EVOLUTIONARY INSTABILITY OF GENOMIC MUTATION RATE IN RAPIDLY ADAPTING ASEXUAL MUTATOR *ESCHERICHIA COLI* POPULATIONS

Mitra M. Eghbal

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

2017

Supervisor of Dissertation

_____

Paul Sniegowski, Professor, Department of Biology, University of Pennsylvania

Graduate Group Chairperson

_____

Michael Lampson, Associate Professor, Department of Biology, University of Pennsylvania

Dissertation Committee:

Timothy Linksvayer, Associate Professor, Department of Biology, University of Pennsylvania

Mechthild Pohlschröder, Professor, Department of Biology, University of Pennsylvania

Paul Schmidt, Professor, Department of Biology, University of Pennsylvania

Sarah Tishkoff, Professor, Department of Biology, University of Pennsylvania

EVOLUTIONARY INSTABILITY OF GENOMIC MUTATION RATE IN RAPIDLY ADAPTING ASEXUAL MUTATOR *ESCHERICHIA COLI* POPULATIONS

This thesis is dedicated to whoever reads it, be they compelled by curiosity or by coercion.

Reading my thesis entails a dedication of your time. It is only right that I reciprocate that dedication.

ACKNOWLEDGMENTS

In a certain light, mutations could be viewed as a form of biological luck (or lack thereof). Studying mutation rate evolution for several years really makes you appreciate the rarity of good luck. Several years is also enough time for me to realize my own good luck. In my time as a graduate student, I've been lucky to be in contact with a number of people who influenced my life for the better.

Thank you to the members of my lab. Although, at almost any given point in time, we were a small lab, there was a changing cast of many delightful characters over the years.

I'll start off with the grad students. My senior graduate student lab mates, Eugene and Chris, were great resources (for science and for mischief), especially when I first arrived in the lab. Years after his graduation, Eugene still offers me excellent advice from afar.

In the latter years of my schooling, I was especially grateful for the company of Tanya and Ben. Their perpetual curiosity facilitated the exploration of a number of scientific questions together. I also appreciated the fact that we had complementary areas of scientific expertise and were able to assist each other as needed. Tanya and Ben, may you dream evermore of many dragons and many robots (respectively), who will welcome you as one of their own and extend the respect and affection that you both deserve.

Our lab was lucky to be visited by a master's student (Damien) and rotation students (Alexandra and Ozan) whose intellectual enthusiasm and warm-heartedness were greatly appreciated, and with whom I continued to have many productive and enjoyable chats about science (and science fiction) for years even after they departed our lab.

I am honored that our lab was visited by a philosophy PhD student, Emily, who was eager to learn about our work and also eager to offer us a glimpse of her academic world. Exposure to the philosophy of science community has helped me see many more complexities behind that hazy beast that we call the 'scientific process'. Our lab was fortunate to have a representative philosopher as approachable, articulate, and enjoyable as Emily.

Although we only had one postdoc in our lab in the time I've been here, she certainly had an impact on my research. Kathleen was very helpful in getting me oriented in the wild frontiers of genomic sequencing. The current state of genomics in our lab wouldn't be the same without her.

I am grateful to Jude for his work on the hard selection project for over a year and for keeping my mind fresh with his passion for science. I can hardly believe that he's already heading to college this fall, but I know that Penn is lucky to have this bright and hard-working fellow. I am also grateful to Holly and Clement for their diligent work on the soft selection fluctuation assays. Spending several weeks of a summer doing labwork is a sacrifice for any high schooler, yet it was a sacrifice they made very good-naturedly.

In the past year, I have had the fortune of working with several other talented and industrious ‘apprentices’ (as I like to think of them): Arlene, Meredith, and Breanna. Performing hundreds of fitness assays is no quick task, and if they hadn’t helped out, well, I don’t know where I’d be. Probably out of my mind.

My lab mentees have given me great cause for optimism with regard to the future of science, and I am proud of them all.

Thank you also to Kate, Matt, Angela, Dave, Dan, Ankur, Brooks, and Emilie for being a pleasure to work alongside and for being such good representatives of the next generation of science.

I am deeply grateful to my committee, both as a group and individually: Mecky, Paul, Tim, and Sarah. I really appreciated their thought-provoking questions and their huge range of expertise. Thank you for the many hours that you spent listening to me talk about my research and offering your advice. My thesis (and, in fact, my career) has definitely benefited from your help.

Many thanks to Phil, especially for his help on the fluctuation assay component of my thesis. His excellent advice from his perspective as a theorist was truly appreciated by this experimentalist, and his unflappably mellow, friendly, and unpretentious demeanor was a welcome counterweight to the stresses of graduate research.

I want to thank the Kohli lab (especially Rahul, Matt, and Charlie) and the other members of the bacterial mutagenesis journal club. They were instrumental in helping me think more broadly about all facets of mutagenesis, which prepared me for the next step of my scientific career. I want to thank Sam from the Goulian lab for kindly sharing her impressive knowledge of molecular microbiology techniques. Thanks to Joe, Luigi, and Camilo for being great postdoc role models.

I want to thank the whole Pohlschröder lab, past and current, for microbiological solidarity and for being such outstanding lab neighbors and generously lending us their supplies and equipment on many occasions.

Thank you to the NASA Astrobiology Institute team, especially Frank, for funding and for their helpful input on my work at a variety of conferences.

I couldn't have wished for a better graduate coordinator than Colleen. In grad school, it's easy for anyone to lose track of all the years, just slipping by under piles of work. Colleen kept everyone on track and celebrated our every milestone. Sometimes I feel like she is our Ambassador to Reality (and a very caring and organized one at that).

I want to thank Reggie, the Biology IT Department, Business Office, Greenhouse, and Academic Office. Their competence and helpfulness made research infinitely more bearable. So did Thor.

I was lucky to have some great supervisors in my time as a teaching assistant. I would like to especially thank Brenda, Linda, Kim, and Greg for modeling such good teaching. While TAing for introductory biology, I was lucky to work alongside some professional TAs, better known as the "super-TAs"… and they are, indeed, super TAs: Jessica, Staver, and Lori. They taught me lots of useful tricks of the trade. Svetlana was also very helpful to me during my time as an introductory biology TA.

I was fortunate to work part-time with the Penn Center for Teaching and Learning (CTL) for a year. Through them, I learned more about pedagogy and had the opportunity to interact with other academic educators across a variety of disciplines. I want to thank all the Biology professors and staff who volunteered their time to help with the CTL workshops.

I want to thank my graduate student cohort for their companionship over the years: Hung, Aaron, Chris, Emily, Lee, Stephanie, Wei, Leo, and George. I especially want to thank Steph, Lee(roy), and Wei for scientific advice and also for communal enjoyment of such things as trashy sci-fi films and tasty victuals.

In fact, I would like to thank all the graduate students of the department for their camaraderie. Thanks to Jean (and, really, the whole Kim lab) for his kind help back when I was first wrangling with a Qubit. Thanks to Aurora, Vinayak, Xiaqing, Katherine, Erin, Jill (whom I see as an honorary grad student), and Dan for their company. Thanks to

Rohini and Alex for enjoyable discussions about biology education and for involving me in their outreach programs. Thanks to Vince and Yang for their collaboration on representing our department in SASgov, where, for a year, we had the opportunity to be involved a very different side of grad student life.

It's possible that I forgot to list a few people, so just realize that any omissions are unintentional.

I saved my acknowledgements for my adviser for near the end of this section, because, well, I have the most to say about him. So: I am thankful to my adviser Paul for his compassion, wide-ranging intellect, and contagious love of science in all its forms. True to his role of adviser, he has provided me with copious amounts of advice, both on research matters and on the general state of being a graduate student – a state that is exciting, nerve-wracking, and perplexing all at once, and he would always acknowledge that this state is indeed all of those things, rather than making me feel like I was the only one who felt that way. His enthusiastic teaching and his prudent approach to research are appreciated. I have enjoyed many hours of intellectually stimulating conversation with Paul, and am grateful for his willingness to wade through countless drafts of my writing, accompanied by critique and caffeine (and sometimes candy, cookies, and even consolation). Because of him, I will never regard a microbe, a split infinitive, or a Poisson distribution in the same way again, and am happy and proud to call him teacher, mentor, and friend. And, perhaps, akin to one of those kindly wizards of legend who offers guidance to puzzled wayfarers. Or… something like that. Thank you.

I am grateful to my friends outside the department for helping me with maintaining so-called “work-life balance”. Many, many thanks to my partner Ethan for his endless support. Your kindness is so tremendous that it sometimes feels unreal. I am grateful to walk this planet with you. I have a lot more to tell you that I’ll tell you in person. Wait, one thing, though -- it makes me sad to realize that you are literally one of the only people among the many named in this section who will physically accompany me into the next stage of my life. That said, if I were forced to pick *just one person*, it would’ve been you anyway. (N.B.: This does not diminish the fact that I will still miss everyone else, too.)

ABSTRACT

EVOLUTIONARY INSTABILITY OF GENOMIC MUTATION RATE IN RAPIDLY ADAPTING ASEXUAL MUTATOR *ESCHERICHIA COLI* POPULATIONS

Mitra Eghbal

Paul Sniegowski

Alleles conferring higher mutation rates (mutators) can fix in asexual populations through a process called 'mutator hitchhiking'. Theory predicts that repeated mutator hitchhiking can occur in an adapting asexual mutator population. I tested this prediction in two settings: a *mutL*$^-$ population under lethal selection and a *mutS*$^-$ population under soft selection. In both experiments, the starting mutation rate was 100-fold higher than wild-type. In the lethal selection experiment, two replicate populations were exposed to a sequence of three different antibiotics. In both replicates, all survivors sampled after the final antibiotic exposure had undergone further genomic mutation rate increases. Whole-genome and Sanger sequencing revealed that an identical spontaneous 1-bp insertion in *mutT* (a known mutator gene) rose to probable fixation in both populations. Complementation tests demonstrated that the *mutT*$^-$ defect was responsible for the increased mutation rates. In the soft selection experiment, 30 isogenic populations were propagated in limited glucose media. After 900 generations, five clones were isolated from each population. Relative to the ancestor, 9% of the clones had increased mutation rates and 68% had unchanged mutation rates; surprisingly, 23% had decreased mutation rates. Most populations (21 of 30) had at least one clone whose mutation rate either

increased or decreased. One population exhibited apparent fixation for a mutator and one other population exhibited apparent fixation for an antimutator. Some of the sequenced clones with altered mutation rates had mutations in known (anti)mutator loci. I conclude that the mutators likely arose by hitchhiking and that the antimutators likely confer pleiotropic direct fitness benefits. Competitions between the evolved clones and the ancestor demonstrated that all clones and populations had increased in fitness since generation 0. No relationship was detected, however, between mutation rate and relative fitness. These experiments provide evidence supporting the prediction of repeated mutator hitchhiking. More broadly, the work described in this dissertation reveals multiple ways in which the mutation rates of asexual populations may be evolutionarily unstable, with potential implications for evolving asexual systems, including infectious agents and cancer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I: Introduction

Overview

All genetic variation ultimately originates from mutations. Mutations are spontaneous in origin, and their effects on their bearers can vary dramatically. Certain types of mutations can modify the rate of mutation acquisition. These mutation rate modifiers are known as 'mutators' and 'antimutators': mutators increase mutation rates, and antimutators decrease mutation rates. Consequently, mutation rates can vary across both species and populations (Drake 1991; Lynch 2010). Not only can mutation rates affect the course of evolution, mutation rates themselves are subject to evolution. The random nature of mutations raises the question of how mutation rates evolve, in light of the fact that many mutations have little to no phenotypic effect. Of the mutations that do have any phenotypic effect, most are deleterious and few are beneficial. What is the evolutionary counterweight against the negative effects of the accumulation of deleterious mutations?

Why the mutation rate is non-zero

Nearly a century ago, Alfred Sturtevant asked why the mutation rate does not evolve to zero (Sturtevant 1937). In the subsequent eighty years, many researchers have sought answers to this question; at present time, at least three possible hypotheses have been extensively explored, with varying degrees of favor among the scientific community. Firstly, mutations are the original source of all evolutionary novelty. Beneficial mutations are crucial for adaptive progress in a changing environment (Kimura 1967; Ishii et al. 1998; Lynch 2011). One might hypothesize, therefore, that selection would favor non-zero mutation rates. Secondly, there has been debate over whether the energy cost of

replication fidelity forms a lower bound to mutation rates (Kondrashov 1995; André and Godelle 2006; Lynch 2008). Quick replication is especially important for prokaryotes, and high replication fidelity would ostensibly reduce the speed of replication, so one might then expect that prokaryotes have lower replication fidelity (Sniegowski and Raynes 2013). In reality, prokaryotes typically have lower mutation rates than eukaryotes (base-substitutional rates of *E. coli* and archaea vs. mammal, chicken, *Drosophila*, and yeast: see Lynch 2008; *Drosophila* vs. microorganisms: John Maynard Smith 1978); thus, it is unlikely that the cost of replication fidelity is actually the ultimate inhibitor of zero mutation rates (Lynch 2010). Thirdly, it has been suggested that genetic drift imposes a lower bound to mutation rate evolution, because genetic drift may limit the ability of natural selection to reduce mutation rate (Lynch 2010; Sung et al. 2012; reviewed in Sniegowski and Raynes 2013).

Mutator hitchhiking: sexual vs. asexual

Mutators have been detected in clinical and natural populations across a variety of microbial species (Gross and Siegel 1981; LeClerc et al. 1996; Oliver et al. 2000; Denamur et al. 2002; Baquero et al. 2004). *Escherichia coli*, with its small genome, short generation time, and ability to be cultivated in large numbers, is a common model organism in experimental evolution. It has been the subject of many evolutionary investigations, including fitness landscapes (Gordo and Campos 2013), epistasis (Wang et al. 2013), novel modes of metabolism (Blount et al. 2008), antibiotic resistance (Lenski and Hattingh 1986), and mutation rate (Sniegowski et al. 1997). In February 1988, Richard Lenski began propagating 12 isogenic populations of asexual *E. coli* in minimal

glucose media; every few hundred generations, samples of each evolving population were preserved in glycerol and archived in a frozen fossil bed (Lenski and Travisano 1994). Mutator alleles spontaneously emerged and fixed in three of the evolving populations at varying time points before 10,000 generations; the mutation rates were elevated by one to two orders of magnitude above the ancestor, which was attributed to the mutator alleles being mutations in genes of the methyl-directed mismatch repair pathway (Sniegowski et al. 1997). Tens of thousands of generations after this initial discovery, genomic and phenotypic studies confirmed the presence of mutators in three additional populations (Barrick et al. 2009). In adapting asexual populations, mutator alleles can increase in frequency because there is no recombination to break the linkage disequilibrium between the mutator allele and the beneficial mutation(s) whose existence is facilitated by the mutator allele; this process is called 'mutator hitchhiking' (Sniegowski et al. 1997). Thus, the immediate fitness advantage indirectly elevates the mutation rate, and the fitness cost of the deleterious mutations may only be exacted after the mutator allele has fixed in the population (Denamur and Matic 2006; Gerrish et al. 2007).

In most theoretical work that explores the evolutionary trajectory of mutation rates, the mutation rate modifiers themselves are treated as having no direct effects on fitness (Holsinger and Feldman 1983; Tenaillon et al. 1999; Gerrish et al. 2007; Raynes et al. 2013; Raynes and Sniegowski 2014). Yet although there is extensive literature on mutators hitchhiking through *indirect* selection (Chao and Cox 1983; Sniegowski et al. 2000; Shaver et al. 2002; Palmer and Lipsitch 2006), there remains the formal possibility that mutator alleles may confer *direct* fitness benefits in some scenarios. The theoretical

treatment of mutator alleles as having no direct fitness effects receives some support from empirical studies that found no direct benefit to several mutator alleles in methyl-directed mismatch repair genes, including *mutH, mutL, mutS,* and *uvrD* (Mao et al. 1997; Shaver et al. 2002). However, there is evidence for direct selection for a *mutS*⁻ mutator allele in an evolving population of *Pseudomona aeruginosa* (Torres-Barceló et al. 2013); it was proposed that the decreased fidelity of DNA repair allowed for greater resistance against the harmful effects of hydrogen peroxide. The question of whether mutation rate modifiers are truly neutral with regard to direct fitness effects is an empirical question that has yet to be addressed in many contexts. In real genomic contexts, the existence of pleiotropy may increase the likelihood that a mutation rate modifier possesses characteristics besides affecting the mutation rate; if these other properties happen to confer fitness benefits, then a mutator allele could hypothetically rise to fixation via these direct fitness effects, rather than via mutator hitchhiking (Raynes and Sniegowski 2014).

Experimental work has demonstrated that, under long-term selection, a single-mutator strain (bearing one mutator allele) can outcompete a wild-type strain, *if* the mutators are seeded into the population above a threshold frequency at the start of propagation (Chao and Cox 1983). It is also known that double-mutators (bearing two mutator alleles) can outcompete single-mutators under short-term selection – if the double-mutators are seeded into the population above a threshold frequency at the start of propagation (Gentile et al. 2011). Though these findings are intriguing, the artificial seeding leaves unanswered the question of whether additional mutator alleles can *spontaneously* arise and fix in a preexisting mutator population. Theoretical predictions suggest that

additional mutator alleles may indeed emerge and rise to fixation in adapting asexual mutator populations (Gerrish et al. 2007), but this prediction of recurrent mutator hitchhiking has not been fully tested experimentally. Because recurrent mutator hitchhiking is predicted to occur more readily in a changing environment, such an experiment would ideally begin in a novel environment and the experimental populations would adapt for hundreds of generations (Gerrish et al. 2007).

Antimutators: the reduction of mutation rates

The work in this dissertation is exclusively concerned with asexuals. In a sexual population, mutators may be expected to decline in frequency, because frequent recombination has the potential to erode linkage disequilibrium between mutators and the beneficial mutations that they facilitate, and the preponderance of deleterious mutations will exact their fitness cost on the population (Kimura 1967; Raynes et al. 2011). How, then, do mutation rates decline in asexual populations? A case has been made against the existence of strong antimutators, both from an enzymatic and an evolutionary perspective (Drake 1993). However, these arguments lie in the context of wild-type backgrounds; we cannot assume that this reasoning automatically applies to all potential antimutators that may land on mutator backgrounds. Furthermore, an apparent antimutator mutation arising in a mutator background may not be an antimutator in a wild-type context; in *E. coli,* an example of this context-dependence is *mutY*$^{-}$, which is a mutator allele in a wild-type background, but can be an antimutator allele if it lands on a *mutT*$^{-}$ mutator background (Fowler et al. 2003). It has been suggested that certain antimutator alleles may increase in frequency by conferring direct fitness benefits; for example, decreased mutation rates

were observed in an experimentally evolved population of *mutS*$^{-}$ *E. coli* that was originally isolated from a clinical patient (Turrientes et al. 2013). It was proposed that the bacteria had spontaneously acquired a new allele that rose to fixation by increasing adaptation to an aerobic environment, but had the additional effect of lowering the genomic mutation rate (Turrientes et al. 2013).

