare distributions of neutral mutants at both higher and lower deleterious mutation rates.

Table 4.1 gives an overview of the design of the simulation study.

Number of replicates	Deleterious Mutation Rates (U_d)	Mean effect size of deleterious mutation (s_d)	Comparison
500	0.9 0.002	0.03	Different deleterious mutation rates, under strong purifying selection
500	0.9 0.002	0.001	Different mutation rates, under weak purifying selection

Table 4.1 Design of Simulation Study

Results

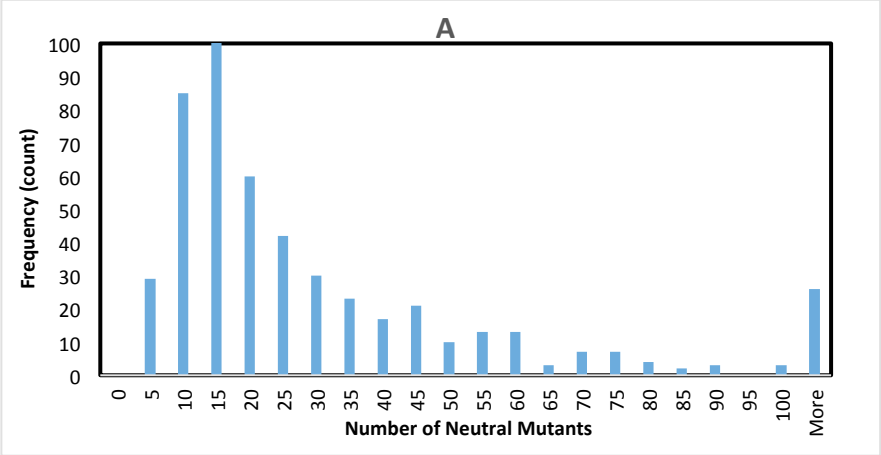
Table 4.2 shows the average number of neutral mutants and the variance in the number of mutants under the different combinations of parameter values. It is evident from the table that under strong purifying selection and a high deleterious mutation rate, the average number of neutral mutants and the variance among the number of mutants tended to decrease in my simulations.

Focal Mutation Rate	Deleterious Mutation Rate (U_d)	Mean Effect Size of Deleterious Mutation (s_d)	Average Number of mutants	Variance of Number of mutants
1.5×10^{-6}	0.9	0.03	42.07	14312.55
1.5×10^{-6}	0.0002	0.03	141.8	4074785.83
1.5×10^{-6}	0.9	0.001	103.37	1288999.4
1.5×10^{-6}	0.0002	0.001	69.76	126644.04

Table 4.2. Summary of Simulation Data

An effective way of visualizing the effect of deleterious mutations on neutral lineages is to study the distribution of neutral mutants at the end of exponential growth. Figures 4.1 and 4.2 present the output of my simulations as such a visualization. A visual comparison of the distributions under different mutation rates given strong purifying selection (high s_d), suggests that number of "jackpots" (defined as any replicate with greater than 100 neutral mutants) is severely reduced at higher deleterious mutation rates, consistent with the hypothesis that overload of deleterious mutations may remove some neutral lineages from the population (Figure 4.1). This visually striking effect was missing when the effect size of deleterious mutations was kept negligible (Figure 4.2). To assess whether the A and B distributions shown in Figures 4.1 and 4.2 are significantly different from each other, I used a Kolmogorov-Smirnov test, which uses a cumulative distribution approach to determine if two samples come from the same underlying distribution. The results from the Kolmogorov-Smirnov test are shown in Table 4.3. The D statistic represents the maximum vertical deviation between two cumulative frequency plots of the two distributions that are being compared. In the comparison, the D statistic is greater than the critical D statistic at a p value cut-off of 0.001 for $s_d = 0.03$, but not for $s_d = 0.001$, indicating that we can reject the null hypothesis that the two distributions come from the same underlying distribution in the former case. These results indicate that even moderately sized deleterious mutations can substantially change the

shape of the distribution of neutral mutants in a growing hypermutable population.



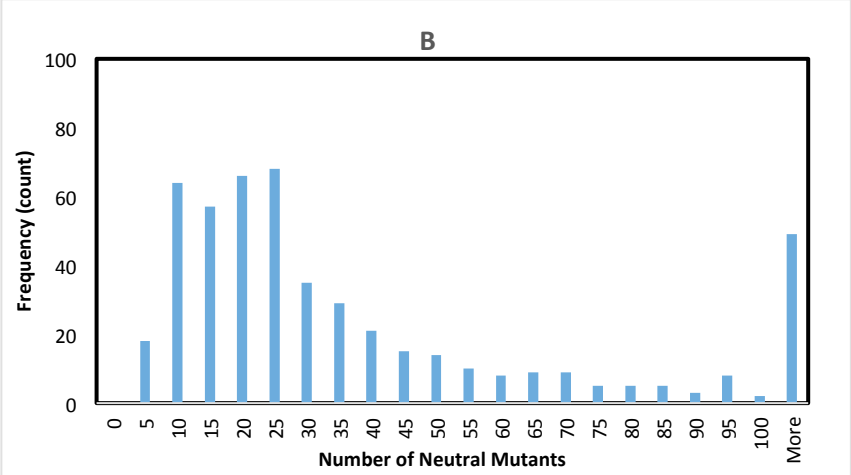
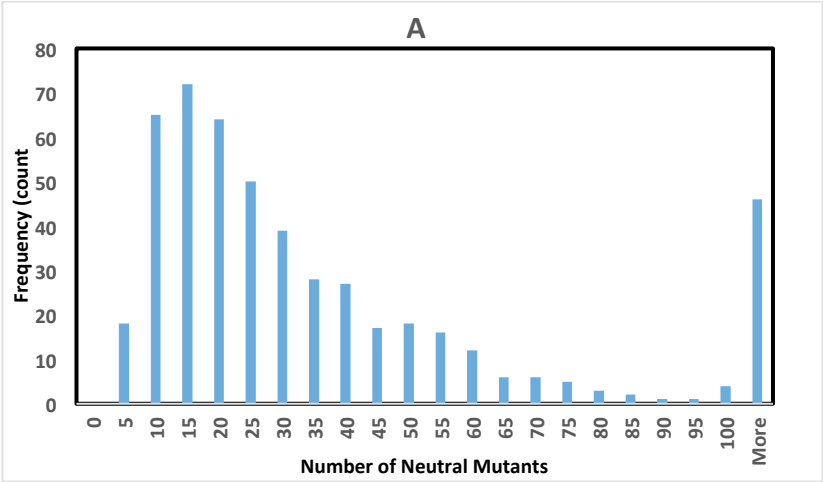


Figure 4.1 Distributions of the number of neutral mutants at the end of the exponential growth period in the simulations for both high (A) and low (B) deleterious mutation rates, given a moderately strong average effect of deleterious mutations (s_d = 0.03). These distributions are significantly different under the Kolmogorov-Smirnov test (p < 0.001).



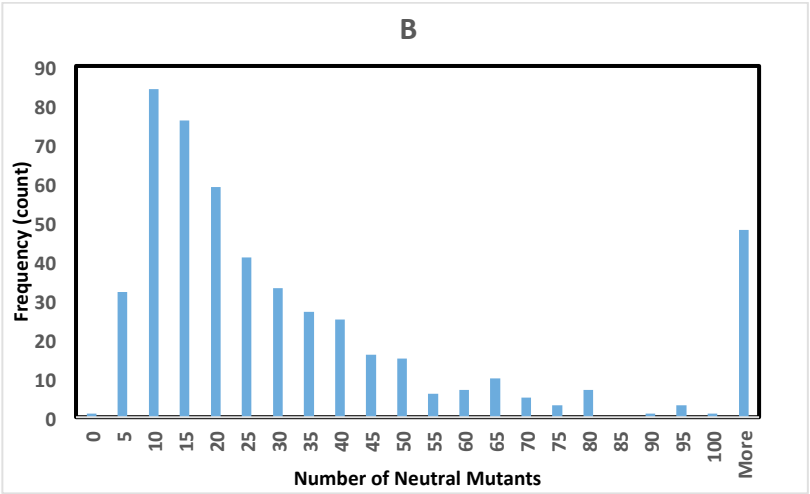


Figure 4.2 Distributions of the number of neutral mutants at the end of the exponential growth period in the simulations for both high (A) and low (B) deleterious mutation rates, under a weak individual effect of deleterious mutations (s_d = 0.001). These distributions are not statistically different under the Kolmogorov-Smirnov test.

Mean Effect Size of Deleterious Mutation, s_d	**$D_{statistic}$**	**$D_{critical}$**
0.001	0.076	0.1233
0.03	0.154***	0.1233

Table 4.3. Results of Kolmogorov-Smirnov test. *** indicates a significant difference (p < 0.001).

Discussion

Using computer simulations, I have shown that the size distribution of neutral mutant lineages in an exponentially growing asexual population can be substantially affected by the accumulation of linked deleterious mutations. In particular, for an experimentally realistic combination of genomic deleterious mutation rate and mean effect of deleterious mutations, the average and variance in the number of neutral mutant lineages (Table 4.2) and the number of jackpots (Figure 4.1A) are decreased by the presence of linked deleterious mutations.

It has long been known that deleterious mutations can accumulate stochastically in a finite population. In the absence of genetic recombination, this process is accelerated and can lead to a monotonous decline in fitness if beneficial mutations are very rare or absent. This phenomenon was first described by Muller (Muller 1964), and subsequently termed “Muller’s Ratchet” by Felsenstein (Felsenstein 1974). The phenomenon of Muller’s Ratchet implies the loss of the least loaded class of individuals and thus can only be accelerated as the mutation rate is increased (Gessler, 1995). Moreover, because deleterious mutations can escape selection if their effects are smaller than the reciprocal of effective population size (Wright, 1931; Ohta and Kimura 1971), decline in fitness and concomitant decrease in population size can lead to a positive feedback

effect that culminates in a “mutational meltdown” of a population (Lynch et al. 1993).

Both the mutational meltdown model and the more recent “lethal mutagenesis” model for the extinction for asexual populations (Bull, Sanjuán, and Wilke 2007; Bull and Wilke 2008) consider populations in which beneficial mutations (including compensatory mutations) and reversions are absent. Recent theoretical (Bachtrog and Gordo, 2004; Poon and Otto, 2000) and empirical (Silander, Tenaillon, and Chao 2007; McDonald et al. 2012) work, however, indicates that the presence of such beneficial mutations can stall or even reverse the loss of fitness in finite asexual populations. A more complete model of the potential for high mutation rates to drive populations extinct thus requires consideration of beneficial mutations.

The work presented in this chapter was stimulated by considering the potential for increased accumulation of deleterious mutations in a small subpopulation relative to the majority background in an asexual population. Such a small subpopulation could be represented by the lineage of a neutral or even a beneficial mutation. Ongoing analytical work (Gerrish et al. 2016, in preparation) considers the effect of differential accumulation of deleterious mutations on the lineage of a beneficial mutation in an asexual population (“lineage contamination”). Here, I have used simulations to show that deleterious

mutations can indeed affect the spread of a neutral lineage. In other simulations (not shown) I have shown that the spread of an isolated beneficial mutation in a population can be similarly inhibited, and ultimately lower the probability of fixation of adaptive alleles. It is hence of some interest to ask if this effect can be demonstrated in experimental populations. In some preliminary experimental work, I attempted to test if a known beneficial mutation is inhibited in its spread in populations with high genomic mutation rates relative to those with low mutation rates. Although there was some evidence in favor of this hypothesis in these experiments, an alternative explanation based on increased clonal interference (Gerrish and Lenski 1998; Park and Krug 2007) among multiple beneficial mutations at high mutation rates could not be ruled out. Future work using an isolated beneficial mutation in a highly adapted background (to minimize clonal interference) may provide more informative results.