Deleterious mutations contribute to genetic load, which can be described as the difference between the mean fitness of a population and the fitness of the population if all its members had an ideal genotype (Whitlock and Davis 2011). Thus, an accumulated deleterious mutational load has the potential to decrease fitness – although deleterious mutational load does not always impair fitness in the short-term (Bull and Wilke 2008). The reduction of deleterious mutational load can be facilitated by lowered mutation rates in mutator populations; the phenomenon can accompany situations such as repeated population bottlenecks and/or increased adaptation to the environment over time (Tenaillon et al. 2000). The outcome of repeated population bottlenecking has been explored through mutation accumulation (MA) experiments, where a genetically homogenous starting population is propagated at tiny population sizes (Lynch et al. 2016). When 40 double-mutator *E. coli* populations were propagated at very small effective population sizes, several populations eventually went extinct; most of these extinct populations had evolved higher mutation rates at some point prior to extinction (Singh et al. 2017). However, the surviving populations generally had lowered mutation rates (Singh et al. 2017). In these experiments, the extreme bottlenecking intensified the accumulation of mutations; because most mutations with any phenotypic effect are

deleterious, there may have been strong selective pressure to lower the mutational load (Singh et al. 2017). Furthermore, the effective population sizes were so small that the mutators had only minimal advantage at rapidly acquiring beneficial mutations (Singh et al. 2017).

However, MA experiments are not the only setting in which antimutators can emerge – the spread of antimutators was also observed in Lenski's long-term evolution experiment, where the population sizes were sufficiently larger than in an MA experiment (Papadopoulos et al. 1999). One population acquired a *mutT*$^-$ mutator allele at some point between 20,000 and 30,000 generations of evolution (Barrick et al. 2009). After ~40,000 generations, this *mutT*$^-$ population became fixed for a *mutY*$^-$ allele, which lowered the genomic mutation rate – not to wild-type levels, but by 40-60% (Wielgoss et al. 2013). Because the population had been adapting for many generations, the average effect of a beneficial mutation had decreased over time. Thus, the fitness advantage of a low genomic mutation rate increased, because the effect of the deleterious mutational load had exceeded the potential for continued adaptation (Wielgoss et al. 2013). Selection for reduced deleterious mutation load (and an accompanying reduction in mutation rate) has also been inferred in populations of mutator (*msh2*) yeast, *Saccharomyces cerevisiae* (McDonald et al. 2012).

The broad relevance of mutation rates

In addition to its significance within evolutionary theory and experimental evolution, mutation rate evolution is relevant to fields as diverse as biomedicine and astrobiology.

Mutators have been found in antibiotic-resistant and pathogenic populations of bacteria (LeClerc et al. 1996; Oliver et al. 2000; Schaaf et al. 2002), RNA viruses (Furio et al. 2005; Vignuzzi et al. 2005; Elena and Sanjuán 2005), and coevolving populations of bacteria and viruses (Pal et al. 2007). Human tumor cells have long been thought to be genetically unstable (Nowell 1976), and there is indeed some evidence for elevated mutation rates in certain cancers (Modrich 1995; Sprouffske et al. 2012). Some theoretical work has gone so far as to speculate that the emergence of a mutator is required for the start of multistage carcinogenesis (Loeb 1991) (but see Shibata and Lieber 2010). The topic of mutator evolution is especially important for understanding the very first years of evolution on Earth, where early life-forms may have lacked contemporary mechanisms for genetic proofreading and recombination (Bernstein, Byers, and Michod 1980), even as they were being bombarded by mutagenic UV light penetrating the still-fragile atmosphere, which may have put them at risk for 'mutational meltdown' over extended periods of time (Hessen 2008). It has even been proposed that the damage repair mechanisms in cells have evolved to cope with levels of radiation almost one order of magnitude greater than the contemporary level of radiation (Karam and Leslie 1996). Additionally, the significantly warmer temperatures on the young Earth may have strongly increased the frequency of mutagenic events; for example, the cytosine deamination reaction is known to be sensitive to temperature and is responsible for many spontaneous mutations (Lewis et al. 2016).

Intriguingly, the study of mutation rates may also be useful in understanding the *demise* of certain populations: much thought has been given to the possibility that evolving

mutation rates may become intolerably high and lead to extinction (Gerrish et al. 2007; Tejero et al. 2016). The potential for high mutation rates to cause 'error catastrophe' (where the number of mutations exceeds the point of viability) has been particularly of interest in the study of riboviruses and retroviruses, which possess the highest known mutation rates and are therefore thought to exist at the brink of error catastrophe (Graci and Cameron 2002; Anderson et al. 2004); indeed, the antiviral ribavirin has been demonstrated to induce error catastrophe in poliovirus via mutagenesis that is moderate in degree (~4-fold increase in mutation rate), yet lethal in its outcome (Crotty et al. 2001).

The factors that complicate experimental investigations of mutation rates

For many reasons, mutation rate is a challenging topic of study. Firstly, mutation rate is a trait that affects other traits. In hastening the rate of change of the genome, a mutator may change its own significance, due to the emerging new mutations, and possibly the potential epistatic interactions and pleiotropic effects accompanying those new mutations. Secondly, the genomic mutation rate ultimately affects the whole genome – and yet, because mutations are rare, only a small number of sites will be mutated, even when the mutation rate is high (Zhu et al. 2014). Among that small number of mutations, most will have no phenotypic effect.

This raises a third point: the rarity of mutations can be a source of frustration to researchers because the mutation rate makes itself known only over multiple generations. Hence, measuring mutation rates is difficult, whether you measure them by phenotypic assays (e.g., fluctuation assays) or by whole genome sequencing. Mutation rates are

already time-consuming to measure in *E. coli*, but they are even harder to measure in organisms with larger genomes and/or longer lifespans. As an extreme case of a challenge in this field, it is worth noting the multiple attempts that have been made at measuring human mutation rates. Human pseudogene sequences have been compared to the respective chimpanzee homologous sequences, yielding an estimate of ~2.5 x 10^{-8} mutations per nucleotide site per generation (or 175 mutations per diploid genome per generation) (Nachman and Crowell 2000), though this measurement cannot fully account for the possibility of convergent evolution between humans and chimpanzees. Y chromosome sequences were compared between two men who shared a distant common ancestor, yielding an estimation of 3.0 x 10^{-8} mutations per nucleotide per generation – yet there were only a scant few mutations between them, and the Y chromosome (small, largely non-recombinant, and confined to a single sex) is hardly representative of the whole human genome (Xue et al. 2009). The genomes of two family triads were sequenced, each triad consisting of two parents and one child, for an average of 1.0 x 10^{-8} mutations per nucleotide per generation – but two generations is a mere snapshot of human evolution, and there was variation in mutation rate between the two families (Conrad et al. 2011). The fact that the three cited studies produced somewhat different estimations is telling of the challenges behind measuring mutation rates.

But even in a simpler organism such as *E. coli,* there are challenges in measuring mutation rates. Fluctuation tests are a common method for measuring mutation rates: the individual of interest is grown to a large population size in permissive media, and then exposed to a selective agent. The numbers of surviving individuals, who have

spontaneously acquired a mutation conferring resistance at some point since the start of the test, are informative of the mutation rate (Luria and Delbrück 1943). However, when performing a fluctuation test, one must assume that resistance to the selective agent does not adversely affect fitness, even in a permissive environment (Pope et al. 2008). Yet there is some evidence suggesting that certain types of resistance are associated with a fitness deficit (for effects of rifampicin resistance in *Mycobacterium tuberculosis*, see Billington, McHugh, and Gillespie 1999). Additionally, a fluctuation test measures the mutation rate in only a small fraction of the genome – namely, whatever sites that are capable of conferring resistance to the selective agent, if mutated.

Fourthly, although mutations are random, there are certain scenarios in which a higher mutation rate may offer a greater advantage than in other scenarios. When a population is under selection in a new environment, high mutation rates allow organisms to acquire new mutations that may increase fitness; after the population becomes better adapted, the number of available beneficial mutations decreases and the usefulness of a high mutation rate declines (Wielgoss et al. 2013; Tenaillon et al. 2016). Furthermore, under lethal selection, the need for beneficial mutations is more urgent than it would be under soft selection. The broad range of scenarios inviting the study of mutation rates could be viewed not so much as a difficulty, but rather as an intellectual challenge inherent to the field. The existence of mutations in every known living organism means that mutation rates could be hypothetically studied in any organism – yet, precisely because of the universal presence of mutations, these studies would take place in highly dissimilar experimental contexts.

The current state of progress in the study of mutation rates

Despite the difficulties inherent to this field of research, this is an exciting time for the study of mutation rate evolution. With the surge of computational power, intensive genomic sequencing is increasingly feasible. And it is perhaps fortunate that mutation rate evolution was a topic of theoretical inquiry long before the data deluge, because population geneticists have had decades to develop the quantitative tools required for analyzing these data. But there remain certain gaps in our understanding of mutation rate evolution. Much of our current knowledge of the evolution of mutators comes from populations of wild-type origin that evolved mutator phenotypes. As discussed, this topic has recently been investigated theoretically and experimentally.

This dissertation explores what happens to a population isogenic for a mutator at the start of selection – i.e., the mutator has fixed long before the potential to acquire mutations beneficial in that environment have been exhausted. Does the outcome matter to some degree on the type of selection? In the next chapter (which has been submitted as a manuscript for publication), I address the effects of repeated lethal selection, via a sequence of different antibiotics, on mutator (*mutL*$^-$) populations. I hypothesized that the successive bouts of lethal selection would increase the frequency of any spontaneously arising additional mutator alleles, due to the great need for evolutionary novelty (in this case, antibiotic resistance mutations) in a harsh, changing environment. I measured the mutation rates of clones surviving this selective regime, and I investigated the identity and effects of the new mutations via whole-genome sequencing and plasmid

complementation. In the third chapter, I address the effects of long-term soft selection, via limited glucose growth media, on 30 mutator (*mutS*$^{-}$) populations with isogenic starting genotypes. After ~900 generations of daily propagation, five clones were randomly isolated from each of the populations. Their mutation rates and fitnesses (relative to ancestor) were measured, and the genomes of 20 of the clones were sequenced. In the fourth chapter, I compare and contrast the outcomes of these two scenarios, and I discuss their relevance to biomedicine, astrobiology, experimental evolution, and fundamental evolutionary thought.

Chapter II: Mutator populations of *Escherichia coli* substitute additional, spontaneously originated mutator alleles under lethal selection.

Abstract

Recent theory suggests that rapidly adapting asexual populations will accumulate multiple mutator alleles, evolving progressively higher genomic mutation rates. If this theory is correct, then it should be possible to observe the substitution of new, spontaneously originated mutator alleles into asexual populations that already exhibit high mutation rates. To test this prediction, we exposed populations of mutator *E. coli* deficient in mismatch repair (MMR) to selection on three antibiotics in succession. The antibiotics were ordered from least strongly selective to most strongly selective. Our results support the prediction: all sampled surviving clones exhibited mutation rates approximately an order of magnitude higher than that of the mutator ancestor. Moreover, two clones' genomes were sequenced and found to harbor a spontaneous 1-bp frameshift mutator mutation in a repeat region of *mutT*, in addition to the ancestral MMR mutator allele. Complementation tests with a $mutT^{+}$ plasmid suggested that the $mutT^{-}$ mutation was indeed responsible for the elevation in mutation rates. Sanger sequencing revealed that all of the other seven clones that were randomly sampled from the final plate also harbored the same $mutT^{-}$ mutation, suggesting that the new mutator allele had risen to fixation in the population. When an independent replicate of the selection series was performed, the single surviving clone had a mutation rate elevated relative to that of the mutator ancestor. In a surprising incidence of convergent evolution, this clone also had the same 1-bp frameshift mutation in the same repeat region of $mutT^{-}$. In all ten clones

across the two experimental replicates, the ancestral MMR mutator allele was retained. Few clones with increased mutation rates are observed after only one or two rounds of lethal selection, or after three rounds of lethal selection if exposure to the most selective antibiotic instead occurs first.

Background

Because mutations affecting the phenotype are far more likely to be deleterious than beneficial (Fisher 1930), natural selection is generally expected to favor the evolution of low mutation rates (Sturtevant 1937). However, selection can elevate the mutation rate of a population given sufficient genetic linkage: under these circumstances, mutator alleles can hitchhike to fixation in wild-type populations with the new beneficial mutations that they facilitate (Raynes & Sniegowski 2014). Recent theory suggests, moreover, that a rapidly adapting asexual population is susceptible to progressive upward evolution of its mutation rate via repeated hitchhiking of mutator alleles (Gerrish et al. 2007; André and Godelle 2006). A prediction of this theory is that additional spontaneously arising mutator alleles should hitchhike to fixation in a rapidly adapting asexual population that is already fixed for one mutator allele. This prediction remains empirically untested. To test this prediction, we exposed mutator populations of *E. coli* to selection on three different antibiotics in succession and characterized the mutation rates and genomes of survivors.

Methods

Ancestral strain

E. coli K-12 strain ES568 (= PS2533) was acquired from the Yale University *E. coli* Genetic Stock Center and used as the ancestor for the selection experiments. This strain is defective for mismatch repair: it harbors the *mutL13* allele containing the point mutation A120T. We have previously estimated that the genomic mutation rate of strain ES568 is approximately 100-fold higher than that of wild-type *E. coli* (Gentile et al. 2011).

Exposure to successive rounds of lethal selection

Populations of the ancestral strain were exposed to lethal concentrations of fosfomycin (100 μg/mL), rifampicin (100 μg/mL), and streptomycin (100 μg/mL)) on lysogeny broth (LB) agar plates (Miller 1972) in succession, as follows (see figure 1): the ancestor was inoculated from frozen stock into liquid LB (without antibiotics) and grown overnight at 37 °C with shaking at 120 rpm. 100 μL (approximately 10^8 cells) of the resulting culture was then spread on a single LB agar plate (without antibiotics) and grown overnight at 37 °C, yielding a lawn. The lawn was replica-plated to an LB-fosfomycin plate and the fosfomycin plate was incubated overnight at 37 °C. The resulting fosfomycin-resistant colonies were suspended in 150 μL of sterile saline diluent pipetted directly onto the plate, and the cell suspension was spread across the same plate. The plate was incubated overnight again, resulting in a lawn of fosfomycin-resistant cells. This lawn was then replica-plated to an LB-rifampicin plate, and the process was repeated as described above, with the resulting lawn of rifampicin-resistant cells being replica-plated to an LB-streptomycin plate. The three antibiotics were deployed in increasing order of selectivity (in that spontaneous acquisition of streptomycin resistance is typically rarer than it is for

rifampicin resistance, which is typically rarer than it is for fosfomycin resistance). However, the above protocol was also performed with the antibiotic exposures in reverse order (i.e. decreasing order of selectivity): streptomycin, rifampicin, fosfomycin. Resistances to these three antibiotics are conferred through biochemically distinct modes of action (Nilsson et al. 2003; Goldstein 2014; Springer et al. 2001).

Colonies representing resistant clones arising at all stages of the experiments were randomly sampled from the antibiotic plates (if a given plate contained very few colonies, all colonies were sampled), inoculated into liquid LB (without antibiotics), grown overnight, and archived in 15% glycerol at -80 °C.

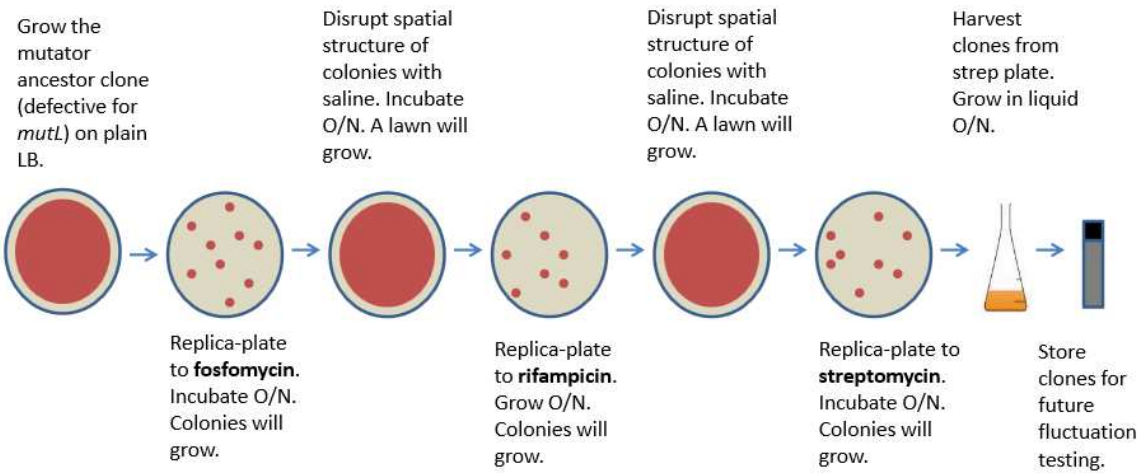


Figure 1: Overview of the antibiotic selection experimental protocol.

Mutation rate estimation

Mutation rates to nalidixic acid resistance were measured using a modified Luria-Delbrück fluctuation assay (Luria and Delbrück 1943; Gerrish 2008) in the evolved clones, the ancestor, and clones transformed with plasmids. Resistance to nalidixic acid can be conferred through mutations in *gyrA* and *parC* (Nakamura et al. 1989; Saenz et al. 2003), which prevent nalidixic acid from binding to DNA gyrase and topoisomerase IV (Saenz et al. 2003). As a control, the mutation rate of the ancestor was always measured alongside those of the evolved and transformed clones. For each assay, 10 mL of liquid LB was inoculated from a frozen stock of the clone of interest and incubated overnight at 37 °C with shaking at 120 rpm. The resulting culture was then diluted 100,000-fold, and 100-μL aliquots (representing approximately 1000 cells) of this dilution were transferred to each of a set of replicate flasks (three replicates for the evolved and transformed clones and six replicates for the ancestor), each containing 30 mL of LB. The resulting cultures were incubated for 48 h at 37 °C with shaking at 120 rpm. To enumerate mutants, 100 μL from each 30 mL culture was spread on an LB plate containing 25 μg/mL of nalidixic acid. To estimate the total population sizes of the cultures, another 100 uL from each culture was diluted 1,000,000-fold and 300 μL of this dilution was spread on an LB plate without nalidixic acid. All plates were incubated at 37 °C. Colonies were enumerated on both sets of plates at 24 h. To account for the possibility of slowly growing resistant mutants, the LB-nalidixic acid plates were incubated another 24 h and colonies were again enumerated. The inferred total number of mutants in each culture never approached the number expected if a pre-existing mutant had been seeded into the culture (see table

S3), indicating that all observed mutants originated from mutations that occurred during culture growth as assumed by the fluctuation assay.

Maximum likelihood mutation rates and associated 95% confidence limits were calculated from the 48 h LB-nalidixic acid counts and the 24 h permissive LB counts with a program kindly provided by Dr. Philip Gerrish (Gerrish 2008); significance of mutation rate differences was inferred from non-overlap of 95% confidence limits (Zheng 2015).