The results I have presented in this chapter may also have some implications for the interpretation of fluctuation assays. A key assumption of the fluctuation assay is that the mutation of interest is itself neutral (Luria and Delbrück 1943; Lea and Coulson 1949) Violations of this assumption can lead to a higher mean mutant count in the assay if the mutation is beneficial and a lower mean mutant count and reduced variance if the mutation is deleterious, with concomitant effects on the mutation rate calculated from the data. My simulations show that at a high background deleterious mutation rate, even an intrinsically neutral mutation can

behave as though it is deleterious. This suggests that mutation rates estimated in fluctuation assays on hypermutable strains may in fact be underestimates. It would be interesting to see whether this prediction is borne out in future experimental work.

References

Bachtrog D, Gordo I (2004) Adaptive evolution of asexual populations under Muller's ratchet. Evolution 58: 1403–1413.

Barton, N H. 2010. “Genetic Linkage and Natural Selection.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 365 (1552): 2559–69. doi:10.1098/rstb.2010.0106.

Birky, C W, and J B Walsh. 1988. “Effects of Linkage on Rates of Molecular Evolution.” *Proceedings of the National Academy of Sciences of the United States of America* 85 (17): 6414–18. doi:10.1073/pnas.85.17.6414.

Bull, J J, R Sanjuán, and C O Wilke. 2007. “Theory of Lethal Mutagenesis for Viruses.” *Journal of Virology* 81 (6): 2930–39. doi:10.1128/JVI.01624-06.

Bull, James J., and Claus O. Wilke. 2008. “Lethal Mutagenesis of Bacteria.” *Genetics* 180 (2): 1061–70. doi:10.1534/genetics.108.091413.

Campos, Paulo R a, and Lindi M. Wahl. 2010. “The Adaptation Rate of Asexuals: Deleterious Mutations, Clonal Interference and Population Bottlenecks.” *Evolution* 64 (7): 1973–83. doi:10.1111/j.1558-5646.2010.00981.x.

Charlesworth, B, M T Morgan, and D Charlesworth. 1993. “The Effects of Deleterious Mutations on Neutral Molecular Varaition.” *Genetics* 134 (4): 1289–1303.

Charlesworth, Brian. 2012a. “The Effects of Deleterious Mutations on Evolution at Linked Sites.” *Genetics* 190 (1): 5–22. doi:10.1534/genetics.111.134288.

———. 2012b. “The Role of Background Selection in Shaping Patterns of Molecular Evolution and Variation: Evidence from Variability on the Drosophila X Chromosome.” *Genetics* 191 (1): 233–46. doi:10.1534/genetics.111.138073.

Eyre-Walker, Adam, and Peter D Keightley. 1999. “High Genomic Deleterious Mutation Rates in Hominids.” *Nature* 397 (6717): 344–47.

———. 2007. “The Distribution of Fitness Effects of New Mutations.” *Nature Reviews. Genetics* 8 (8): 610–18. doi:10.1038/nrg2146.

Felsenstein, J. 1974. “The Evolutionary Advantage of Recombination. Genetics.” *Genetics*.

Gentile, Christopher F, Szi-Chieh Yu, Sebastian Akle Serrano, Philip J Gerrish,

and Paul D Sniegowski. 2011. “Competition between High- and Higher-Mutating Strains of Escherichia Coli.” *Biology Letters* 7 (3): 422–24. doi:10.1098/rsbl.2010.1036.
Gerrish, Philip J, Alexandre Colato, Alan S Perelson, and Paul D Sniegowski. 2007. “Complete Genetic Linkage Can Subvert Natural Selection.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (15): 6266–71. doi:10.1073/pnas.0607280104.
Gerrish, Philip J, and Richard E Lenski. 1998. “The Fate of Competing Beneficial Mutations in an Asexual Population S N Ln N :,” 127–44.
Gerrish, Philip, Sophie Pénisson_, Tanya Singh, and Paul Sniegowski. 2016. “Dynamics and Fate of Beneficial Mutations under Lineage Contamination by Linked Deleterious Mutations. In Preparation.”
Gessler, D. D. n.d. “The Constraints of Finite Size in Asexual Populations and the Rate of the Ratchet.” *Genet. Res.* 66 (2): 41–253.
Hill, W. G., and A Robertson. 1966. “The Effect of Linkage on Limits to Artificial Selection.” *Genetical Research* 8: 269–94.
Johnson, Toby, and Nick H. Barton. 2002. “The Effect of Deleterious Alleles on Adaptation in Asexual Populations.” *Genetics* 162 (1): 395–411.
Keightley, Peter D, and Sarah P Otto. 2006. “Interference among Deleterious Mutations Favours Sex and Recombination in Finite Populations.” *Nature* 443 (7107): 89–92. doi:10.1038/nature05049.
Lea, DE, and CA Coulson. 1949. “The Distribution of the Numbers of Mutants in Bacterial Populations.” *Genetics* 49: 264–85.
Loewe, Laurence, Brian Charlesworth, Carolina Bartolome, and Veronique Noel. 2006. “Estimating Selection on Non-synonymous Mutations.” *Genetics* 172 (2): 1079–92. doi:10.1534/genetics.105.047217.
Luria, S E, and M Delbrück. 1943. “Mutations of Bacteria from Virus Sensitivity to Virus Resistance.” *Genetics* 28 (6): 491–511. doi:10.1038/nature10260.
Lynch, M, J Blanchard, D Houle, T Kibota, S Schultz, L Vassilieva, and J Willis. 1999. “Spontaneous Deleterious Mutation.” *Evolution* 53 (3): 645–63.
Lynch, M, R Bürger, D Butcher, and W Gabriel. 1993. “The Mutational Meltdown in Asexual Populations.” *The Journal of Heredity* 84 (5): 339–44.
Lynch, M. 2007. *The Origins of Genome Architecture.* Sunderland:Sinauer Associates.
McDonald, Michael J., Yu Ying Hsieh, Yen Hsin Yu, Shang Lin Chang, and Jun Yi Leu. 2012. “The Evolution of Low Mutation Rates in Experimental Mutator Populations of Saccharomyces Cerevisiae.” *Current Biology* 22 (13): 1235–40. doi:10.1016/j.cub.2012.04.056.
Mes, Ted H M. 2008. “Microbial Diversity – Insights from Population Genetics.” *Environmental Microbiology* 10 (1). Blackwell Publishing Ltd: 251–64. doi:10.1111/j.1462-2920.2007.01449.x.
Muller, H.J. 1964. “Muller, H.J., 1964. The Relation of Recombination to