Genomic sequencing

Genomic DNA from two random clones (designated $FRS1_3$ and $FRS1_6$) that survived exposure to all three antibiotics in the first replicate series (concluding with streptomycin) was extracted for sequencing using the QIAGEN DNeasy Blood & Tissue Kit (69504) according to the manufacturer's instructions. Library preparations and genome sequencing were performed by the University of Pennsylvania's Next Generation Sequencing Core using the Illumina Nextera XT DNA library preparation kit (FC-131-109), the Illumina Nextera XT index kit (FC-131-1002), and Illumina NextSeq (150 paired-end) for a mean 325-fold coverage. Mutations were detected using the computational pipeline breseq version 0.27 (Deatherage and Barrick 2014). The sequences of the evolved clones were aligned against the reference sequence, *E. coli* K-12 MG1655 (NC_000913.3, downloaded from https://www.ncbi.nlm.nih.gov/nuccore/NC_000913.3?report=gbwithparts&log$=seqview on July 15, 2016).

Complementation tests

Cells were rendered competent as follows: liquid inocula of the single-mutator ancestor, evolved clone FRS1$_3$, and evolved clone FRS1$_6$ were centrifuged for 10 minutes at 13,000 rpm at 4 °C. Supernatants were discarded and cells were resuspended in 100 mM $MgCl_2$ and incubated on ice for 20-30 minutes. Cells were then centrifuged for 10 minutes at 7,000 rpm at 4 °C. Supernatants were discarded and cells were resuspended in 100 mM $CaCl_2$ with 15% glycerol before they were frozen at -80 °C.

Competent cells (the single-mutator ancestor, evolved clones FRS1$_3$ and FRS1$_6$, and a positive control of ThermoFisher Scientific *E. coli* Subcloning Efficiency™ DH5α™ Competent Cells) were transformed with a *mutT*$^+$ *amp*^{R} plasmid and an empty *amp*^{R} plasmid, yielding a total of eight transformant groups. The *mutT*$^+$ plasmid pSK25 (Bhatnagar and Bessman 1988) was isolated from a strain of REL606 *E. coli* (Jeong et al. 2009) via a QIAprep spin miniprep kit from QIAGEN (product number 27104). The empty plasmid pBR322 was purchased from ThermoFisher Scientific (product number SD0041).

Transformations were conducted as follows: competent cells were thawed on ice and incubated in the presence of plasmid DNA for 30 minutes, then heat-shocked for 20 seconds in a 42 °C water bath, then incubated on ice for 2 minutes. 950 μL of liquid LB was added to each transformation reaction before the reactions were incubated at 37 °C

for 1 hour at 120 rpm. The transformation reactions were plated on LB containing ampicillin at a concentration of 50 μg/ml. Two colonies were harvested from each ampicillin plate, inoculated into liquid LB-amp, grown overnight, and archived in 15% glycerol at -80 °C. Mutation rates of the transformants were measured via fluctuation assays using nalidixic acid as the selective agent; all LB-based media contained ampicillin.

Polymerase chain reaction (PCR) and Sanger sequencing

Genomic DNA of all surviving clones of the FRS triple selection series was extracted for sequencing using the QIAGEN DNeasy Blood & Tissue Kit (69504) according to the manufacturer's instructions. DNA concentration was quantified with a Qubit fluorometer. Primers for the *mutT* repeat region of interest were designed as follows: forward, GCGCACATGGCGAATAAAC; reverse, TTCATTGGCTGGCGGAAA. Primers for the *mutL* gene were designed as follows: forward, CAGCAACAACAGCGAAGAA; reverse, CGGCCCCATCAAAAAAAT. Each PCR reaction was assembled as follows: 25 uL iProof master mix, 18 uL nuclease-free water, 1 uL 10 uM respective (*mutT* or *mutL*) forward primer, 1 uL 10 uM respective reverse primer, and 5 uL of 10 ng/ul DNA extraction. The *mutT* PCR thermocycler program was 98 degrees C for 3 minutes, then 30 rounds of the following: 98 degrees at 10 seconds, 55 degrees at 30 seconds, and 72 degrees at 10 seconds; the final extension period was 72 degrees for 5 minutes. The *mutL* program was 98 degrees C for 3 minutes, then 30 rounds of the following: 98 degrees for

10 seconds, 54 degrees for 30 seconds, and 72 degrees for 2 minutes 5 seconds; the final extension period was 72 degrees for 5 minutes.

After PCR was completed, each PCR sample clean-up was assembled as follows: 17 uL of PCR sample, 0.3 uL Exonuclease I, 0.3 uL Antarctic phosphatase, and 2 uL nuclease-free H2O, which was run in a PCR thermocycler for 15 minutes at 37 degrees C, then 15 minutes at 80 degrees. Cleaned-up PCR samples were then submitted for Sanger sequencing. The chromatograms were viewed with the software Chromas.

Results

Two replicate runs were performed. In the first run, numerous clones survived and nine random clones were sampled; in the second run, one clone survived and was sampled. Mutation rates were significantly and similarly elevated (approximately one order of magnitude above that of the mutator ancestor and 1,000-fold above wild-type) in all 10 of the sampled endpoint clones (figure 2a; appendix, table S1). Genomic sequencing of two clones from the first experimental run confirmed the presence of a second mutator allele in addition to the original *mutL* mutation: a frameshift single-base insertion (C) in a 6-bp repeat region of *mutT* (see appendix, figure S2) (Fowler and Schaaper 1997). Expression of wild type MutT in these two clones restored the ancestral mutation rate, confirming the new *mutT* allele as a mutator (figure 2b). PCR and Sanger sequencing of the ten surviving clones sampled across the two independent experimental runs showed that all clones had the same 1-bp insertion in the same repeat region of *mutT*. In all ten surviving clones across the two experimental replicates, the original *mutL*$^-$ mutation was retained.

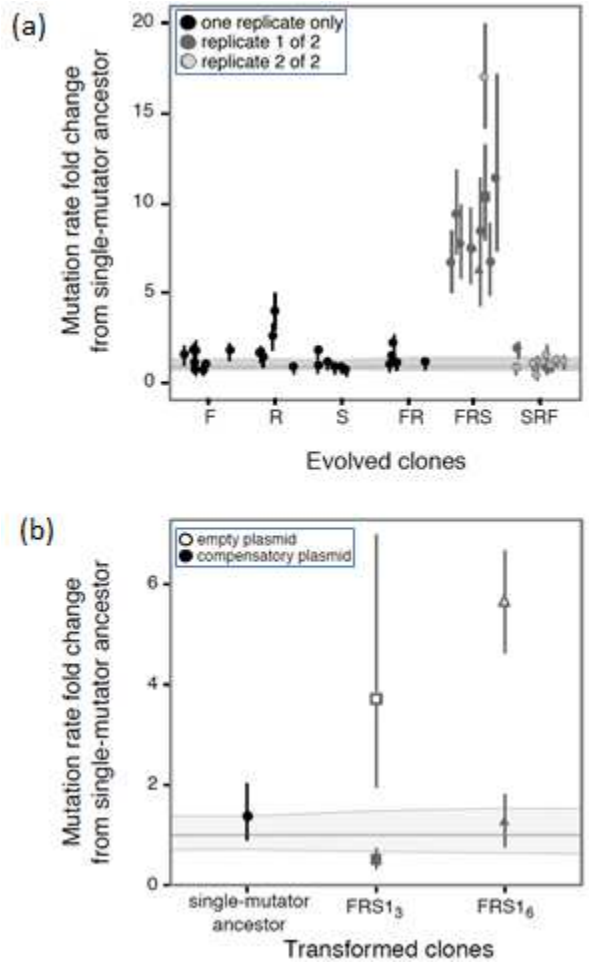


Figure 2a-b. Mutation rates to nalidixic acid resistance for all measured clones. The thick horizontal gray line represents the ancestral mutation rate to nalidixic acid resistance relative to itself; the thin lines above and below it represent the 95% confidence limits on the ancestral mutation rate. Error bars on the individual points represent 95% confidence limits on the mutation rates obtained from fluctuation assays, normalized by dividing each by the ancestral mutation rate. (See Table S1 for absolute mutation rate and confidence limit values.)

(2a): Mutation rates in clones post-antibiotic exposure, plotted relative to the mutation rate of the single-mutator ancestor. F: fosfomycin exposure only; R: rifampicin only; S: streptomycin only; FR: fosfomycin,

then rifampicin; FRS: fosfomycin, then rifampicin, then streptomycin; SRF: streptomycin, then rifampicin, then fosfomycin. The data include numerous different groups of fluctuation assays in which the ancestor was always measured alongside the evolved clones. Although the ancestral mutation rate estimate was approximately constant, its confidence intervals varied across assays. For each treatment (F, R, S, etc.), the 95% confidence limits plotted for the ancestral mutation rate represent the highest and lowest values obtained from the relevant group of assays. For FRS and SRF, the shading difference indicates clones isolated from two different replicate experiments. In FRS, the square and triangle represent two clones ($FRS1_3$ and $FRS1_6$, respectively; see appendix) whose genomes were sequenced.

(2b): Transformation with a *mutT*$^+$ plasmid restores the ancestral mutation rate in clones that had evolved elevated mutation rates. Here the thick gray line represents the mutation rate of the ancestor transformed with an empty control plasmid. Filled circle: mutation rate in the ancestral clone transformed with a *mutT*$^+$ plasmid. Squares and triangles represent mutation rates (relative to the ancestral rate) in two evolved clones with elevated mutation rates that were both found to harbor a *mutT* frameshift. Solid square and triangle: mutation rates after transformation with *mutT*$^+$ plasmid; open square and triangle: mutation rates after transformation with empty control plasmid.

Genomic sequencing uncovered non-synonymous mutations putatively responsible for antibiotic resistance in the two clones, as expected: $FRS1_3$ has a mutation (T36P) in *ptsH*, encoding a phosphocarrier protein, and $FRS1_6$ has a mutation (Y121D) in *uhpA*, encoding a transcriptional regulator; mutations in these genes can potentially cause fosfomycin resistance (Nilsson et al. 2003; Waygood et al. 1987). Furthermore, both clones exhibit substitutions in *rpoB*, encoding the β subunit of RNA polymerase, that may confer rifampicin resistance (Jin and Gross 1988): $FRS1_3$ has the mutation S512P and $FRS1_6$ has the mutation Q513P. Finally, both clones have the mutation K43T in *rpsL*; mutations in the 40-43 region of *rpsL* can confer streptomycin resistance (Timms et al. 1992). For additional information on the mutations in both clones, see figure S1 and tables S2a-c.

Additional runs were performed with selection on single antibiotics and selection in the reverse order of antibiotics (i.e., streptomycin, rifampicin, fosfomycin); these produced significantly lower frequencies of clones with elevated mutation rates compared to the ancestor (see figure 2a and appendix).

To test to what degree one round of lethal selection resulted in further increases in mutation rate, we characterized the mutation rates of randomly sampled clones surviving exposure to fosfomycin only (eight clones), rifampicin only (five clones), and streptomycin only (six clones). Two of these 19 single-exposure clones (both from the single-rifampicin exposure) displayed statistically significant increases in mutation rate compared with the mutator ancestor (figure 2a; table S1). The difference in the

prevalence of increased mutation rates between this outcome and the outcome of the original three-antibiotic experiment (10 of 10 with further increases in mutation rate) is highly significant (Fisher exact test: two-tailed $p = 3x10^{-6}$). To test to what degree two rounds of lethal selection resulted in increased mutation rates, we characterized the mutation rates of five random clones surviving exposure to fosfomycin and then rifampicin. One clone showed a statistically significant increase in mutation rate compared with the mutator ancestor (figure 2a; table S1). The difference in the prevalence of increased mutation rates between this outcome and that of the original three-antibiotic experiment (10 of 10 with further increases in mutation rate) is significant (Fisher exact test: two-tailed $p = 0.004$). Finally, to test whether the order of exposure to the three antibiotics affected the prevalence of increased mutation rates, we characterized mutation rates of randomly sampled clones surviving sequential exposure to the same three antibiotics, but in reverse order, from most to least selective: streptomycin, rifampicin, fosfomycin. This reverse-order series was repeated for a total of two experimental replicates. Nine clones were randomly sampled from the final plate of the first replicate, and two clones were randomly sampled from the final plate of the second replicate. One of these 11 reverse-order clones displayed a statistically significant increase in mutation rate compared with the mutator ancestor (figure 2a; table S1). The difference in the prevalence of increased mutation rates between this outcome and that of the original three-antibiotic experiment (10 of 10 with further increases in mutation rate) is highly significant (Fisher exact test: two-tailed $p = 3x10^{-5}$).

Discussion

Previous experimental work has provided qualified support for the theory that mutation rate evolution in rapidly adapting asexual populations can be upwardly biased (Gerrish et al. 2007; André and Godelle 2006) by showing that a double-mutator *E. coli* strain (bearing two mutator alleles with cumulative effects on the genomic mutation rate) introduced into experimental populations at substantial frequency can hitchhike and displace a single-mutator strain (Gentile et al. 2011). The generality of that previous result was limited by the fact that new mutator alleles must originate in natural populations as rare spontaneous mutations rather than via experimental introduction at substantial frequencies. The results we have presented here overcome this limitation by demonstrating that additional, spontaneously arising mutator alleles can hitchhike to apparent fixation in populations already fixed for one mutator allele. Only one clone survived our second experimental run and it exhibited an elevated mutation rate in comparison with the ancestor, consistent with fixation of a new mutator allele in this population. All nine random clones sampled from the first run exhibited elevated mutation rates, strongly suggesting high frequency or fixation of an additional mutator or mutators in this population; the probability that all nine randomly sampled clones from this population would exhibit elevated mutation rates is < 0.05 for new mutator frequencies less than or equal to 0.7. Genomic sequencing and complementation test results indicate that a *mutT* mutator mutation was responsible for the high mutation rates of two clones sampled from the first experiment. Sanger sequencing of the other seven clones (that were randomly isolated from the same antibiotic plate) showed that they harbored the same 1-bp insertion in the same repeat region of *mutT*. A parsimonious explanation for the elevated mutation rates observed in all nine clones in the first

experimental run is that they are caused by the same mutator allele. To our knowledge, ours is the first observation of a second spontaneous mutator hitchhiking to fixation in a mutator population, though Kinnersley et al. have reported a double mutator genotype present at polymorphic frequency in one mutator experimental *E. coli* population (Kinnersley et al. 2014).

The single surviving clone from our second replicate run was also shown to have a 1-bp insertion in the same repeat region of *mutT*; we note that the magnitude of the added mutator effect is similar between the two experimental runs. This genomic parallel between the survivors of the two existing experimental replicates has precedence in the literature: interestingly, the frameshift producing the new *mutT* mutator in our first experiment lies in the same cytosine string as a *mutT* mutation that spread in a bacterial population descended from a wild-type ancestor that had evolved experimentally for 20,000+ generations (Barrick et al. 2009). This parallel is consistent with the higher mutation rate expected in such repeated sequences (Shaver and Sniegowski 2003).

Exposure of *E. coli* to sublethal concentrations of fosfomycin may mildly enhance the frequency of mutants resistant to lethal rifampicin concentrations (Thi et al. 2011). Although there is no evidence that exposure to lethal fosfomycin concentrations enhances mutagenesis in fosfomycin-resistant survivors, it is nonetheless possible that such an effect could have enhanced the likelihood of mutator hitchhiking in our experiments. Additionally, fosfomycin was the least selective of our antibiotics: when comparing the number of colonies after single-exposure to each antibiotic, fosfomycin exposure resulted

in the highest number of survivors. The series beginning with streptomycin (the most selective antibiotic) did not result in detectable increases in mutation rate (figure 2a). A low number of survivors would conceivably decrease the likelihood of a new mutator being present on the streptomycin plate, making that first selective bout a lost opportunity for increased mutation rates.

There are reasonable scenarios in which natural populations will encounter repeated lethal or very strong selection, as employed here and in similar previous work (Mao et al. 1997) —most notably, through host-pathogen interactions and coevolution (Pal et al. 2007; Graves et al. 2013). It is currently unclear whether any natural populations under such strong selection have substituted multiple new mutator alleles; this remains an avenue for further study. In ongoing work, we are also testing whether non-lethal selection can cause a rapidly adapting asexual mutator population to substitute additional spontaneously originated mutators. An obvious further step would be to test whether rapidly and continually adapting populations that are initially fixed for low, wild-type mutation rates can substitute multiple mutator alleles, as predicted by theory (Gerrish et al. 2007). Simulations of this process under non-lethal selection, however, suggest that it is likely to take a prohibitively long time to unfold in the laboratory (Gerrish et al. 2007). Using repeated rounds of lethal selection would be quicker, but is potentially limited by the number of lethal selective conditions to which cells would not exhibit cross-resistance.

Our results have potential biomedical implications apart from their relevance to the general understanding of mutation rate evolution and the fate of asexual populations. Mutator lineages have been identified in pathogenic bacterial populations (LeClerc et al. 1996) and are also implicated in certain forms of cancer, both in clinical studies (Loeb and Loeb 2000) and mathematical models (Solé and Deisboeck 2004). Both therapeutic intervention and the lethal, relentlessly specific host immune response in pathogen infections and cancers could conceivably result in the substitution of additional mutator alleles into such populations, as shown by our results. The implications of such progressive elevation of the mutation rate in clinical contexts may be frightening or encouraging. Pathogens or tumors with extremely high mutation rates might be highly treatment-resistant because they easily acquire further resistances (Sprouffske et al. 2012). Alternatively, additional mutagenesis could potentially lead to the demise of such populations via the influx of deleterious mutations; the latter possibility has been considered for RNA viruses, which already tend to have extraordinarily high genomic mutation rates (Drake and Holland 1999).

Data accessibility

Source data will be available in Dryad (DOI:10.5061/dryad.k76d4) upon publication of the manuscript that is associated with this chapter (http://datadryad.org/review?doi=doi:10.5061/dryad.k76d4) (Eghbal et al. 2017). The Genbank accession number for the *E. coli* (MG1655) genome is NC_000913.3; *mutT* spans nucleotides 111044-111433. Mutations were detected via the computational pipeline breseq (Deatherage and Barrick 2014).

Acknowledgements

We thank Andreas Wagner for sharing the ancestral genomic sequence (PS2533).

Chapter III: Investigating mutation rate evolution in experimental mutator *Escherichia coli* populations under long-term soft selection.

Abstract

In this study of asexual mutators under long-term soft selection, 30 near-isogenic *mutS*- populations were propagated via daily batch transfer in minimal glucose media. After ~900 generations of evolution, five clones were randomly isolated from each population and their mutation rates were measured via fluctuation assays. Change in mutation rate was determined by non-overlap of 95% confidence intervals with the mutator ancestor. There was extensive polymorphism for mutation rate across populations and within many populations: 9% had increased mutation rates relative to the mutator ancestor, 23% had decreased mutation rates relative to the mutator ancestor, and 68% had unchanged mutation rates. Most populations had at least one clone whose mutation rate had either increased or decreased. Additionally, in one population, all five clones' mutation rates were decreased, and in another population, all five clones' mutation rates were increased, suggesting potential fixation of a new antimutator and mutator, respectively. The fitnesses of all evolved clones and populations were elevated relative to the ancestor. No relationship was detected between mutation rate and relative fitness. Whole-genome sequencing of 20 evolved clones, representing a range of mutation rates and populations, revealed that several of the sequenced clones with altered mutation rates had mutations in known (anti)mutator loci. PCR and Sanger sequencing of a few of the suspected (anti)mutator loci in some of the clones whose genomes had not been sequenced suggested, in several cases, a relationship between the presence of several suspected

(anti)mutator mutations and changes in mutation rates. I propose that the increased mutation rates were the result of hitchhiking of new mutator alleles and that the decreased mutation rates were most likely the result of new antimutator alleles that rose in frequency due to the conferring of pleiotropic fitness benefits.

Introduction

In 1937, Alfred Sturtevant proposed that the mutation rate itself can evolve, and he inquired why the mutation rate does not decline to zero (Sturtevant 1937). Indeed, theory predicts that in an unchanging environment, barring any physicochemical limitations, the mutation rate will evolve to zero (Feldman & Liberman 1986). Yet the mutation rate has never been observed to evolve to zero in any laboratory or natural population. This is unsurprising if we consider the fact that beneficial mutations, despite their rarity, are the ultimate source of all evolutionary novelty. The drift-barrier hypothesis and the cost of replication fidelity have also been proposed as explanations for why mutation rates are greater than zero; refer to Chapter I for a more detailed description of these hypotheses (Lynch 2010; Kondrashov 1995).

The genomic mutation rate among most DNA-based haploid microbes is relatively constant (Drake 1991; Sniegowski and Raynes, 2013), yet variant mutation rates have been detected among many strains and individuals. In an asexual population, selection could hypothetically impel a spontaneously arising mutator (a variant allele that raises genomic mutation rate) to fixation through one of two different processes: by conferring a direct pleiotropic benefit to fitness or by mutator hitchhiking. Thus far, a mutator has

been shown to confer a direct fitness benefit in only one laboratory setting, involving *Pseudomonas aeruginosa* subjected to oxidative stress (Torres-Barceló et al. 2013). In an experimental population with a wild-type mutation rate, a spontaneously emerging mutator allele can hitchhike to fixation; due to the absence of recombination, the new mutator remains linked to the beneficial mutations it causes, thereby rising in frequency in the population (Sniegowski et al. 1997). The deleterious mutational load accumulates gradually and may not be realized until much later (Gerrish et al. 2007). In an adapting population, selection could conceivably impel a spontaneously arising antimutator to fixation through one of at least two different mechanisms: the antimutator could confer a direct pleiotropic benefit to fitness or there may be a strong advantage to reducing the genetic load. Direct fitness benefits stemming from an emergent antimutator allele have been proposed to exist in only one experimental scenario, involving a patient isolate of mutator *Escherichia coli* that underwent a reduction in mutation rate, despite the ancestral *mutS*⁻ mutator allele remaining intact in the descendants (Turrientes et al. 2013). The benefit of reducing the genetic load has been observed in a mutator population that was invaded by an antimutator allele after tens of thousands of generations of propagation under soft selection exerted by a limited glucose environment (Wielgoss et al. 2013).

Theoretical models suggest that a continually adapting asexual population may undergo repeated mutator hitchhiking events (André and Godelle 2006; Gerrish et al. 2007). Experimental work demonstrates that in an adapting asexual population, a double-mutator strain will ultimately supplant a single-mutator strain, if the double-mutator is

seeded in at the start of propagation above a certain threshold frequency (Gentile et al. 2011). In a *mutY*$^{-}$ mutator population evolving in a limited glucose environment for 765 generations, a *mutM*$^{-}$ mutator (causing a 10-fold increase) was shown to have spontaneously arisen in at least one clone, although it was not present across the whole population (Kinnersley et al. 2014). The spontaneous emergence and fixation of a second mutator allele (i.e. hypermutator) has been observed in asexual populations under repeated exposure to lethal antibiotics (Eghbal et al. 2017, in review). However, the advantage of a hypermutator phenotype can be readily discerned in this scenario, as the populations were under immense and immediate pressure to acquire antibiotic resistance. This finding raises the question of whether a hypermutator could spontaneously emerge and rise to fixation in a preexisting mutator asexual population under a non-lethal, soft selective regime.

The answer is far from obvious: a higher mutation rate need not always lead to a higher rate of adaptation, in part due to the potential for clonal interference (Sprouffske et al. 2012). Furthermore, Muller's ratchet, characterized by the gradual but irreversible accumulation of deleterious mutations in asexuals, has been shown to cause fitness deterioration in bacteria (Andersson and Hughes 1996). An extremely elevated mutation rate carries with it the risk of attrition of genomic information (Eigen and Schuster, 1977). Although mutators have been observed in carcinogenic human cell populations and in virulent microbial populations, pathogenic cells are not universally mutators (Jones et al. 2008; Tomlinson et al. 1996).

Here we describe the evolutionary paths of 30 mutator populations defective for mismatch repair (*mutS*$^-$) that underwent 900 generations of propagation in minimal glucose medium. We randomly isolated five clones from each evolved population and we measured their mutation rates and their fitnesses relative to their ancestor. We sought to characterize the genomic changes in 20 evolved clones across eight populations, which represented a broad range of mutation rates.

Methods

Origin of experimental strains

We established 30 populations of high-mutating asexual *E. coli*. These starting populations were isogenic except with respect to *ara*, an operon that codes for an enzyme enabling the breakdown of arabinose. 15 populations have the Ara$^-$ phenotype and the other 15 populations have the Ara$^+$ phenotype. The Ara$^+$/Ara$^-$ distinction is neutral with regard to fitness in glucose minimal medium (the medium of propagation). The Ara$^-$ colonies are red on tetrazolium arabinose (TA) agar, and the Ara$^+$ colonies are white. These color differences were used as markers in competitive fitness assays, as described later.

The strain PS174, which was previously constructed in the Sniegowski lab, possesses the *mutS*$^-$ mutator allele – identical in sequence to the *mutS*$^-$ allele that spontaneously emerged in an evolving population in Richard Lenski's long-term evolution experiment (LTEE) (Sniegowski et al. 1997) – on the Lenski ancestral background (strain B REL606). Therefore, although *mutS*$^-$ can cause mismatch repair defects in both

experimental contexts, our experiment is not simply a continuation of Lenski's LTEE. Upon initiation of this project, PS174 was plated onto minimal arabinose agar, and an Ara^+ revertant colony (PS2717) was isolated. Thus, PS174 is the ancestor of all of the evolving Ara^- strains in our project, and PS2717 is the ancestor of all of the evolving Ara^+ strains.

Daily propagation of experimental strains

The 30 strains were propagated daily in two 96-well plates. One plate contained the 15 Ara^- populations, and the other plate contained the 15 Ara^+ populations. Each well bore 1.5 mL of Davis minimal broth medium (Carlton and Brown 1981) with glucose at 166.7 μg/mL. To inhibit cross-contamination, the rows of populations were spaced apart from each other across each plate. To inhibit external contamination, all populations were located away from the periphery of each plate.

The populations were housed inside a 37 degree C incubator shaking at 120 rpm. Each day, 15 μL from each population were transferred into fresh medium in the corresponding well on a fresh sterile plate. All transfers were conducted inside a sterile laminar flow hood. Cross-contamination (the most likely source of contamination) could be easily detected by plating the populations on TA agar.

Each day, 12 extra wells of plain DM166.7 were added to the lower border of both plates to hinder evaporation of the experimental wells; these wells of plain DM166.7 were also transferred daily in an identical manner as the experimental populations. Samples from

each population were periodically frozen in 15% glycerol and stored at -80 degrees C (on average, every ~90 generations). Liquid from the plain wells was periodically plated on permissive LB agar plates to detect any global contamination in the medium, as the plain wells ideally should not yield any growth. In the few instances that a plain well yielded growth, the populations of that entire plate were restarted from their most recent frozen timepoints.

Clonal isolation of the evolved experimental strains

At approximately 900 generations, five clones were randomly isolated from each of the 30 evolving populations. The isolation was accomplished by plating each of the evolved populations on permissive LB agar, incubating the plate overnight, and isolating a colony at random with a sterile implement; the dilution-streaking was repeated three times in total, to ensure the purity of the isolated colony.

Fluctuation tests (FTs) to measure mutation rates

Each clone's mutation rate was measured by fluctuation tests, using the antibiotic nalidixic acid as a selective agent. In our modern adaptation of the classic fluctuation test (Luria and Delbrück 1943), each strain begins as a small number of cells in multiple replicate cultures. After several generations of growth, the cultures are exposed to a selective agent, such as an antibiotic. Most cells are sensitive to the agent and therefore are killed. However, a fraction of cells survive the selective agent, because they have spontaneously acquired mutations conferring resistance. If a resistance mutation emerges early in the lineage of cells (pre-exposure), there will be a greater number of resistant

descendants than if the mutation had occurred later in the lineage. The numbers of colonies on the plates are used to calculate the mutation rate for each strain. In the FTs described herein, mutations conferring resistance to nalidixic acid served as a proxy for the genomic mutation rate. A diagram of the experimental procedure can be found in Figure 3.

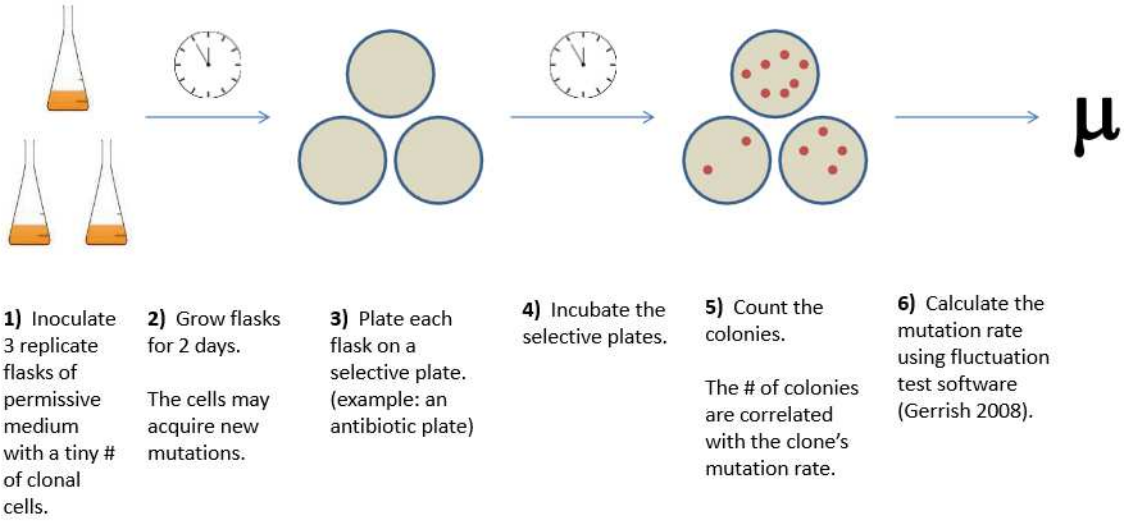


Figure 3: A diagram illustrating the methodology of a fluctuation assay, which is used for estimating the genomic mutation rate of a clone.

The FT methodology and analytical software can be found in Gerrish 2008 (Gerrish 2008). For each clone whose mutation rate was being measured, an ancestral clone was assayed alongside for the sake of comparison. (The mutation rate measurement may vary slightly from day to day, due to factors such as minor changes in the environment.) The mutation rates of the evolved clones were compared to that of the mutator ancestor; statistical significance of differences in mutation rates was inferred from non-overlap of 95% confidence intervals (Zheng 2015).

On day 1 of the FT, flasks of DM1000 (Davis minimal broth medium containing glucose at 1mg/mL) were seeded with one frozen clone stock per flask; the cultures were grown overnight to a total of 1 x 10^{10} cells/flask. On day 2, cells from each flask were serially diluted through tubes of sterile saline. From the last dilution, a small number of cells were transferred into each of three new flasks of DM1000 (30 mL per flask), for a total of three replicates per evolved strain. (The ancestor was assayed with six replicates.) The dilution assured that each replicate flask started with a small population, making it highly unlikely that a mutant is seeded in the starting population purely by chance.

The flasks incubated for 48 hours. A diluted fraction of each replicate flask's culture was plated on permissive LB agar, to estimate the total population size of each flask. Additionally, an undiluted fraction of each replicate flask's culture was plated undiluted on LB agar containing nalidixic acid, to estimate the number of resistant cells in each flask. All plates were incubated at 37 degrees C. After 24 hours from the start of plate

incubation, the colonies on both the permissive and selective LB plates were enumerated. The selective plates were incubated for another 24 hours, and then their colonies were enumerated a second time, so as to include the slower-growing resistant colonies.

After performing fluctuation assays on the 150 clones, we decided to perform a closer characterization of the Ara-3 population. Hence, we randomly isolated an additional five clones from the Ara-3 population (again at ~900 generations).

Relative fitness competitions

To measure whether evolved strains (from generation ~900) had higher fitnesses relative to their ancestral clones (from generation 0), we performed fitness competition assays, both for the evolved clones and the evolved populations. These evolved clones were identical to the clones whose mutation rates were measured via fluctuation tests.

On day 1, the frozen evolved clones/populations and ancestral clones were started from frozen stock in flasks containing 10 mL of DM1000 (one strain per flask). On day 2, the inocula of each evolved strain (clone or population) were transferred to separate wells in 96-well plate containing DM166.7, after dilution through sterile saline; the arrangement and volume of the wells in the 96-well plates was identical to the setup of the 96-well plates during the daily propagation, aside from the total number of wells on each plate. On day 3, each of the acclimated strains were diluted with saline, then mixed 1:1 with the diluted ancestral clone of the opposite $Ara^{-/+}$ marker in a fresh well of DM166.7 in a new 96-well plate. Immediately after setting up the day 3 wells, 100 µL of the mixed culture

from each well was diluted through saline and plated on TA agar. Each strain competition was performed in triplicate. On day 4, 100 μL of each competition well was diluted through saline and plated on fresh TA plates. Additionally, the red and white colonies on the day 3 TA plates were counted. On day 5, the red and white colonies on the day 4 TA plates were counted. The red and white colony counts from day 4 and day 5 were used to calculate the selection coefficient (s) for each replicate, using the following equation from classical population genetics theory for haploids:

$$s = (1/t) * (\ln[p'(1-p)/p(1-p')])$$

where t = 6.64 (the number of generations per day), p is the percentage of colonies from evolved clones on day 4 (out of the total number of colonies on day 4), and p' is the percentage of colonies from evolved clones on day 5 (out of the total number of colonies on day 5). The mean selection coefficient was then calculated for the three replicates of each evolved strain. Each mean selection coefficient calculation was added to the working selection coefficient dataset if the standard deviation of the mean selection coefficient for each evolved strain divided by the mean of the three replicates was ≤0.15. The working selection coefficient dataset can be found in the Appendix; the other selection coefficient calculations are located in a separate table in the Appendix.

Whole genome sequencing

To identify new mutations that emerged at some point during the 900 generations, we performed whole genome sequencing on select clones after ~900 generations. We extracted genomic DNA from 20 evolved clones that represented a broad range of mutation rates across eight different populations, and also the two ancestral clones (Ara$^-$

and Ara^{+}) for comparison. (The evolved clones were among the set of 155 clones whose mutation rates and relative fitnesses were measured.) Extractions were performed using the QIAGEN DNeasy Blood & Tissue Kit (69504). DNA concentrations were quantified with a Qubit Fluorometer. Library preparation and sequencing were conducted with the Illumina Nextera XT DNA library preparation kit (FC-131-109), the Illumina Nextera XT index kit (FC-131-1002), and Illumina NextSeq (150 paired-end) with a mean coverage of ~325-fold. The sequenced genomes were aligned against the reference sequence, *E. coli* B strain REL606. All new mutations were identified with the computational pipeline breseq (version 0.27) (Deatherage and Barrick 2014). The list of the 22 sequenced clones can be found in Table 1.

Parent strain #	Strain #	Identity	MR change	# new mutations	Mutations detected in candidate mutation rate modifiers
PS174	PS3840	Ara-3, c1	Unchanged	54	***recC*** *(N354S (T-->C))*, ***topA*** *(I106T)*, **mutS** (R324R), **dnaK** (K55K), **mukB** (E1108E)
PS174	PS3886	Ara-3, c2	Unchanged	43	***rep (R615C)***
PS174	PS3887	Ara-3, c3	Unchanged	49	None found.
PS174	PS3888	Ara-3, c4	Higher	63	***miaA (E151G)*****, mukB** (E1108E)
PS174	PS3889	Ara-3, c5	Lower	60	***katG*** *(1 bp del)*, ***recC*** *(D646E)*
PS174	PS4326	Ara-3, c6	Unchanged	41	***topA*** *(A364V)*
PS174	PS4322	Ara-3, c7	Higher	49	***recC*** *(D324G)*, ***topA*** *(T110A)*, **mukB** (E1108E)
PS174	PS4323	Ara-3, c8	Unchanged	42	***nrdA (Y379H)*****, *topA*** *(A364V)*, **sbcC** (intergenic) (+302/+14).
PS174	PS4324	Ara-3, c9	Higher	53	***gyrB*** *(M461I)*, **mukB** (E1108E), **sbcC** (intergenic region) (+314/+2)
PS174	PS4325	Ara-3, c10	Lower	58	***recC (D646E),*** **recB (E1149E), oxyR** (intergenic)
PS174	PS3895	Ara-5, c3	Unchanged	65	***uvrC (L79S)*****, *priA (C30R)*****, *recN (V486A)***
PS174	PS3850	Ara-13, c1	Higher	56	***nuoH (G106D)***
PS174	PS3939	Ara-13, c5	Lower	52	***recC*** *(L286P)*, ***topA*** *(T613A)*
PS174	PS3941	Ara-14, c3	Higher	134	***gyrA (Y448C)*****, *dnaE (A518V)*****, nfi** (intergenic region) (+10/-33)
PS174	PS3944	Ara-15, c2	Lower	56	None found.
PS2717	PS3855	Ara+3, c1	Lower	64	None found.
PS2717	PS3956	Ara+3, c2	Lower	46	None found.
PS2717	PS3856	Ara+4, c1	Higher	45	None found.
PS2717	PS3960	Ara+4, c2	Higher	48	***recN*** *(G467S)*
PS2717	PS3964	Ara+5, c2	Lower	59	***radA*** *(1 bp del)*, ***nrdE*** *(Y16C)*, **rep** (D218D)
PS174	PS3870	Ara- ancestor c	N/A	20	None found.
PS2717	PS3869	Ara+ ancestor c	N/A	21	**gyrA** (R365R)

Table 1: Summary of whole genome sequence findings for clones propagated under soft selection for 900 generations. The genomes of 20 evolved clones and the two ancestral clones were sequenced. Data on the number of new mutations found via the computational pipeline breseq are shown. Also listed are the new mutations in candidate mutation rate modifier loci, along with the specific residue that was mutated, if

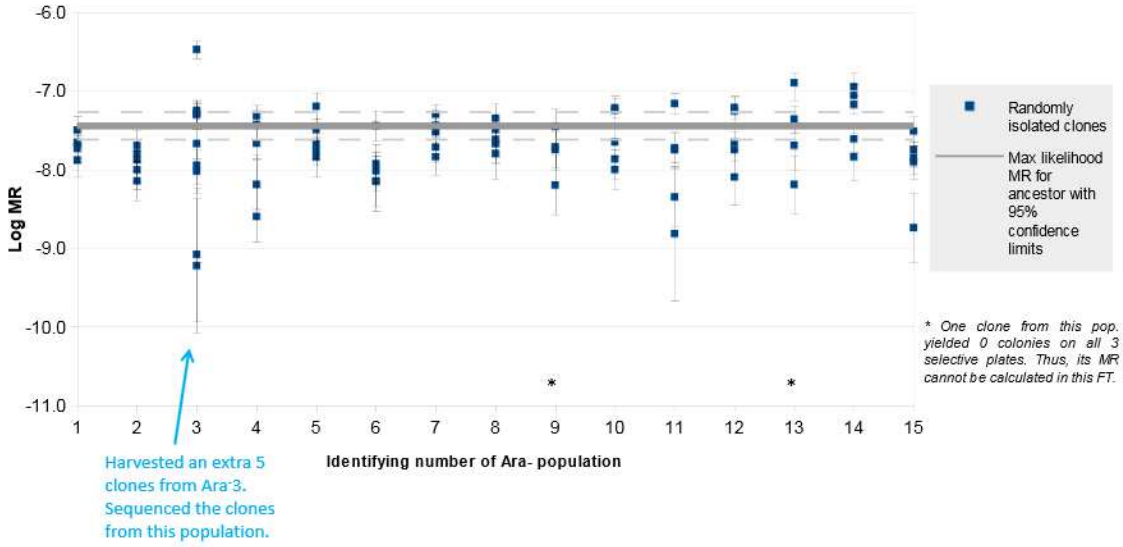
applicable. If a mutation was a non-conservative substitution or an indel in a coding region, it is underlined and italicized. If a mutation was only nonsynonymous or intergenic, it is only italicized.