Mutational Advance. Mutat." *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* 1: 2–9.
Ohta, Tomoko, and M Kimura. 1971. "On the Constancy of the Evolutionary Rate of Cistrons *" 25 (813).
Park, Su-Chan, and Joachim Krug. 2007. "Clonal Interference in Large Populations." *Proceedings of the National Academy of Sciences of the United States of America* 104 (46): 18135–40. doi:10.1073/pnas.0705778104.
Peck, Joel R. 1994. "A Ruby I N the Rubbish: Beneficial Mutations, Deleterious Mutations."
Pénisson, Sophie, Paul D Sniegowski, Alexandre Colato, and Philip J Gerrish. 2013. "Lineage Dynamics and Mutation–selection Balance in Non-Adapting Asexual Populations." *Journal of Statistical Mechanics: Theory and Experiment* 2013 (01): P01013. doi:10.1088/1742-5468/2013/01/P01013.
Piganeau, Gwenaël, and Adam Eyre-Walker. 2003. "Estimating the Distribution of Fitness Effects from DNA Sequence Data: Implications for the Molecular Clock." *Pnas* 100 (18): 10335–40. doi:10.1073/pnas.1833064100.
Poon, a, and S P Otto. 2000. "Compensating for Our Load of Mutations: Freezing the Meltdown of Small Populations." *Evolution; International Journal of Organic Evolution* 54 (5): 1467–79. doi:10.1554/0014-3820(2000)054[1467:CFOLOM]2.0.CO;2
Robertson, A. 1960. "A Theory of Limits in Artificial Selection." *Proceedings of the Royal Society of London B: Biological Sciences* 153 (951): 234–49.
Silander, Olin K., Olivier Tenaillon, and Lin Chao. 2007. "Understanding the Evolutionary Fate of Finite Populations: The Dynamics of Mutational Effects." *PLoS Biology* 5 (4): 922–31. doi:10.1371/journal.pbio.0050094.
Sniegowski, Paul D, and Brian Charlesworth. 1994. "Transposable Element Numbers I N Cosmopolitan Inversions From a Natural Population of," no. 1988.
Wright, Stephen I., and Peter Andolfatto. 2008. "The Impact of Natural Selection on the Genome: Emerging Patterns in Drosophila and Arabidopsis." *Annual Review of Ecology, Evolution, and Systematics* 39 (1): 193–213. doi:10.1146/annurev.ecolsys.39.110707.173342.

CHAPTER 5

Summary

Deleterious mutations are more common than beneficial mutations (Eyre-Walker and Keightley 1999) As long as populations are finite, some deleterious mutations will continue to accumulate by virtue of being invisible to selection if they decrease fitness by a fraction smaller than the inverse of the effective population size (Wright 1931; Ohta and Kimura 1971). Moreover, in the absence of recombination, deleterious mutations can accumulate by a stochastic process (Muller 1964) that has been called “Muller’s Ratchet” (Felsenstein 1974). The rate at which Muller’s ratchet advances in populations is primarily dependent on the effective population size and the deleterious mutation rate; a small population size and a high mutation rate can both cause mutations to accumulate more quickly. The influence of deleterious mutations—especially in the presence of linkage to beneficial mutations—is an area of considerable ongoing interest in evolutionary genetics (Charlesworth, Morgan, and Charlesworth 1993; Barton 2010; Fontanari, Colato, and Howard 2003; M Lynch et al. 1993).

In the first part of my dissertation, I carried out mutation accumulation experiments with an *E. coli* strain that had an extraordinarily high genomic mutation rate. In these experiments, replicate populations were propagated at a very low effective population size. Consistent with the theory described above, measures of fitness declined in these populations and multiple populations went extinct during their propagation. I used fluctuation assays to measure mutation rates in both the extinct populations (in archives frozen shortly before extinction) and the surviving populations. Strikingly, the populations that went extinct exhibited mutation rates identical to or higher than the ancestral mutation rate shortly before their extinction, whereas many of the surviving populations had evolved lower mutation rates. Although the evolution of decreased fitness and even extinction was anticipated in these experiments, evolution of genomic mutation rates was unanticipated. Indeed, mutation accumulation experiments to date have in general assumed that the mutation rate is constant throughout propagation; my experiments clearly indicate that this need not be the case.

Because they are likely to involve loss rather than gain of function, mutator mutations are expected to be more common than antimutator mutations. In experimental and natural populations, moreover, many more instances of the evolution of increased mutation rates (Chao and Cox 1983; Cox, Degnen, and Scheppe 1972; Mao et al. 1997; Sniegowski, Gerrish, and Lenski 1997; Shaver et al. 2002) have been observed than evolution of decreased mutation rates (but

see Wielgoss et al. 2012; McDonald et al. 2012). Thus it was surprising to observe the evolution of reduced genomic mutation rates in the short timescale that spanned my mutation accumulation experiment. A genomic analysis of the populations that survived (Chapter 3) has provided some candidate mutations that may be responsible for the lower mutation rates, but the multitude of mutations observed in these populations makes interpretation of these data difficult. Further experimental work will be necessary to test the effects of these candidate mutations on the mutation rate directly.

In Chapter 2, I suggested that the evolution of decreased mutation rates in the surviving mutation accumulation populations was a consequence of selection based on avoidance of mutational load, which is likely to be a strong factor in these hypermutable populations. Computer simulations that I carried out supported this interpretation. Because beneficial mutations are rare to begin with, and because few beneficial mutations of sufficient magnitude to overwhelm drift were expected to arise in the context of the mutation accumulation protocol, it is questionable whether genetic hitchhiking (see Chapter 2) explains the evolution of increased mutation rates in some other populations—most notably, most of those that went extinct. Instead, it is possible that these increases in mutation rate were a consequence of mutation pressure and genetic drift (Lynch et al. 1993; Lynch 2010; Gerrish et al. 2007). If so, this would to my knowledge be the first experimental observation of the evolution of mutation rates as a

consequence solely of those factors, and this finding suggests an avenue for further work in this area.

In the second part of my dissertation, I carried out a computer simulation study of the effects of deleterious mutations on the spread of a lineage in an exponentially growing asexual population. As noted in Chapter 4, this effect has been tentatively called "lineage contamination" (Gerrish et al. 2016, in prep). Lineage contamination is clearly distinct from background selection, although both effects are likely to have been occurring simultaneously in some previous simulation work (Johnson and Barton 2002; Peck 1994). My investigation of lineage contamination was stimulated by theoretical (Pénisson et al. 2013; Bull and Wilke 2008; Bull, Sanjuán, and Wilke 2007) and experimental (Gentile et al. 2011) work on the fate of fitness in populations with very high mutation rates--hypermutable populations. I hypothesized that the distribution of a neutral (and, by extension, beneficial) mutant in a hypermutable asexual population is distorted by deleterious mutations that accumulate differentially in the small subpopulation represented by the neutral mutation. My simulation results indicated that, for experimentally reasonable values of the deleterious mutation rate and average effect of deleterious mutations, lineage contamination can substantially depress both the mean and variance of the number of neutral mutants after growth. This observation can be extended as the basis to investigate the fate of small beneficial lineages that arise spontaneously in an asexual population, and I am

presently carrying out simulation work associated with a larger study (Gerrish et al. 2016, in prep) to analyze this effect. The ultimate goal of that larger work is to derive the conditions (deleterious and beneficial mutation rates and distributions of their effects) under which beneficial mutations will fail to spread in asexual population due to the influence of linked deleterious mutations, thus halting or even reversing adaptive evolution. Existing models of mutation-driven extinction in asexual populations (Lynch et al. 1993; Bull and Wilke 2008; Bull, Sanjuán, and Wilke 2007) largely ignore beneficial mutations; thus, Chapter 4 and my ongoing collaborative work address a large gap in the field.

The influence of deleterious mutations on populations has been a persistent theme in evolutionary genetics since its earliest days (Sturtevant 1937). By and large, however, the classical early population genetic models of Wright, Fisher and Haldane considered the fates of alleles at individual loci, rather than the effects of linkage. The second half of the 20th century saw a strong shift toward interest in the effects of linkage on the dynamics of beneficial and deleterious mutations in populations (Hill and Robertson 1966; Charlesworth, Morgan, and Charlesworth 1993; Gerrish and Lenski 1998; Campos and Wahl 2010; Charlesworth 2012). In part this was probably driven by analyses of linked molecular regions from early sequencing studies; in part it is likely to have been a consequence of the rise of microbial experimental evolution studies, almost all of which have been carried out on asexual populations. In this dissertation, I have

explored two novel aspects of the effects of deleterious mutations in hypermutable asexual populations: the surprising evolution of mutation rates in very small populations, and the influence of deleterious mutations on the spread of a sublineage within a growing population. Further work in both areas is likely to contribute materially to our growing understanding of the influence of deleterious mutations on the evolutionary process.

References

Barton, N H. 2010. “Genetic Linkage and Natural Selection.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 365 (1552): 2559–69. doi:10.1098/rstb.2010.0106.

Bull, J J, R Sanjuán, and C O Wilke. 2007. “Theory of Lethal Mutagenesis for Viruses.” *Journal of Virology* 81 (6): 2930–39. doi:10.1128/JVI.01624-06.

Bull, James J., and Claus O. Wilke. 2008. “Lethal Mutagenesis of Bacteria.” *Genetics* 180 (2): 1061–70. doi:10.1534/genetics.108.091413.

Campos, Paulo R a, and Lindi M. Wahl. 2010. “The Adaptation Rate of Asexuals: Deleterious Mutations, Clonal Interference and Population Bottlenecks.” *Evolution* 64 (7): 1973–83. doi:10.1111/j.1558-5646.2010.00981.x.

Chao, L, and E C Cox. 1983. “Competition Between High and Low Mutating Strains of Escherichia Coli Author (S): Lin Chao and Edward C . Cox.” *Evolution* 37 (1): 125–34

Charlesworth, B, M T Morgan, and D Charlesworth. 1993. “The Effects of Deleterious Mutations on Neutral Molecular Varaition.” *Genetics* 134 (4): 1289–1303.

Charlesworth, Brian. 2012. “The Role of Background Selection in Shaping Patterns of Molecular Evolution and Variation: Evidence from Variability on the Drosophila X Chromosome.” *Genetics* 191 (1): 233–46. doi:10.1534/genetics.111.138073.

Cox, E C, G E Degnen, and M L Scheppe. 1972. “Mutator Gene Studies in Escherichia-Coli - Muts Gene.” *Genetics* 72 (4): 551–67.

Eyre-Walker, Adam, and Peter D Keightley. 1999. “High Genomic Deleterious Mutation Rates in Hominids.” *Nature* 397 (6717): 344–47.