Results

Mutation rates

The fluctuation assays showed that (relative to the mutator ancestor at generation zero) 9% of the evolved clones had higher mutation rates, 23% had lower mutation rates, and 68% had similar mutation rates (Figure 4: Ara^- and Ara^+ mutation rate graphs). Of the 30 evolved populations at generation ~900, 21 populations had at least one clone whose mutation rate differed from that of the single-mutator ancestor, meaning that only nine populations had zero clones with altered mutation rates. In the Ara+3 population, all five clones displayed mutation rates lower than ancestor. In the Ara+4 population, all five clones displayed mutation rates higher than ancestor. Two clones (one from the Ara-9 population and one from Ara-13) did not yield any colonies on the selective plates and therefore did not have measurable mutation rates; in qualitative classifications, their mutation rates were listed as 'lowered'. A table of fluctuation assay values can be found in the Appendix.

a)


-6.0
-7.0
-8.0
-9.0
-10.0
-11.0
Log MR
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Identifying number of Ara- population
Randomly isolated clones
Max likelihood MR for ancestor with 95% confidence limits
* One clone from this pop. yielded 0 colonies on all 3 selective plates. Thus, its MR cannot be calculated in this FT.
Harvested an extra 5 clones from Ara-3. Sequenced the clones from this population.

b)

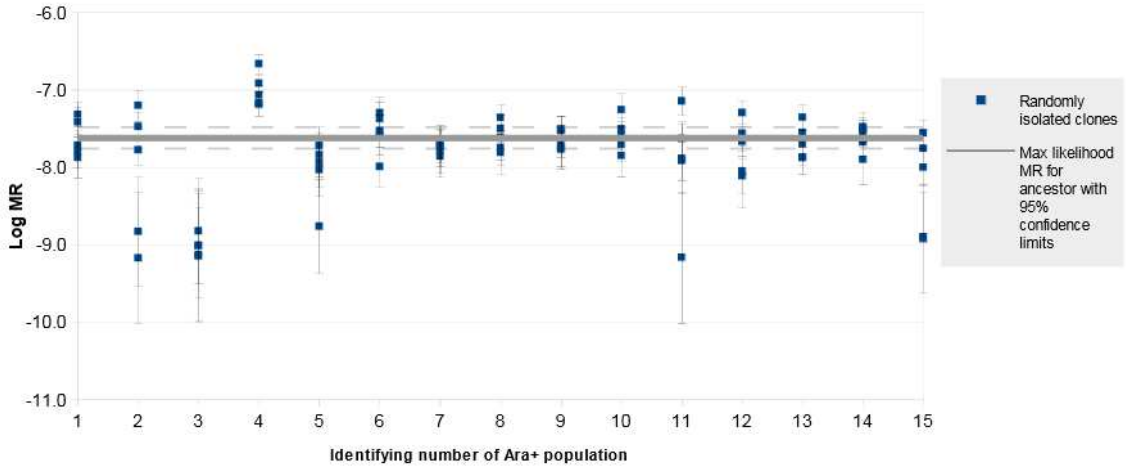


Figure 4a-b: Mutation rates for 80 evolved Ara$^-$ clones and 75 evolved Ara$^+$ clones. After 900 generations of propagation under soft selection, five clones were randomly isolated from each population (plus an additional five from population Ara-3). Their mutation rates were measured with nalidixic acid fluctuation assays. For each evolved clone, significance of mutation rate differences was determined by overlap between the 95% confidence intervals (CIs) of the clone and the ancestor, whose mutation rate was measured in the same assay batch as the evolved clone. Mutation rates are displayed on the y-axis as log values. Each clone is shown as a blue point; the brackets on each point represent 95% CIs of the clone's mutation rate. The Ara$^-$ clones are shown in Fig. 4a, and the Ara$^+$ clones are shown in Fig. 4b. For each of the two graphs, a median ancestral mutation rate (one for Ara$^-$ and one for Ara$^+$) is shown as a thick horizontal line; the dotted horizontal lines are the 95% CIs of the ancestor. A complete list of mutation rate values can be found in the Appendix.

c)

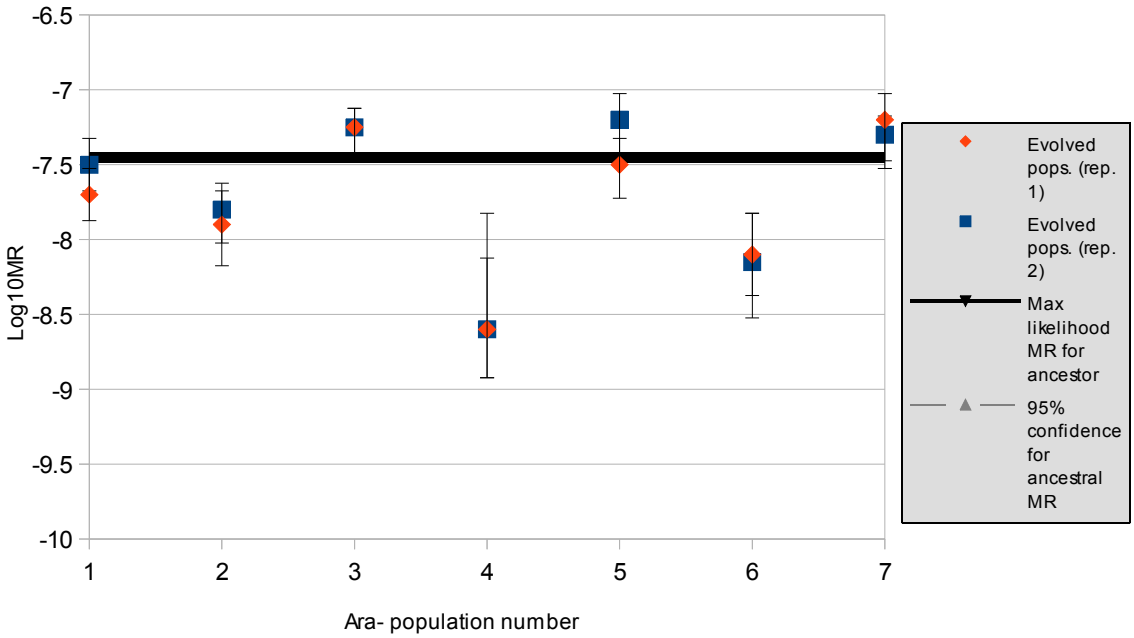


Figure 4c: Two sets of replicate fluctuation assays for a subset of the Ara$^-$ evolved clones. The mutation rates of seven clones (one from each population of Ara-1 through Ara-7) were re-measured with fluctuation assays. These second measurements are plotted here with the same x-coordinate as their respective first measurement. The two measurements of each clone were assessed for overlap across the 95% confidence intervals. The ancestral Ara$^-$ mutation rate is plotted as a thick gray horizontal line.

Relative fitness

All measured evolved populations and clones showed increased fitness relative to ancestor, although to varying degrees. In the evolved populations, the average selection coefficient was 0.28, with a standard deviation of 0.08, a median of 0.26, and a range of 0.14 to 0.41 across all populations. (See Figure 5.) Across all 155 clones, the average selection coefficient was 0.24, with a standard deviation of 0.06, a median of 0.24, and a range of 0.09 to 0.41 across all populations. (See Figure 6.) No relationship between mutation rate and selection coefficient of the clones was detected: Spearman's correlation coefficient (r_s) = 0.071; p-value = 0.808

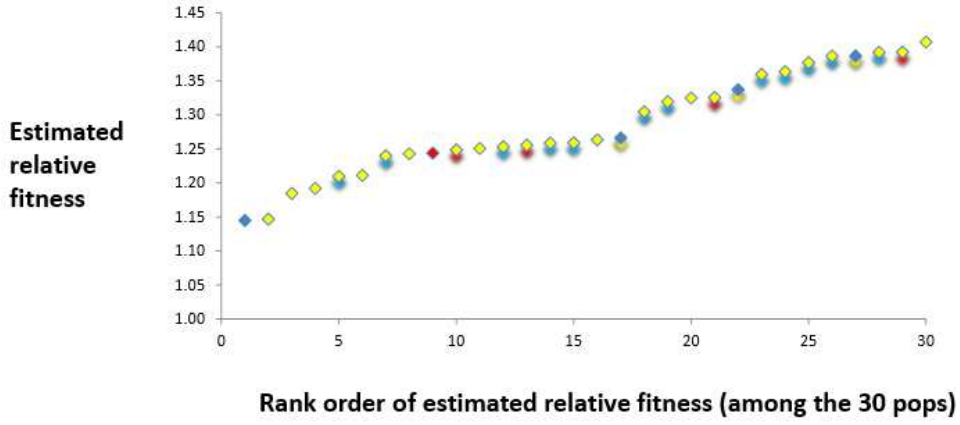


Figure 5: Estimated relative fitnesses of all evolved populations. After 900 generations of propagation under soft selection, the 30 populations' fitnesses were measured relative to that of the mutator ancestor of the opposite $Ara^{-/+}$ genotype, through overnight growth competitions, which were performed in triplicate. The estimated relative fitnesses are shown here; each point represents one evolved population. They are listed in increasing numerical order of relative fitness. The colors of each point represent the mutation rate profile of each population's five clones at 900 generations. Blue represents a lowered mutation rate, red represents an elevated mutation rate, and yellow represents an unchanged mutation rate; all mutation rates are depicted relative to the ancestor at generation 0. The 'shadows' on some of the points represent mutation rate polymorphism; for example, if a yellow point has a blue shadow, that means that most of the clones in the population had unchanged mutation rates, but a minority had lowered mutation rates. The estimated selection coefficient values can be found in the Appendix.

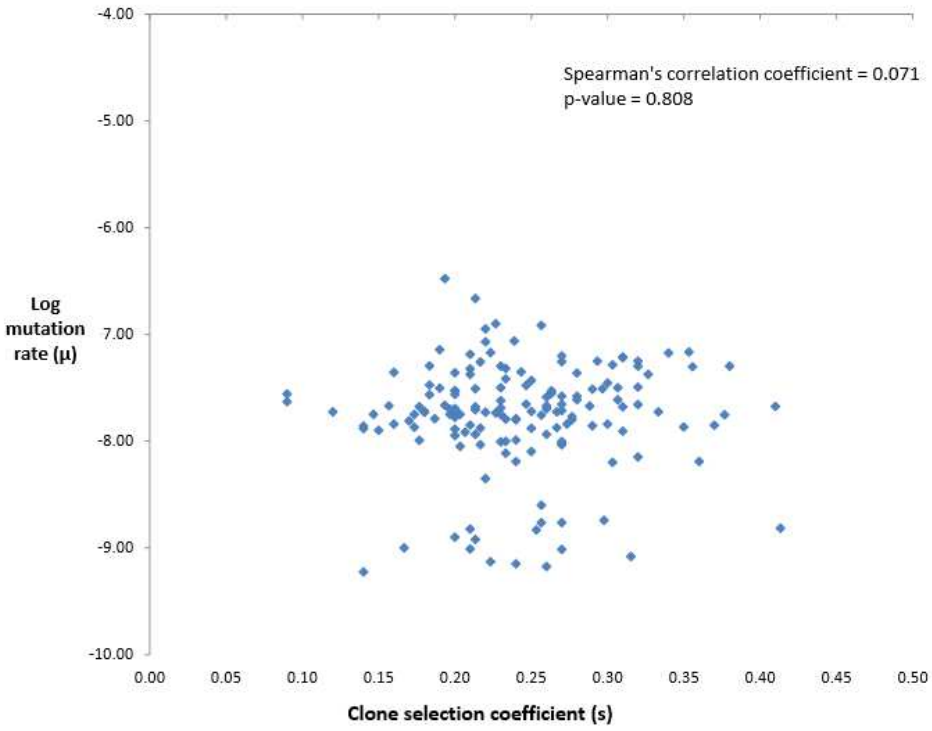


Figure 6: Mutation rate vs. selection coefficient of evolved clones after 900 generations of soft selection, across 30 populations. After 900 generations of propagation under soft selection, five clones from each of the 30 populations (plus five additional clones from population Ara-3) were randomly isolated. The fitnesses of the 155 clones were measured, relative to that of the mutator ancestor of the opposite $Ara^{-/+}$ genotype, through overnight growth competitions, which were performed in triplicate. The selection coefficients of the clones are shown here, relative to the mutation rate of the respective clone. The selection coefficient values can be found in the Appendix.

Whole genome sequencing

Lists of new mutations detected by breseq for each sequenced strain are in the Appendix. To narrow down the list of likely causative (anti)mutators, I constructed a candidate gene table of loci already described in other literature as affecting the mutation rate (see table in Appendix) (Turrientes et al. 2013). If breseq detected that an evolved clone had a new mutation in one of these loci, it was marked in Table 1. As seen in Table 1, multiple nonsynonymous mutations were detected in multiple loci that have been described in the literature as potentially affecting the mutation rate. The supplemental genomic data and genomic sequences will be deposited in Dryad Digital Repository. Tables of breseq outputs for each sequenced clones are located in the Appendix.

I sought to determine whether there was a correlation between mutation rate and number of detected new mutations at generation ~900. The Spearman's correlation coefficient (r_s) was -0.06, with a p-value of 0.60, which does not suggest a relationship between mutation rate and number of detected new mutations at generation ~900. The dataset used in this analysis consisted of the 2 ancestral clones and 19 sequenced evolved clones (one sequenced evolved clone was excluded from this analysis because its mutation rate was too low to be measured). A visual representation of the changes in mutation rate versus number of new single nucleotide polymorphisms (SNPs) is shown in Figure 7. A visual representation of the number of new SNPs in each clone, grouped by mutational class, is shown in Figure 8.

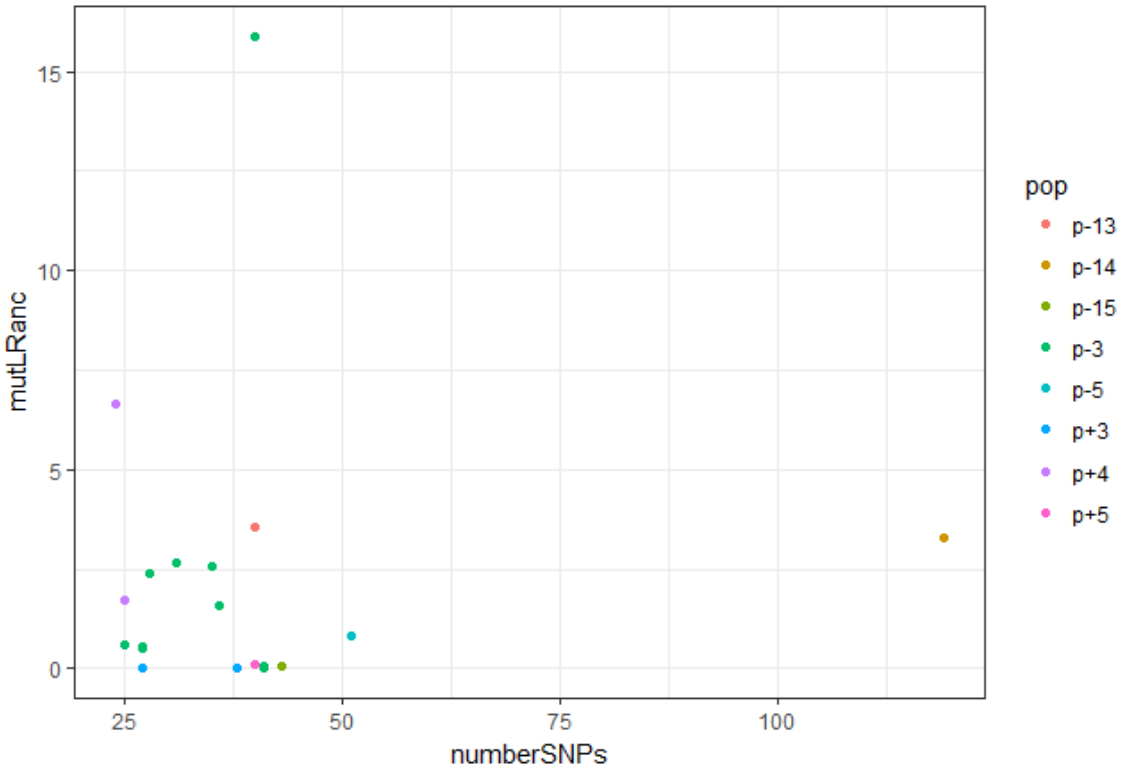


Figure 7: Mutation rate (-fold change relative to mutator ancestor) versus number of new single nucleotide polymorphisms (SNPs) per sequenced clone. After 900 generations of propagation under soft selection, five clones were randomly isolated from the evolving 30 populations, and their mutation rates were measured. The genomes of a small subset of clones (representing a variety of mutation rates) were sequenced. The x-axis shows the number of single-nucleotide polymorphisms (SNPs) detected in each sequenced clone. The y-axis shows the change in mutation rates of each sequenced clone relative to the mutator ancestor at generation 0. The colors of the points represent the populations that the clones were isolated from.