Felsenstein, J. 1974. “The Evolutionary Advantage of Recombination. Genetics.” *Genetics*.

Fontanari, J F, a Colato, and R S Howard. 2003. “Mutation Accumulation in

Growing Asexual Lineages.” *Physical Review Letters* 91 (21): 218101. doi:10.1103/PhysRevLett.91.218101.
Gentile, Christopher F, Szi-Chieh Yu, Sebastian Akle Serrano, Philip J Gerrish, and Paul D Sniegowski. 2011. “Competition between High- and Higher-Mutating Strains of Escherichia Coli.” *Biology Letters* 7 (3): 422–24. doi:10.1098/rsbl.2010.1036.
Gerrish, Philip J, Alexandre Colato, Alan S Perelson, and Paul D Sniegowski. 2007. “Complete Genetic Linkage Can Subvert Natural Selection.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (15): 6266–71. doi:10.1073/pnas.0607280104.
Gerrish, Philip J, and Richard E Lenski. 1998. “The Fate of Competing Beneficial Mutations in an Asexual Population S N Ln N :,” 127–44.
Gerrish, Philip, Sophie Pénisson_, Tanya Singh, and Paul Sniegowski. 2016. “Dynamics and Fate of Beneficial Mutations under Lineage Contamination by Linked Deleterious Mutations. In Preparation.”
Hill, W. G., and A Robertson. 1966. “The Effect of Linkage on Limits to Artificial Selection.” *Genetical Research* 8: 269–94.
Johnson, Toby, and Nick H. Barton. 2002. “The Effect of Deleterious Alleles on Adaptation in Asexual Populations.” *Genetics* 162 (1): 395–411.
Lynch, M, R Bürger, D Butcher, and W Gabriel. 1993. “The Mutational Meltdown in Asexual Populations.” *The Journal of Heredity* 84 (5): 339–44.
Lynch, Michael. 2010. “Evolution of the Mutation Rate.” *Trends in Genetics* 26 (8). Elsevier Ltd: 345–52. doi:10.1016/j.tig.2010.05.003.
———. 2016. “Mutation and Human Exceptionalism: Our Future Genetic Load.” *Genetics* 202 (3): 869–75. doi:10.1534/genetics.115.180471.
Mao, E F, L Lane, J Lee, J H Miller, Emily F Mao, Laura Lane, Jean Lee, and Jeffrey H Miller. 1997. “Proliferation of Mutators in A Cell Population . Proliferation of Mutators in A Cell Population” 179 (2): 417–22.
McDonald, Michael J., Yu Ying Hsieh, Yen Hsin Yu, Shang Lin Chang, and Jun Yi Leu. 2012. “The Evolution of Low Mutation Rates in Experimental Mutator Populations of Saccharomyces Cerevisiae.” *Current Biology* 22 (13): 1235–40. doi:10.1016/j.cub.2012.04.056.
Muller, H.J. 1964. “Muller, H.J., 1964. The Relation of Recombination to Mutational Advance. Mutat.” *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* 1: 2–9.
Ohta, Tomoko, and M Kimura. 1971. “On the Constancy of the Evolutionary Rate of Cistrons *” 25 (813).
Peck, Joel R. 1994. “A Ruby I N the Rubbish: Beneficial Mutations, Deleterious Mutations.”
Pénisson, Sophie, Paul D Sniegowski, Alexandre Colato, and Philip J Gerrish. 2013. “Lineage Dynamics and Mutation–selection Balance in Non-Adapting Asexual Populations.” *Journal of Statistical Mechanics: Theory and*

Experiment 2013 (01): P01013. doi:10.1088/1742-5468/2013/01/P01013.
Robertson, A. 1960. “A Theory of Limits in Artificial Selection.” *Proceedings of the Royal Society of London B: Biological Sciences* 153 (951): 234–49.
Shaver, Aaron C., Peter G. Dombrowski, Joseph Y. Sweeney, Tania Treis, Renata M. Zappala, and Paul D. Sniegowski. 2002. “Fitness Evolution and the Rise of Mutator Alleles in Experimental Escherichia Coli Populations.” *Genetics* 162 (2): 557–66.
Sniegowski, Paul D, Philip J Gerrish, and Richard E Lenski. 1997. “Evolution of High Mutation Rates in Experimental Populations of E . Coli.” *Nature* 606: 703–5.
Sturtevant, a. H. 1937. “Essays on Evolution. I. On the Effects of Selection on Mutation Rate.” *The Quarterly Review of Biology* 12 (4): 464. doi:10.1086/394543.
Wielgoss, Sébastien, Jeffrey E Barrick, Olivier Tenaillon, Michael J Wiser, W James Dittmar, Stéphane Cruveiller, Béatrice Chane-Woon-Ming, Claudine Médigue, Richard E. Lenski, and Dominique Schneidera. 2012. “Mutation Rate Dynamics in a Bacterial Population Re Fl Ect Tension between Adaptation and Genetic Load.” *Proc Natl Acad Sci (USA)* 110 (1): 222–27. doi:10.5061/dryad.hb3b5.

APPENDIX

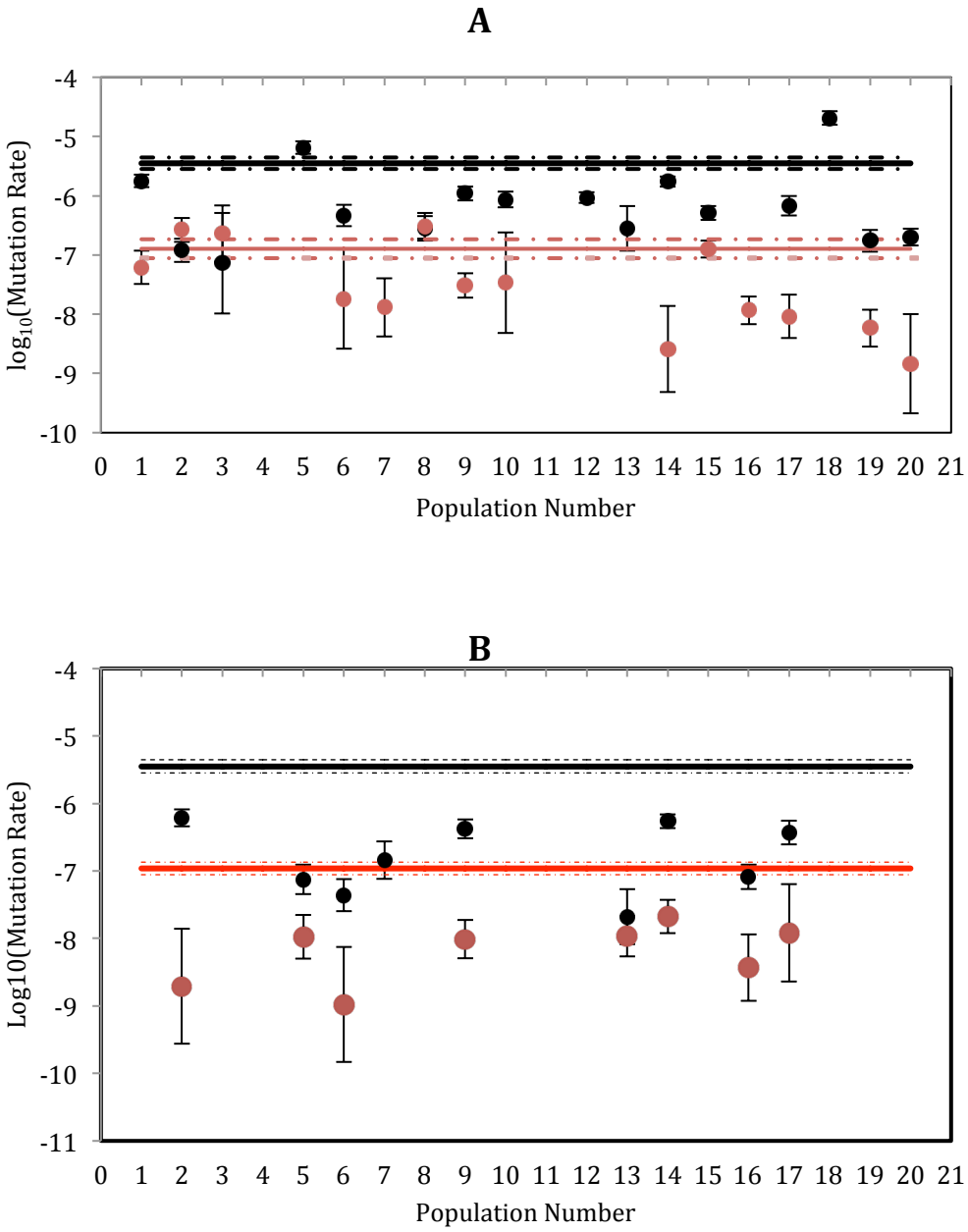


Figure A1. Nalidixic acid mutation rates (BLACK) and Streptomycin mutation rates (RED) along with associated 95% confidence intervals. The solid lines, black and red, represent the nalidix acid resistance mutation rate and the streptomycin resistance mutation rate of the ancestor, PS2534.