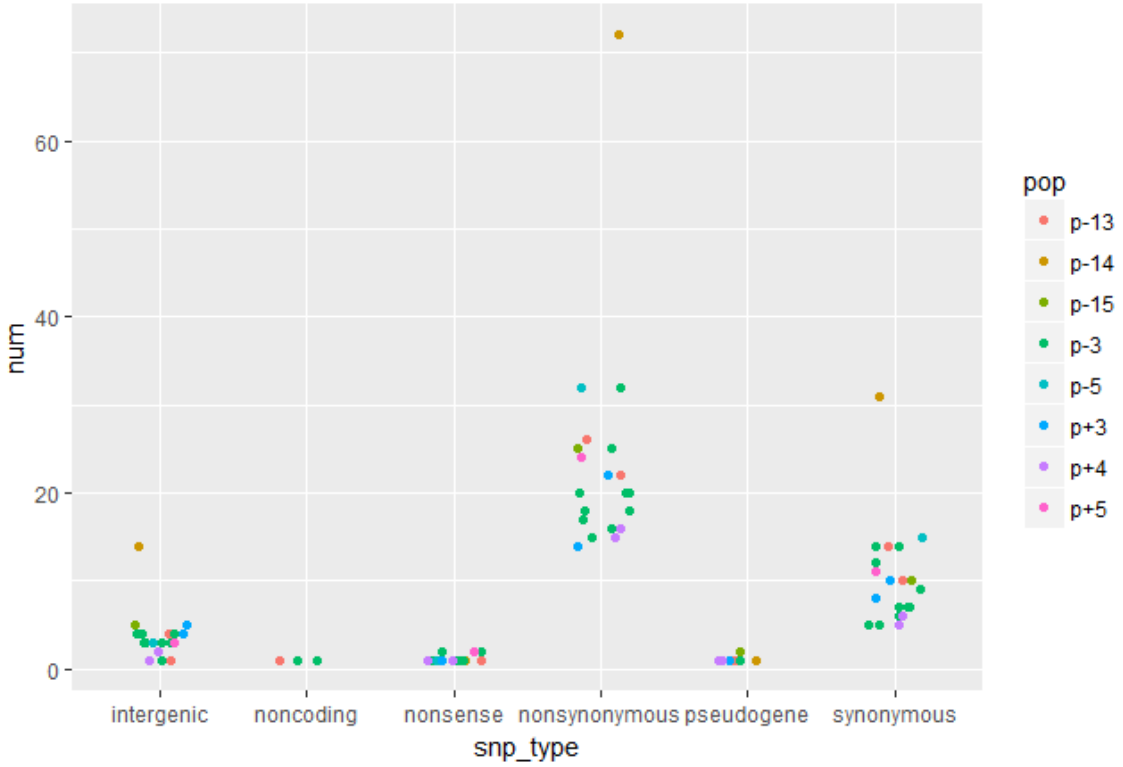


Figure 8: Number of single-nucleotide polymorphisms (SNPs) per clone, classified by type of mutation. SNPs of a variety of types were found across the sequenced clones (harvested after 900 generations of propagation under soft selection). On the y-axis, "num" is the number of SNPs of each mutational class. The types of SNPs are listed along the x-axis.

Sanger sequencing of suspected (anti)mutator loci

In the Ara-13 population, the *recC* (L286P) and *topA* (T613A) mutations were present in clone 5 (as discovered through WGS), which had the most strongly lowered mutation rate in the population; Sanger sequencing showed that these two mutations were absent in the other four clones of this population, which include a clone with an elevated mutation rate (clone 1), two clones with unchanged mutation rates (clones 3 and 4), and a clone with a moderately lowered mutation rate (clone 2).

In the Ara+3 population, whose five clones all had lowered mutation rates, the *lpxA* (D180G) mutation was present in the two clones whose genomes had been sequenced (clones 1 and 2); Sanger sequencing suggested that the mutation was present in the other three clones as well. Furthermore, the WGS showed that clones 1 and 2 had a mutation in *nadR* (T28A); Sanger sequencing showed that this T28A mutation was also present in two of the three other clones in the Ara+3 population. Although clone 3 did not have the T28A mutation, the Sanger sequencing targeting that locus revealed that clone 3 had a +1 C insertion in a string of cytosines in codons 59-60 of *nadR*. Hence, all five clones had a mutation at some point in the NadR N-terminus.

In the Ara+4 population, where all five clones had elevated mutation rates, a mutation in *smf* was found to be present in the two clones whose genomes had been sequenced; through Sanger sequencing, this mutation was shown to be present in at least two of the remaining three clones. (Clone 4's Sanger sequencing chromatogram was ambiguous; thus, the PCR will be optimized and re-run.)

In the Ara+5 population, the *nrdE* (Y16C) and *radA* (1 bp deletion) mutation that were present in clone 2 (as shown by WGS) were absent in the other four clones of the population (as shown by Sanger sequencing). Clone 2 is the only clone with a lowered mutation rate in the Ara+5 population; all other Ara+5 clones have unchanged mutation rates.

Discussion

Mutation rate diversity

Across the evolved clones in this study, we see evidence of mutation rate instability – both upward and downward instability. Among the five clones randomly isolated from the evolved Ara+3 population, all have significantly lowered mutation rates. Among the five clones randomly isolated from the evolved Ara+4 population, all have significantly elevated mutation rates. This raises the question of whether these two populations are fixed for (an) antimutator(s) and (a) mutator(s), respectively. Whole-population genomic sequencing is underway for all 30 evolved populations and may help determine whether any putative mutation rate modifiers are fixed in these two populations. The genomes of two clones from each of those two populations were already sequenced (as part of the batch of 20 evolved clones that were sequenced). Additionally, we are in the process of sequencing the other 3 clones from each of the two putative fixation populations. The population libraries are being prepared with the PCR-free Illumina TruSeq kits; the use of PCR would risk causing non-random apparent changes in the intra-population SNP distribution.

In the Ara+3 population (where all five randomly isolated clones were found to have decreased mutation rates), there were no mutations in *known* mutation rate modifier loci in either of the two sequenced clones. However, there were several non-conservative mutations in common between the two clones in loci coding for the following: acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase (D180G), pyruvate dehydrogenase (ubiquinone) (1 bp deletion), galactarate transporter (P369S), porin (E218K), and trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR (T28A). Intriguingly, none of these mutated loci have been explicitly characterized as antimutator in the existing literature.

In the Ara+4 population (where all five randomly isolated clones were found to have mutation rates elevated above the single-mutator ancestor), the sole mutation in any known mutation rate modifier among the two sequenced clones was in *recC* – and this mutation was only present in clone 2 (G467S). There are at least two possible explanations for this observation. One possibility is that the observed hypermutator phenotype is actually caused by different mutations in different individuals. Another possibility is that the responsible mutation has not been previously characterized as a mutator in the existing literature. Between the two sequenced clones, there were several non-conservative mutations in common across loci coding for: OLD family ATP-dependent endonuclease (DUF2813 family protein) (E326G), pyruvate kinase I

(E71K), outer membrane assembly protein AsmA (stop codon at Q314), protein smf (coding mutation at nucleotide 567; deletion of one C), bifunctional (p)ppGpp synthetase II/ guanosine‑3',5'‑bis pyrophosphate 3'‑pyrophosphohydrolase (Y431C, which is in the region homologous to the TGS domain in the RelA (SpoT) protein), and a hypothetical protein (in genomic site 4,308,608; an insertion of one G in the coding region at nucleotide 535). Among these loci, *smf* is particularly interesting, as it may help RecA (a protein involved in DNA repair) load onto single-stranded DNA (Tadesse and Graumann 2007). Sanger sequencing showed that the *smf* mutation is present in at least four out of the five clones in the Ara+4 population (clone 4's Sanger sequencing results were ambiguous, necessitating a second replicate); these sequencing data encourage further exploration into the possibility of *smf* as a potential mutator allele.

Extensive mutation rate polymorphism was revealed by assays of evolved clones within many populations as well as mutation rate divergence across all populations. The mutation rate polymorphism could be interpreted in one (or more) of at least three possible ways. Firstly, in a polymorphic population, the new (anti)mutator allele(s) could be currently in the midst of their trajectory to eventually sweep to fixation. Secondly, there could be balanced polymorphism for mutation rate modifiers in the population, maintained by frequency-dependent selection (Rozen and Lenski 2000); however, that scenario may necessitate either trade-offs between the multiple phenotypes and/or opportunistic exploitation of unoccupied niche space (Kinnersley et al. 2014). Thirdly, consideration of the multiple testing problem is important – due to the large number of

fluctuation assays performed (with change in mutation rate determined by lack of overlap between the 95% confidence intervals of ancestor and evolved clone), there is some chance that several measurements are misleading (Rice 1989; Benjamini and Hochberg 1995). Note, however, that the fluctuation assays were repeated for clones 1 in Ara-1 through Ara-7; the 95% confidence intervals (CIs) for the repeat measurements overlapped with those of the original measurements (refer to Figure 4c). These CIs are slightly more conservative than other CIs cited in the preexisting literature on fluctuation assays (Zheng 2015), so it is in fact possible that some of these evolved clones had subtle but genuine *changes* in mutation rate that were not detected in our own method of fluctuation analysis. Hypothetically, one could test statistical significance of mutation rate differences by repeating each fluctuation assay several times, and these additional measurements could account not only for the multiple testing problem, but also for the rate of experimental error and slight environmental variations -- however, this method is prohibitively time-consuming.

It is worth noting that even outside the field of mutation rate evolution, the effects of hard and soft selection have been compared in experimental settings. In a recent empirical study, polymorphic *E. coli* populations in heterogeneous habitats were exposed to hard and soft selection; soft selection preserved the polymorphism and retarded the process of fixation (Gallet et al. 2017, preprint).

The *mutS*$^{-}$ mutator allele does not have any known direct effect on fitness (Shaver et al. 2002). Considering that direct fitness benefits to mutator alleles have very rarely been

described in the existing literature, we hypothesize that the increased mutation rates in the current study are most likely the result of mutator hitchhiking – which has been well documented in several different evolutionary contexts (Sniegowski et al. 1997, Gentile et al. 2011). Note that in Lenski's LTEE, three out of 12 populations showed evidence of mutator hitchhiking as of 10,000 generations (Sniegowski et al. 1997).

We were surprised to see so many lowered mutation rates; however, below, I suggest a reasonable explanation for them. But firstly, let us offer evidence *against* one hypothesis for the lowered mutation rates: pressure to reduce mutational load. The 30 populations have been evolving for only ~900 generations, so it is likely that there are many beneficial mutations still available to them. Thus, reduction of deleterious mutational load is unlikely to explain the decreased mutation rates. A more plausible explanation is that the antimutator alleles confer pleiotropic direct fitness benefits that allow them to spread in the population. In this case, the polymorphism and divergence of mutation rates across and within populations could be explained by the potential for fitness gains to be facilitated through mutator hitchhiking *or* through pleiotropic benefits of antimutator alleles. It should be noted that genetic drift is not a sufficient explanation for the mutation rate instability observed in these populations, due to factors such as the effective population size and the high presence of ongoing selection. The effective population size (N_e) in this experiment is 30 million, as it is the product of the number of cells transferred to fresh media each day and the number of generations it takes for this minimum population size to grow to the maximum daily population size:

$$N_e = (5 \times 10^6) \times (\log_2 100) = 3 \times 10^7$$

This is identical (by design) to the effective population size of the Lenski's LTEE, where it has already been argued that the power of adaptation exceeds the potential for drift (Lenski 2004); millions of generations would need to elapse in order for a mutator or antimutator to progress from low to high frequency solely through the power of drift (Crow and Kimura 1970).

Relative fitness gains over 900 generations; is there a relationship between mutation rate and relative fitness?

The relative fitness data were consistent with the original hypothesis of increased fitness. As a point of reference, Lenski's experimental *E. coli* lines had an average selection coefficient of 0.20235 after 1000 generations (Lenski and Travisano 1994). Thus far, we have not detected any relationship between mutation rate and relative fitness (at the population level or at the clone level). Bear in mind that the fluctuation test measures the mutation rate of one clone; the fluctuation test cannot measure the mutation rate of a population. However, we performed fitness competitions on both the evolved clones *and* evolved populations (always relative to the ancestor at generation 0). Refer to Figure 5 for population fitness data; refer to Figure 6 for clone fitness data.

As mentioned in the Results section, we did not detect a relationship between a clone's selection coefficient and the number of new mutations in the sequenced clones. However, note that, for the clones with increased or decreased mutation rates, it is unknown at what generation these changes actually took place. Hence we cannot *assume* that there will be a strong positive correlation between mutation rate (at 900 generations) and the number

of mutations (at 900 generations). Furthermore, note that the subset of evolved sequenced clones were deliberately selected to represent a variety of mutation rates, and thus these data may not be representative of the whole set of 155 evolved clones.

Detected new mutations

The original *mutS*⁻ mutator allele was retained in all of the evolved sequenced clones. Therefore, the changes in mutation rate must have occurred through mutations in other loci. This raises the question of what mutations are causing these changes in mutation rate. What follows is a brief catalog of the most promising candidates for mutation rate modifiers among the mutations that were uncovered after sequencing the genomes of 20 of the evolved clones.

miaA, a tRNA modification gene, can confer mutator status when mutated (Connolly and Winkler 1989; Michaels, Cruz, and Miller 1990). The wild-type version of *miaA* codes for tRNA Δ^2-isopentenylpyrophosphate (IPP) transferase. Mutations in *miaA* may have a variety of effects, including decreasing the stability of anticodon interactions (Vacher et al. 1984). The *miaA* locus is located near the *mutL* locus, near 95 min (Michaels, Cruz, and Miller 1990). Experimental evidence suggests that *miaA* and *mutL* form a complex operon with a weak internal promoter (Connolly and Winkler 1989) that serves diverse cellular functions (Tsui et al. 1996). Hence, it is important to keep in mind the particular characteristics of the *mutS* and *mutL* mutator alleles used in this dissertation work despite the similar change in mutation rate (~100-fold) caused by the two ancestral alleles, *mutL* and *mutS*; had the *miaA* mutant allele landed on the *mutL* background instead, we cannot

be sure that the associated phenotype (if any indeed exists) would have been the same. A non-conservative mutation in *miaA* (E151G) was found in PS3888 (Ara-3, clone 4), which had one of the highest elevations in mutation rate among all 155 evolved clones. Among the list of candidate (anti)mutator loci, *miaA* was the only one mutated in PS3888.

recC is a gene that normally codes for the alpha subunit of the recBC enzyme (Umeno et al. 1983), which is a 5'-3' helicase that binds to double-stranded DNA breaks to prepare them for repair. Multiple nonsynonymous mutations in *recC* were found in several clones from the Ara-3 population, albeit at different residues across different mutation rates: N354S (clone 1 with an unchanged mutation rate), D646E (clone 5 and clone 10, both with lowered mutation rates), and D324G (clone 7, with an elevated mutation rate). A nonsynonymous (and non-conservative) *recC* mutation was also found in the Ara-13 population in clone 5 (L286P), which had a strongly lowered mutation rate. Interestingly, Sanger sequencing showed that this *recC* mutation was absent in the other Ara-13 clones (including the clone with a more moderately lowered mutation rate).

The gene *topA* codes for DNA toposisomerase I, which inhibits hypernegative supercoiling of DNA (Tan et al. 2015). Multiple nonsynonymous *topA* mutations were found in the sequenced clones; similar to *recC*, these mutations were located at different residues across a variety of mutation rates. In the Ara-3 population, sequencing uncovered the *topA* mutations T110A (in clone 7 with an elevated mutation rate), I106T (in clone 1 with an unchanged mutation rate), and A367V (in clones 6 and 8 with an

unchanged mutation rate). In clone 5 (strongly lowered mutation rate) from the Ara-13 population, a *topA* mutation was found (T613A); Sanger sequencing showed that the mutation was absent in the other four clones of this population, as was the case for the *recC* mutation discussed in the preceding paragraph.

A few genes, characterized in the previous literature as potential mutation rate modifiers, possessed non-conservative mutations at different loci across different clones with different mutation rates. These genes include *recC* and *topA*. In each individual case, it is possible that these mutations are not the causative alleles for the changes in mutation rate; however, it is also possible that these represent cases of phenotypic heterogeneity, where different phenotypes are caused by different mutations in the same gene.

After examining the WGS data, we must remain open to the possibility that some of the mutation rate changes were caused by more than one mutation rate modifier (even within the same population) and/or that some of these mutation rate modifiers have not been well-characterized in the existing literature. For example, in the Ara+4 population, there are no mutations in any *known* mutator loci (from the candidate list) in common between the two sequenced clones. In a population with polymorphism for mutation rate, such as Ara-3, clones of similar mutation rate classification (higher, lower, unchanged) do not share any non-conservative mutations in *known* mutation-rate-affecting loci with each other (refer to Table 1). It should be noted that many of the elevated mutation rates were not *profound* changes from the ancestral single-mutator mutation rate (refer to Figure 4). In the existing literature, most of the well-characterized mutator loci have strong effects

on mutation rate: as a point of comparison, before 40,000 generations of evolution, six of Lenski's 12 LTEE populations had acquired mutations in known mutator loci *mutS, mutT,* or *mutL* (Sniegowski et al. 1997; Barrick et al. 2009; Wielgoss et al. 2013). Thus, it is possible that many of the elevated mutation rates observed in these evolved clones may be caused by less well-characterized mutator alleles, each allele of minor strength by itself (Philip Gerrish, personal communication). One advantage to studying as many as 30 evolving populations is that the experiment serves as a "detector array" for changes in mutation rate and maximizes the possibility of identifying subtle, lesser-known mutation rate modifiers. Previous simulations have suggested that mutators of moderate strength (~100-fold) are the most likely to hitchhike (Tenaillon et al. 1999); however, that work was performed with wild-type starting populations in mind, and therefore its applicability to mutator starting populations is subject to debate.

Implications and Applications of Research

This study is a 'snapshot' of 30 asexual populations at a critical time in their evolutionary trajectory. Sufficient time has passed that the populations have likely undergone extensive adaptation to their new environment – as noted previously, in Lenski's LTEE, the fitness of the Ara-1 population relative to ancestor was 1.23 at 900 generations (Lenski and Travisano 1994). At the same time, it is unlikely that the potential for acquiring beneficial mutations has yet been exhausted (Wielgoss et al. 2013).

In certain aforementioned clones, we observed a possible discrepancy between the mutation rate at 900 generations and the number of genomic mutations accumulated at

the point of 900 generations. This is not a biological contradiction: it could be explained by the possibility of a mutator strain acquiring a recent antimutator allele. Whatever the ultimate explanation, this finding will hopefully be relevant to other researchers who measure mutation rates, be it in an experimental evolution laboratory or in a clinical setting. More specifically: despite the contrast between the long history of the laborious fluctuation assay versus the technological advances and popularity of sequencing, WGS cannot fully replace the role of fluctuation assays in measuring mutation rate, because WGS cannot measure the *current* mutation rate of an individual. The mutation rate leaves its imprints on the genome -- but even though the mutation rate can change genotypes, the mutation rate itself can still be seen as a phenotype, and the fluctuation test is a phenotypic assay. There is a relative paucity of literature on the statistics of fluctuation assays; it remains to be seen whether future additions to this literature will be able to improve the sensitivity of fluctuation analysis.

Advances in the study of mutation rate evolution may have biomedical applications. Long-term evolution experiments offer an opportunity to consider the emergence of antibiotic resistance, which has been a pervasive medical problem since at least the 1950s with the arrival of penicillin resistance, and it continues to escalate in intensity to crisis-level proportions in recent decades (Ventola 2015). The evolution of antibiotic resistance is not caused solely by lethal selection occurring at high doses of antibiotics; recent studies have begun to appreciate the fact that extremely low, sublethal doses of antibiotics can also select for antibiotic resistance (Baquero et al. 1997; Hughes and Andersson 2012). It has been suggested that sublethal selection is likely to enrich for

mutators that will have increased probabilities of becoming multidrug-resistant and therefore surviving future encounters with lethal selection (Sandegren 2014). This scenario may be relevant to situations such as unintentional human consumption of trace amounts of antibiotics in food originating from livestock ingesting antibiotic-treated feed or exposure to runoff from animal facilities (Sandegren 2014).

Studies of asexual mutator replicators are also relevant to other areas of biomedicine. Within the human immune system, somatic hypermutation of immune cells aids adaptation in dealing with invading pathogens, by diversifying the V (variable) region of antibodies (Li et al. 2004). It has been observed that mice with genetic defects in mismatch repair have a lower rate of somatic hypermutation (Martin and Scharff 2002). If we view human tumor cells as asexual replicators, investigations of soft selection in asexual bacteria may also be relevant to the evolution of cancer: nutrient competition in the tumor microenvironment can propel the progression of cancer (Chang et al. 2015). The parallel between cancer research and the work described here is strengthened by the observation that cancer cells may be associated with mutator phenotypes, particularly those caused by defective mismatch repair – a system that has many homologous features between human and bacteria (Karran 1995).

Considering that early life might have lacked the same DNA repair and recombination mechanisms present in many modern life forms (Bernstein, Byers, and Michod 1980), high mutation rates may have been prevalent in early life. The environmental context of early Earth should be taken into account: the ‘hot-start hypothesis’ suggests that life on

Earth began in a hot environment, and then was forced to adapt to decreased temperatures as the planet cooled (Nguyen et al. 2017); experimental evolution results suggest that high mutation rates are more advantageous in a continually changing environment, as they speed the acquisition of beneficial mutations (Wielgoss et al. 2013). Additionally, increased mutation rates may be associated with species diversification in organisms as varied as plants (Duchene and Bromham 2013; Bromham et al. 2014), birds (Lanfear et al. 2010), and bacteria (Sawabe et al. 2009); they have also been associated with diversification in viruses (Ribeiro et al. 2012). At least in the case of the *Vibrio splendidus* bacterial clade, there is evidence that mutation had a stronger role than recombination in driving the diversification of species millions of years ago (Sawabe et al. 2009).

Future Directions

Confirmation of the presumed direct fitness effects of the new antimutator alleles could be acquired through allelic replacement assays, where a suspected (anti)mutator mutation would be moved to an ancestral single-mutator background from generation 0; its mutation rate and fitness would then be measured relative to the un-manipulated ancestor. If the mutator hitchhiking hypothesis were correct for a given candidate mutator mutation, then the fluctuation tests would reveal that the suspected mutator mutation increases the mutation rate, and the fitness competitions would reveal that the mutation does not increase the fitness directly (relative to the ancestor at generation zero). If the pleiotropic direct fitness benefit hypothesis were correct for a given candidate antimutator mutation, then we would expect that the allele causes the mutation rate to

decrease, yet it directly increases the fitness of its bearer (relative to the ancestor at generation zero). In our laboratory, a preliminary step has been taken toward this goal with the deletion mutation found in the DNA repair gene *radA*. This mutation was uncovered in one clone of the population Ara+5, which had a mutation rate one order of magnitude lower than the other four isolated Ara+5 clones (which had mutation rates unchanged relative to ancestor). This clone was a prime candidate for sequencing because we wanted to determine its potential antimutator alleles. In light of the correlation between a lowered mutation rate and the presence of the *radA* and *nrdE* (ribonucleoside-diphosphate reductase 2 subunit alpha) mutations in the Ara+5 population, allele replacement assays of these two mutations are a potential avenue for detecting new antimutator alleles.

Further explorations of these populations could involve locating the time that the mutation rate modifiers arose, and observe their frequencies change over time; this investigation would entail whole-population genome sequencing across timepoints revived from frozen stock. This observation could even be continued into the future trajectory of these populations, by continuing to propagate the populations and sequencing them at regular timepoints. In this manner, one could observe mutation rate modifiers rising to fixation – or, alternately, being purged from the population, or possibly even remaining polymorphic within the population. As mentioned earlier in this chapter, the pervasive mutation rate polymorphism in these populations at 900 generations raises the question of whether the polymorphism is stable or transient. Also, although our existing research on the populations at 900 generations suggests an absence

of any relationship between relative fitness and mutation rate, whether this lack of relationship is temporally dependent or not could be examined in populations at earlier and later generations.

The gene *mutS* is but one component in the mismatch repair pathway; it may be worth considering whether similar outcomes would have been achieved if the mutator ancestral strain had instead been defective in a gene with a different molecular function in the pathway (*mutL, mutH*). Also, the possibility that a mutator's specific mutation rate (low, moderate, high) might affect the likelihood of adaptation has been explored in theoretical work (Tenaillon et al. 1999), and is recently being explored in a laboratory setting (Sprouffske et al. 2017, unpublished submitted work). Based on the aforementioned existing literature, it seems plausible that a much higher starting mutation rate would have led to limited adaptation in our experiment, because there may be a mutation rate where the advantage of higher rates of beneficial mutations is overshadowed by the disadvantage of higher rates of deleterious mutations (Sprouffske et al. 2017, unpublished submitted work) – but it may be worth investigating whether this hypothesis stands up to empirical testing.

Chapter IV: Conclusions and Future Directions

Overview

This dissertation offers evidence – at least in the context of asexual populations – that challenges the tacit but common supposition of mutation rates as stable and fixed. From a population's genesis to adaptation to extinction, perhaps it is the case that mutation rate instability is more widespread than previously assumed. In this chapter, I compare the outcomes of the hard selection and soft selection experiments. I discuss how the findings fit into the theme of chronic upward instability of the mutation rate, which was a prime theoretical motivator for the initiation of my empirical investigations. Yet I also discuss how these findings add a layer of nuance to that theme of chronic upward instability, in that "downward instability" may be as intriguing a field of study as "upward instability". I discuss the challenges behind measuring mutation rates with fluctuation assays, which was a method critical to both of my investigations. I explore the implications of these results in the settings of biomedicine, evolutionary thought, and astrobiology. I also suggest some directions for future investigations that may follow naturally from the findings described in this dissertation.

Many an introductory biology lecture has begun with a statement similar or even identical to the following: 'mutations are random, and selection acts upon them'. This statement, though not incorrect, raises another question: does the *type* of selection affect the fate of those mutations? The work described in this dissertation (in conjunction with the pre-existing literature) may suggest 'yes'. Furthermore, the type of selection may be

able to change the rate at which the mutations emerge – without changing the content or quality of those mutations. In the work described in Chapters II (lethal selection experiment) and III (soft selection experiment), the mutation rates of the mutator ancestors were similar, but these ancestral mutation rates became destabilized over the course of adaptation, and the patterns of mutation rate evolution of the evolved clones were notably different across the two experiments.

Comparison of the populations studied under hard and soft selection

In the lethal antibiotic regime described in chapter II, the population was confronted with selective agents that were fatal to all individuals except the ones that bore particular beneficial mutations (namely, mutations conferring resistance to certain antibiotics). By the end of the antibiotic regime, all survivors had adapted to *multiple* new environmental variables (as represented by the different antibiotics); the fact that a high mutation rate confers a selective advantage in a rapidly changing environment is supported by preexisting experimental literature (de Visser 2002). The strong increase in mutation rates (caused by apparent fixation of an additional mutator allele) was a logical outcome in the survivors of the series of three different lethal antibiotics.

In the soft selection environment described in chapter III, an individual's acquisition of beneficial mutations could obviously confer a competitive advantage over other members of the same population. However, in the soft selection environment, an individual's survival was less dependent on the rapid acquisition on specific beneficial mutations (e.g. antibiotic resistance mutations as in the hard selection experiments). This may have

allowed for the spread of antimutator alleles conferring pleiotropic direct fitness advantages. Even if an antimutator allele can confer a direct fitness advantage, it is less likely that antimutator alleles would have survived in a lethal environment, because of the risk that their pleiotropic direct advantage may have been outweighed by the reduced probability of acquiring specific beneficial mutations; it is known that higher mutation rates tend to be favored under hard selection (Mao et al. 1997). Although the soft selection experiments were longer-term (~900 generations) than the hard selection experiments, it is unlikely that enough time passed for the deleterious mutational load to accumulate to the point where there was a selective advantage to decreasing the mutation rate, as has been described in other experimental evolution literature (Wielgoss et al. 2013). In the soft selection populations, extensive polymorphism was detected, with possible fixation of a mutation rate modifier in two populations: an antimutator in Ara+3 and a mutator in Ara+4.

Despite the different patterns of mutation rate evolution observed in the two experiments, there are some parallels. In all sequenced genomes of evolved clones (20 from the soft selection and two from the hard selection populations), we did not see any reversion of the ancestral mutator allele. In other words, in the soft selection clones, the original *mutS*$^{-}$ allele was retained (except for one clone with an unchanged mutation rate that had a synonymous mutation in the *mutS*$^{-}$ gene, but located outside of the mutator locus); in the hard selection clones, the original *mutL*$^{-}$ allele was retained. The subsequent changes in mutation rate appear to have been caused by mutations in loci outside the ancestral mutator locus. Our results do not *preclude* the possibility of back mutations in other

hypothetical mutator populations. The possibility of back mutations occurring in mutator populations has been explored in some existing theoretical literature, where it is posited as being the only option (besides recombination and compensatory mutations) that can stall the progressive fitness loss in a population that is caused by Muller's ratchet (Söderberg and Berg 2011).

Discussion

Chronic upward instability of the mutation rate

The experiments described in this dissertation were motivated by theoretical models that explored the possibility of repeated mutator hitchhiking in an asexual population, resulting in extremely high mutation rates (André and Godelle 2006; Gerrish et al. 2007). In light of the obvious fact that deleterious mutations are more common than beneficial mutations, it has been proposed that a population can be driven to extinction by a deleterious mutational load accumulated by an intolerably high mutation rate (Eigen 2002; Gerrish et al. 2013). Although this extinction outcome has not been described in *E. coli*, there is experimental evidence suggesting that viral populations can be rendered inviable via an extremely high mutation rate (Crotty et al. 2002). RNA viruses have some of the highest known mutation rates (Belshaw et al. 2007) and because of this, they are thought to exist on the brink of inviability – indeed, an increase in mutation rate less than one order of magnitude may send an RNA virus population into extinction (Crotty et al. 2001). Studies in yeast have provided evidence that there is a mutation rate in diploid *Saccharomyces cervisiae* after which the population can succumb to replication error-

induced extinction, and that a decline in fitness begins at mutation rates far below the lethal mutation rate (Herr et al. 2014).

In the hard selection experiment described in Chapter II, a spontaneously originated mutator allele (*mutT*$^{-}$) hitchhiked to fixation in a preexisting mutator (*mutL*$^{-}$) population that underwent repeated rounds of lethal selection; this was observed across two independent experimental replicates. In the soft selection experiment described in Chapter III, further increases in mutation rates (including to possible fixation in one population) were observed in several of the 155 clones randomly sampled after 900 generations of propagation; however, decreases in mutation rates were actually more common than increases in mutation rate. Although we did not see (and did not expect to see, in part due to the relatively brief length of time) extinction events in our experiments, our experiments do offer evidence for the general phenomenon of mutation rate instability in evolving asexual mutator populations. Yet it is interesting to note that this mutation rate instability can occur in a *downward* direction, in addition to upward, for the reasons discussed in Chapter III.

Detection of polymorphism

When performing fluctuation assays throughout this thesis, changes in mutation rate were calculated by checking for overlap between the confidence intervals (CIs) of evolved and ancestral strains that were set at a conservative 95%; in a methods paper on analyzing fluctuation assay data, the use of 84% CIs was suggested (Zheng 2015). With that in mind, it is possible that the amount of mutation rate polymorphism in these experiments

may exceed our estimates, possibly even despite the multiple testing problem mentioned in Chapter III. Additionally, even in a liberal interpretation of confidence in a fluctuation test, it is formally possible that a new mutation could boost the mutation rate by a modest amount (say, 70%) that is not detectable with current methods of mutation rate measurement, yet it nevertheless may change its bearer's opportunities for adaptation, if the population is of sufficient size (Jiang et al. 2010). The existence of the phenomenon of stress-induced mutagenesis could hypothetically complicate measurements further, as its occurrence may be able to transiently affect mutation rate by merely a few-fold (or by as much as a thousand-fold or more) (Galhardo et al. 2007).

The properties of specific mutation rate modifiers

One potentially interesting area of investigation is confirming whether the suspected mutation rate modifiers do indeed have an effect on mutation rate, and also whether they have a direct (or indirect) effect on fitness. This can be explored through allele replacement assays, where a suspected mutation rate modifier is placed on the ancestral background; its mutation rate can then be measured via fluctuation assays, and its fitness can be measured relative to the ancestor via growth competition assays. Although most mutation rate modifiers described in the literature are assumed to have no direct effect on fitness (Raynes and Sniegowski 2014) (but see Torres-Barceló et al. 2013), this assumption should, in an *ideal* world, be tested for each suspected mutator. If a widespread mutator allele is shown to have a direct fitness benefit, then it is possible that this direct effect is more responsible for its rise to fixation than mutator hitchhiking. It should be noted that the *mutT*$^{-}$ mutator alleles detected in the hard selection experiments

are not expected to affect fitness directly (Wielgoss et al. 2013). Rather, they probably confer an indirect benefit in that they generate new beneficial mutations that confer antibiotic resistance and allow their bearers to survive in a rapidly changing and potentially lethal environment; in short, they most likely rose to fixation by mutator hitchhiking.

Molecular characteristics of mutators

What are the trends, if any, in the molecular characteristics of spontaneously arising mutator alleles? -- both in their origin (what loci they emerge in) and in the mutational spectra that they inflict upon the genome? Are there 'hotspots' in the genome where mutators are more likely to emerge? Working with numerous evolving populations affords us an opportunity to observe potential convergent or divergent evolution. We should not assume that the studies in this thesis will certainly display rampant convergent evolution; for example, although a long-term multi-patient clinical study showed that 17 bacterial genes had multiple nonsynonymous mutations suggestive of parallel adaptive evolution, this pattern had been observed across 112 pathogenic bacterial isolates harvested from 14 patients over 16 years (Lieberman et al. 2011), which is a larger time scale and sample size than the work described in this thesis.

Nevertheless, the work described in this thesis did show certain parallels to preexisting literature. Namely, in the hard selection experiment in Chapter II, we observed that the same mutation (a +1 C frameshift mutation in a repeat C region of *mutT*) arose to probable fixation in survivors of the antibiotic selection series across two independent

replicates. This same stretch of repeat DNA was mutated in a long-term evolution experiment under soft selection (Barrick et al. 2009). Additionally, previous experimental work in two populations with spontaneously arising mutators showed that the mutator mutations independently arose in the same repeat region of *mutL* (Shaver et al. 2003). In both prokaryotes and eukaryotes, DNA polymerase slippage during replication is more likely to occur in repeat regions (Viguera et al. 2001). Though the trend for higher mutability at repeat regions is well-supported (Levinson and Gutman 1987), it is controversial whether the presence of repeat regions can be considered "adaptive" in general (Shaver et al. 2003). Experimental studies have identified a high number of short close repeats in *E. coli* DNA repair and stress response genes, and it was suggested that these repeats may be able to produce a variety of phenotypes if there is slippage during the synthesis of genetic material (Rocha et al. 2002). However, it is uncertain whether the existence of these repeat regions is necessarily the result of selection for increased mutability at their respective loci (Shaver et al. 2003; but see Field et al. 1999 for more conclusive examples of adaptive hypermutability in repeat-heavy 'contingency loci', which are specific genomic regions in pathogenic microbes).

Regardless of the degree to which this *mutT*$^{-}$ region has adaptive significance, it is interesting to consider whether the hard selection experiment would have yielded similar results had this genomic region been less hypermutable. This question could be answered by replacing the repeat C region with alternate codons. The codons in which the string of Cs resides are as follows: ACC CCC CAA (National Center for Biotechnology Information), which codes for threonine, proline, and glutamine, respectively. In this

hypothetical experiment, the Thr codon could be replaced with ACT, ACA, or ACG; the Pro codon could be replaced with CCT, CCA, or CCG – thus, there would be no more than 2 Cs adjacent in this region, but the MutT protein would remain the same. The hard selection experiment could then be repeated on this engineered strain.

The size of a gene (and, by extension, the size of the 'target' for potential mutation) may be worth considering when attempting to gauge the likelihood of certain new mutators emerging -- although our own work cannot claim to offer immediate support to this connection, in light of the fact that the *mutT* gene is only 390 basepairs long (National Center for Biotechnology Information). The size of a mutator gene is not the same as the number of sites that could cause a mutator phenotype if mutated -- for potentially illustrative examples of mutations in mutator loci causing no apparent effect on the mutation rate, refer to Chapter III, where, in some evolved clones, several mutations were found in loci described in the existing literature as mutators, yet we were unable to detect significant changes in mutation rate relative to ancestor. Of course, probability of emergence does not equal probability of fixation: other factors, such as the potency of the mutator's effect on mutation rate, may affect whether the mutator allele spreads in the population. For example, a strong mutator may give an individual under lethal selection an advantage in acquiring the beneficial mutations requisite for survival; however, there has been some theoretical work suggesting that intermediate-effect mutators may be more likely to hitchhike (Tenaillon et al. 1999).