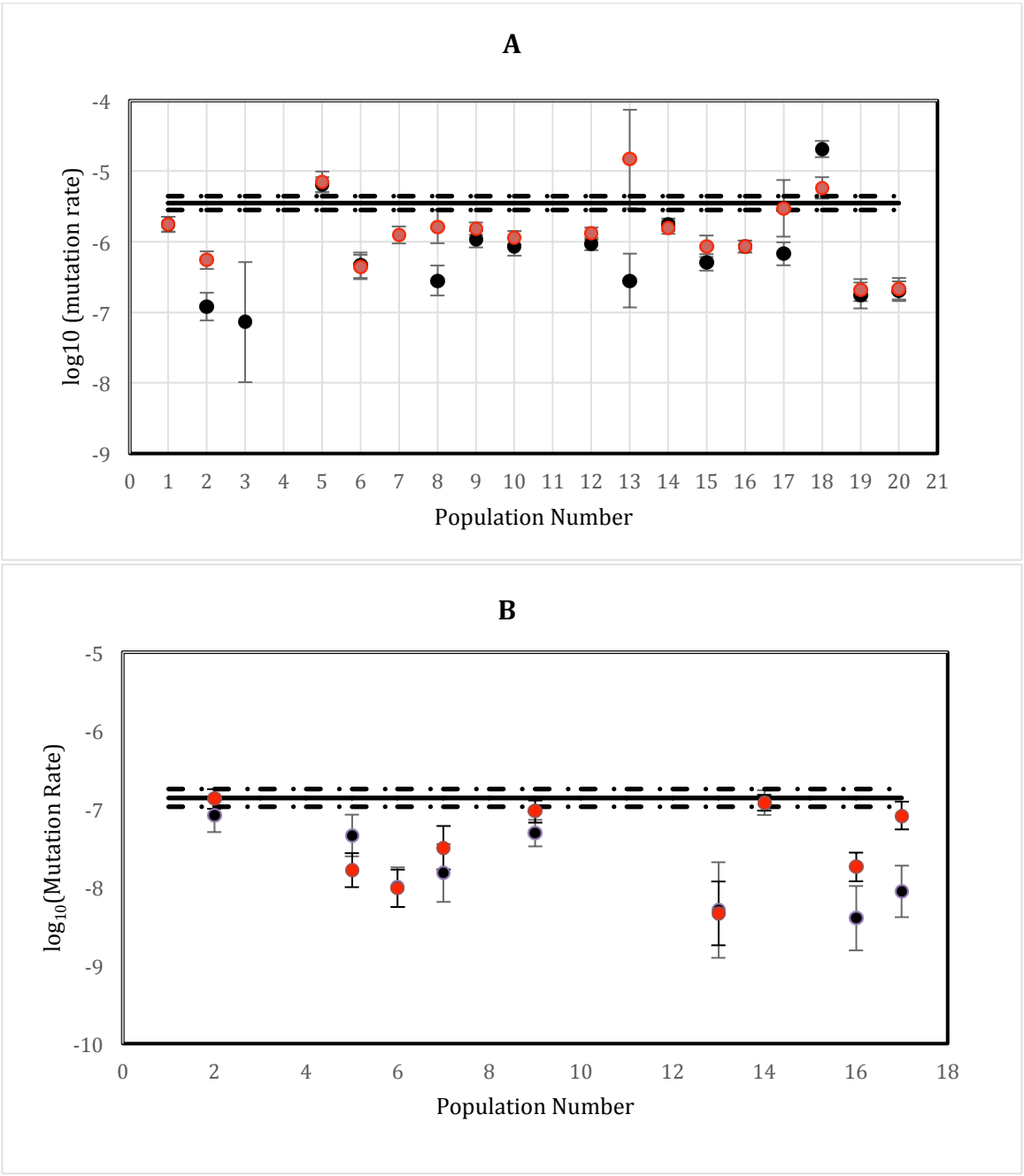


Figure A2. Nalidixic acid mutation rates from replicate 1(BLACK) and replicate 2 (RED) along with associated 95% confidence intervals. The solid lines, black and red, represent the nalidix acid resistance mutation rate and the streptomycin resistance mutation rate of the ancestor, PS2534.

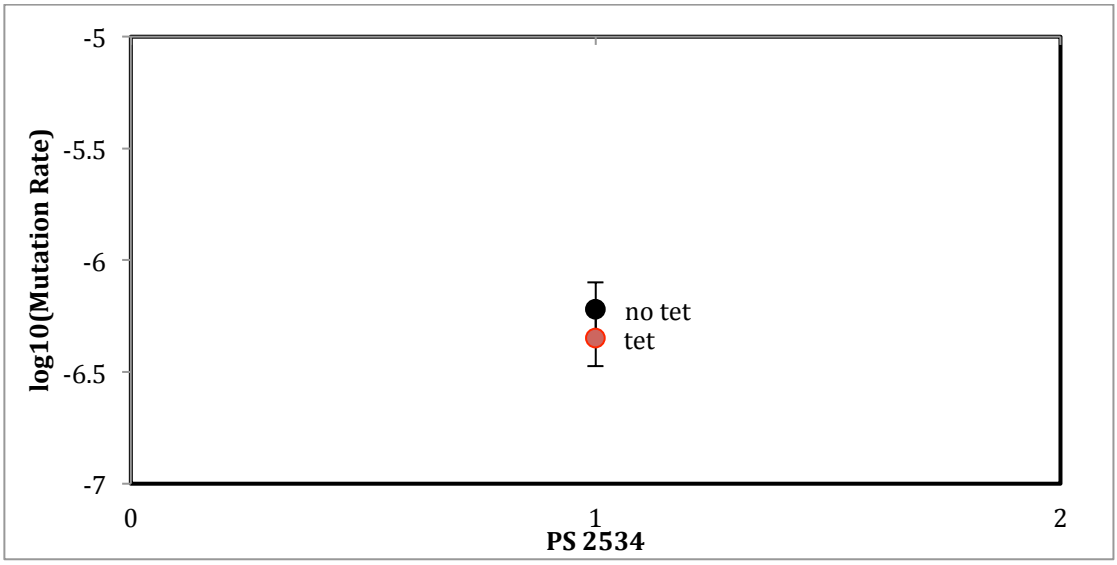


Figure A3. Nalidixic acid mutation rates of the common ancestor PS2534 in tetracycline supplemented medium (BLACK) and medium devoid of tetracycline (RED) along with associated 95% confidence intervals. DM1000 or Davis Minimal Media supplemented with 1000 mg of Glucose was used for this Fluctuation Assay.

GENE NAME	NUMBER OF NONSYNONYMOUS SNPS OBSERVED
crfC	7
dcm	3
dinB	14
dnaA	2
dnaB	2
dnaC	1
dnaE	10
dnaG	5
dnaJ	2
dnaK	1
dnaN	1
dnaQ	2
dnaX	3
gyrA	5
gyrB	8
hda	1
holA	3
holB	3

holD	1
ligA	4
ligB	3
mukB	6
mukE	6
mukF	1
mutH	1
mutL	4
mutM	1
mutS	4
mutY	2
nrdA	2
nrdB	3
nrdE	20
nrdF	3
polA	4
polB	3
priA	8
priC	1
prlC	1
rarA	4
recA	1
recQ	1
rep	5
ruvA	1
ruvB	4
ruvC	1
sbcC	9
sbcD	3
ssb	1
topB	7
tus	3
uvrD	1
ycdX	3
yciV	4
ydaV	1

. Table A.1 List of all mutation rate associated genes where SNPs were discovered for the surviving LB lines and the number of non-synonymous SNPs discovered in those genes.

GENE NAME	NUMBER OF NONSYNONYMOUS SNPS OBSERVED
crfC	1
dam	1
dcm	1
dinB	2
dnaC	1
dnaE	2
dnaG	1
dnaK	1
dnaN	1
dnaX	2
gspB	4
gyrA	3
holA	1
ligA	1
mukB	3
mutH	1
mutL	3
mutS	1
nrdE	5
polA	3
recQ	3
sbcC	3
sbcD	2
uvrD	1
ycdX	1
yciV	2
ydaV	1

Table A.2 List of all mutation rate associated genes where SNPs were discovered for the surviving MG lines and the number of non-synonymous SNPs discovered in those genes.