Because both the hard selection experiment and the soft selection experiment began with preexisting mutator strains, it is interesting to consider to what degree the observed changes in mutation rate (as a result of the acquisition of additional mutation rate modifiers) are influenced by the specific identity of the original ancestral mutator allele (*mutL*$^{-}$ in Chapter II and *mutS*$^{-}$ in Chapter III). Though it was not our intent to address this question in this dissertation, this question is worth contemplating for future studies on repeated mutator hitchhiking events. Although phrases such as "double-mutator" are occasionally seen in the mutation rate evolution literature, including our own, it must be remembered that the number of mutation rate modifiers is a discrete variable, yet the mutation rate is a continuous variable. Thus, the knowledge that two mutator alleles exist in a given individual is not always immediately informative at the practical level; the mutation rates in such individuals may not necessarily be the sum of their individual single-mutator mutation rates, and they may even have their own unique mutational spectra that are not the sum of their single mutator allele. This has been noted in strains mutant for both *mutT*$^{-}$ and *mutY*$^{-}$: each, in the absence of the other allele, is a mutator allele (Fowler et al. 2003). When an evolving *mutT*$^{-}$ mutator strain spontaneously acquired a *mutY*$^{-}$ mutation, the new mutation rate was lower than the mutation rate of the *mutT*$^{-}$ *mutY*$^{+}$ strain, because the mutational spectrum of *mutY*$^{-}$ counteracts the mutational spectrum of *mutT*$^{-}$ (Wielgoss et al. 2013). However, if two hypothetical mutators arising on the same genome had similar spectra, it is worth considering the possibility that they could cause imbalances in the GC content of the genome. It has been proposed that the methyl-directed mismatch repair system serves to stymie gradual escalation of the GC content of the genome (Lee, Popodi, Tang, and Foster 2012). Since higher GC content is

correlated with the presence of higher environmental temperature (Smarda et al. 2014), the type of environment could perhaps, in some cases, affect the viability of a population with a very high mutation rate that has been evolving for many generations.

As mentioned earlier (particularly in light of the genomic data from the soft selection experiment in Chapter III), the mere existence of a mutation in a mutator locus, even if it is a non-conservative mutation, does not automatically mean that the mutation rate will be affected. There are several possible reasons for this non-effect on a mutation rate despite a mutation in a mutator locus. Firstly, it is obvious that not all residues of a protein are equally important to the protein's function. Thus, the location of a mutation in a mutator locus is important for understanding what effect (if any) it will have on the mutation rate. Secondly, if a new mutation affects a process downstream of the same pathway disrupted by the first mutator allele, the mutation rate may not necessarily change. Thirdly, it is possible that the new mutation does indeed affect the mutation rate, but that there was yet *another* new mutation in a different mutation rate modifier locus that counteracts the effects of the other new mutation. This raises the question of the pervasiveness of epistatic interactions in our experiments, as discussed in the next section of this chapter.

The potential for epistasis

It is interesting to contemplate the extent to which new mutations observed in our evolving mutator populations have epistatic interactions with other alleles in their respective genomes. This is of particular relevance in the hard selection experiment

(Chapter II): although *mutT* impairment is expected to increase the genomic mutation rate between 100- and 1000-fold in *E. coli* (Maki and Sekiguchi 1992), the mutation rates in the evolved *mutL*$^{-}$ *mutT*$^{-}$ genotypes in chapter II were estimated to be only about 10- to 20-fold higher than that of our ancestral *mutL*$^{-}$ genotype. Preexisting experimental studies have shown that the A.G mispairing mutations caused by the *mutT*$^{-}$ mutator gene are not recognized by the mismatch repair system (even though the same system can sometimes recognize A.G mispairings from other sources) (Schaaper et al. 1989), and that the mutator effect of the double mutator *mutT*$^{-}$ *mutL*$^{-}$ genotype is approximately additive (Bridges 1996). This strengthens the case for epistasis caused by another unknown genetic locus in the survivors of our hard selection experiment, since the mutation rate of the evolved genotype (*mutL*$^{-}$ *mutT*$^{-}$*)* was only about one order of magnitude higher than the *mutL*$^{-}$ ancestral mutation rate; the proteins coded by the two genes operate in different pathways, and there is no reason to believe that the lower-than-expected resultant mutation rate is due to redundancy between the two genes.

To determine what may be partly suppressing the *mutT*$^{-}$ mutator effect in the hard selection strains, an allele replacement assay could be performed, where the *mutT*$^{-}$ mutation is placed on the ancestral *mutL*$^{-}$ background, along with one of the other mutations uncovered in the whole genome sequencing, which may be suspected as being a (conditional) antimutator. Drake's 1993 statement on the improbability of strong antimutators may be worth revisiting in the context of mutator bacteria: this claim may be true for wild-type mutation rates, but there could be many genes that, when mutated, may act as antimutators in mutator strains. It is worth noting that in one experimental study of

bacteria with an ancestral *mutT*$^-$ mutator genotype, their mutation rates declined after 2,300 generations of evolution, though they still remained higher than wild-type levels (Tröbner and Piechocki 1984). The authors showed that the *mutT*$^-$ allele itself had not changed (Tröbner and Piechocki 1984). Although this study took place in a chemostat, involved a different starting genotype than the *mutL*$^-$ starting genotype of our hard selection experiment, and was propagated for a longer number of generations, it is a precedent for the existence of dampened *mutT*$^-$ mutator phenotypes in evolving populations.

Gaining new perspectives along the evolutionary timeline

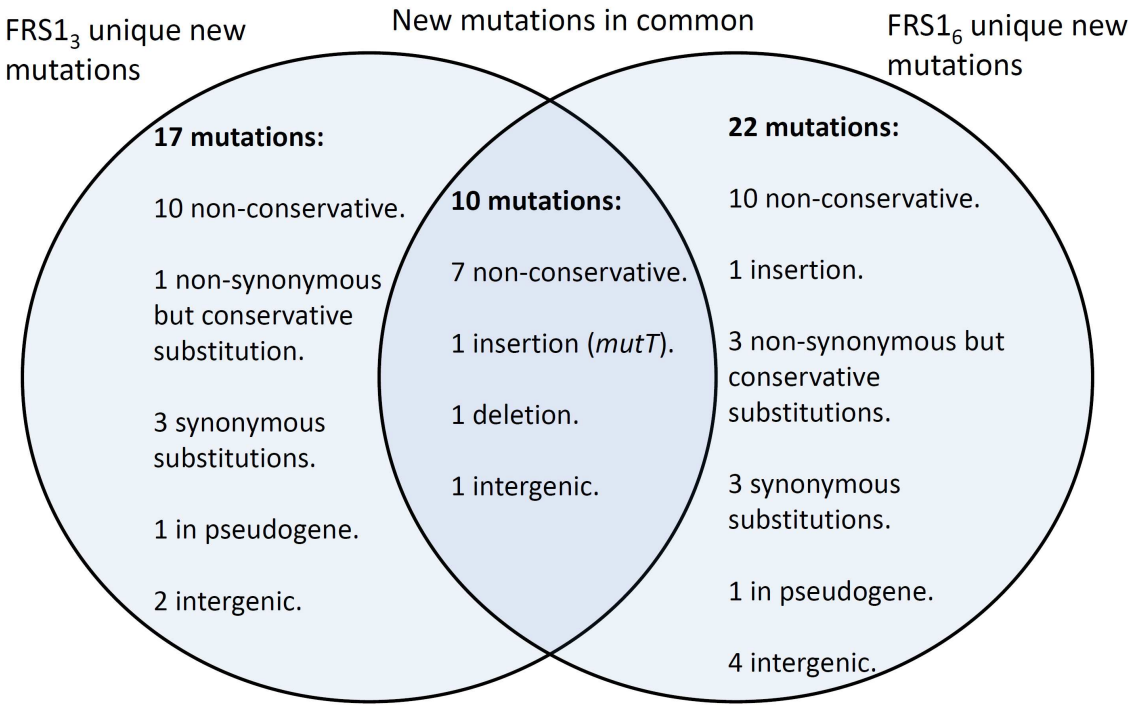
In an evolution experiment, it is important to understand how the observed slice of time fits in with the past and the future. Therefore, it would be illuminating to sequence the frozen timepoints across the course of the soft selection experiment and determine the changes in frequency of candidate mutation rate modifiers over time, especially the time of emergence. (At the present time, whole-genome population sequencing is underway for the 30 evolved soft selection populations to determine the current frequencies.) Furthermore, the timelines for both the hard selection experiment and the soft selection experiment could hypothetically be extended.

In the soft selection experiment, this would be a simple matter of reviving the 30 populations at the latest timepoints and continuing propagation for hundreds or even thousands more generations. Considering the degree of mutation rate polymorphism in many of the current populations, it would be interesting to see whether this mutation rate

polymorphism is stable over time. Another potential avenue of study would be to see if the populations containing clones with elevated mutation rates (described in Chapter III) underwent further mutator hitchhiking events, thereby giving more weight to theoretical predictions of the possibility of repeated mutator hitchhiking events in asexual adapting populations (Gerrish et al. 2007), and/or whether these mutation rates decline over time, either out of the need to reduce deleterious mutational load or because of pleiotropic direct fitness benefits conferred by new antimutator alleles. Extending the timeline of the hard selection experiment would require the challenging task of finding additional selective agents that are not cross-resistant to the three antibiotics used in our selection series; furthermore, they cannot be cross-resistant to nalidixic acid, the selective agent used in the fluctuation tests. If these agents are identified and the experiment extended, however, it would be interesting to see if there is a point in the selective 'gauntlet' after which the entire population became extinct (from accumulated deleterious mutations or from an inability to acquire the necessary resistance mutations), and/or whether additional mutator hitchhiking events would occur. It has been suggested that theoretical models of extinction via mutational degradation tend to depend on the context of species, environment, and timing (Gerrish et al. 2013); with that in mind, it is important to supplement these theoretical models with empirical studies.

APPENDIX

Figure S1.



Venn diagram summarizing the classifications of the new mutations uncovered in the two sequenced evolved clones (FRS1$_3$ and FRS1$_6$; refer to Figure 2a-b) since their divergence from the single-mutator ancestor.

Figure S2.

ATGAAAAAGC TGCAAATTGC GGTAGGTATT ATTCGCAACG AGAACAATGA
AATCTTTATA ACGCGTCGCG CAGCAGATGC GCACATGGCG AATAAACTGG
AGTTTCCCGG CGGTAAAATT GAAATGGGTG AAACGCCGGA ACAGGCGGTG
GTGCGTGAAC TTCAGGAAGA AGTCGGGATT A**CCCCCC**AAC ATTTTTCGCT
ATTTGAAAAA CTGGAATATG AATTCCCGGA CAGGCATATA ACACTGTGGT
TTTGGCTGGT CGAACGCTGG GAAGGGGAGC CGTGGGGTAA AGAAGGGCAA
CCCGGTGAGT GGATGTCGCT GGTCGGTCTT AATGCCGATG ATTTTCCGCC
AGCCAATGAA CCGGTAATTG CGAAGCTTAA ACGTCTGTAG

The nucleotide sequence of the wild-type *mutT* gene. The gene contains a string of six cytosines at positions 182-187 (underlined). After three rounds of antibiotic selection, two clones with further increases in mutation rate were sequenced ($FRS1_3$ and $FRS1_6$, refer to Figure 2a-b and Figure S1); both clones possess a frameshift mutation (+C) within the 182-187 repeat region.

Table S1.

Absolute mutation rate (MR) estimates from all fluctuation assays (Figure 2a-b). Refer to text for details on methodology and interpretation of results. Each yellow row designates the reference ancestor in that category that was used for the ancestral confidence intervals (CIs) in Figure 2a-b. (The reference ancestor was the ancestor with the largest difference between the 95% CIs when expressed as "-fold difference".) Red refers to increased MR; blue refers to decreased MR.

Table S1a: Fosfomycin→rifampicin→streptomycin (FRS) exposure. (Refer to FRS in Figure 2a.)

Strain description	Lab strain #	Fluctuation assay batch ID	*Mutation rate (MR) measurements w/ 95% CIs*				Notes
			Qualitative MR relative to ancestor	Max likelihood MR	Lower 95% CI	Upper 95% CI	
MUTATOR ANCESTOR	PS2533	C	N/A	6.74E-08	4.77E-08	9.51E-08	
MUTATOR ANCESTOR	PS2533	D	N/A	6.42E-08	5.00E-08	8.24E-08	
MUTATOR ANCESTOR	PS2533	F	N/A	9.64E-08	7.36E-08	1.26E-07	
FRS SERIES 1, clone 1	PS4146	D	HIGHER	7.18E-07	4.69E-07	1.10E-06	
FRS SERIES 1, clone 2	PS4147	D	HIGHER	5.29E-07	3.83E-07	7.30E-07	
FRS SERIES 1, clone 3	PS4148	D	HIGHER	6.57E-07	5.07E-07	8.50E-07	
FRS SERIES 1, clone 4	PS4149	D	HIGHER	4.17E-07	3.21E-07	5.42E-07	
FRS SERIES 1, clone 5	PS4150	D	HIGHER	4.19E-07	3.09E-07	5.69E-07	
FRS SERIES 1, clone 6	PS4151	D	HIGHER	3.83E-07	2.73E-07	5.35E-07	
FRS SERIES 1, clone 7	PS4152	C	HIGHER	5.08E-07	3.88E-07	6.65E-07	
FRS SERIES 1, clone 8	PS4153	C	HIGHER	6.19E-07	4.81E-07	7.96E-07	
FRS SERIES 1, clone 9	PS4154	F	HIGHER	7.07E-07	5.31E-	9.43E-	

					07	07	
FRS SERIES 2, clone 1	PS4735	D	HIGHER	1.08E-06	9.07E-07	1.28E-06	From same selection series as PS4705-4709.

MRs of PS4146-51 were measured in fluctuation assay batch K, whose ancestral MR values are not listed here because they were outliers compared to all other measurements of the ancestor's MR. (Even with the batch K ancestral MR values, however, the MRs of these six clones were still higher than ancestor, with non-overlapping CIs.)

The MR of PS4735 was measured in fluctuation assay batch L, whose ancestral MR values are not listed here because a technical error occurred with the population (plain LB) plates of the ancestor, preventing the ancestral MR from being calculated.

Hence, in Figure 2a and Table S1, the ancestral MR from fluctuation assay batch D was used for calculating -fold MR change from ancestor in the aforementioned strains.

Table S1b: Single fosfomycin (F) exposure. (Refer to F in Figure 2a.)

Strain description	Lab strain #	Fluctuation assay batch ID	*Mutation rate (MR) measurements w/ 95% CIs*				Notes
			Qualitative MR relative to ancestor	Max likelihood MR	Lower 95% CI	Upper 95% CI	
MUTATOR ANCESTOR	PS2533	A	N/A	7.23E-08	5.48E-08	9.53E-08	
MUTATOR ANCESTOR	PS2533	E	N/A	7.48E-08	5.62E-08	9.95E-08	
MUTATOR ANCESTOR	PS2533	H	N/A	2.87E-07	2.36E-07	3.49E-07	
MUTATOR ANCESTOR	PS2533	I	N/A	7.50E-08	5.71E-08	9.85E-08	
F SERIES 1, clone 1	PS4678	A	SAME	1.19E-07	8.80E-08	1.60E-07	
F SERIES 1, clone 2	PS4679	A	SAME	1.18E-07	8.77E-08	1.60E-07	
F SERIES 1, clone 3	PS4680	A	SAME	1.16E-07	8.01E-08	1.69E-07	
F SERIES 1, clone 4	PS4681	E	SAME	1.05E-07	7.06E-08	1.56E-07	
F SERIES 1, clone 5	PS4682	E	SAME	6.92E-08	4.72E-08	1.02E-07	
F SERIES 1, clone 8	PS4685	H	LOWER	1.51E-07	1.10E-07	2.06E-07	
F SERIES 1, clone 9	PS4686	H	LOWER	1.57E-07	1.13E-07	2.18E-07	
F SERIES 1, clone 10	PS4687	I	SAME	6.46E-08	4.77E-08	8.76E-08	

Table S1c: Single rifampicin (R) exposure. (Refer to R in Figure 2a.)

Strain description	Lab strain #	Fluctuation assay batch ID	*Mutation rate (MR) measurements w/ 95% CIs*				Notes
			Qualitative MR relative to ancestor	Max likelihood MR	Lower 95% CI	Upper 95% CI	
MUTATOR ANCESTOR	PS2533	E	N/A	7.48E-08	5.62E-08	9.95E-08	
MUTATOR ANCESTOR	PS2533	F	N/A	9.64E-08	7.36E-08	1.26E-07	
R SERIES 1, clone 1	PS4485	E	HIGHER	2.85E-07	2.17E-07	3.73E-07	
R SERIES 1, clone 3	PS4487	E	HIGHER	1.81E-07	1.32E-07	2.48E-07	
R SERIES 1, clone 6	PS4490	F	SAME	6.93E-08	4.56E-08	1.05E-07	
R SERIES 1, clone 7	PS4491	F	SAME	1.44E-07	1.06E-07	1.94E-07	
R SERIES 1, clone 8	PS4492	F	SAME	1.21E-07	8.07E-08	1.82E-07	

Table S1d: Single streptomycin (S) exposure. (Refer to S in Figure 2a.)

Strain description	Lab strain #	Fluctuation assay batch ID	*Mutation rate (MR) measurements w/ 95% CIs*				Notes
			Qualitative MR relative to ancestor	Max likelihood MR	Lower 95% CI	Upper 95% CI	
MUTATOR ANCESTOR	PS2533	D	N/A	6.42E-08	5.00E-08	8.24E-08	
MUTATOR ANCESTOR	PS2533	G	N/A	1.89E-07	1.49E-07	2.40E-07	
S SERIES 1, clone 1	PS4451	D	SAME	1.05E-07	8.24E-08	1.34E-07	From same selection series as PS4533-4542.
S SERIES 1, clone 2	PS4452	D	SAME	4.68E-08	3.10E-08	7.06E-08	From same selection series as PS4533-4542.
S SERIES 1, clone 3	PS4453	D	SAME	6.31E-08	4.45E-08	8.95E-08	From same selection series as PS4533-4542.
S SERIES 1, clone 4	PS4454	D	SAME	4.53E-08	3.25E-08	6.31E-08	From same selection series as PS4533-4542.
S SERIES 1, clone 5	PS4455	D	SAME	3.53E-08	2.19E-08	5.71E-08	From same selection series as PS4533-4542.
S SERIES 1, clone 6	PS4456	G	SAME	1.50E-07	9.80E-08	2.31E-07	From same selection series as PS4533-4542.

Table S1e: Fosfomycin→rifampicin (FR) exposure. (Refer to FR in Figure 2a.)

Strain description	Lab strain #	Fluctuation assay batch ID	*Mutation rate (MR) measurements w/ 95% CIs*				Notes
			Qualitative MR relative to ancestor	Max likelihood MR	Lower 95% CI	Upper 95% CI	
MUTATOR ANCESTOR	PS2533	C	N/A	6.74E-08	4.77E-08	9.51E-08	
MUTATOR ANCESTOR	PS2533	F	N/A	9.64E-08	7.36E-08	1.26E-07	
FR SERIES 1, clone 6	PS4705	C	SAME	8.93E-08	6.74E-08	1.18E-07	From same selection series as PS4735.
FR SERIES 1, clone 7	PS4706	F	HIGHER	1.97E-07	1.49E-07	2.59E-07	From same selection series as PS4735.
FR SERIES 1, clone 8	PS4707	F	SAME	8.27E-08	5.85E-08	1.17E-07	From same selection series as PS4735.
FR SERIES 1, clone 9	PS4708	F	SAME	9.14E-08	6.61E-08	1.26E-07	From same selection series as PS4735.
FR SERIES 1, clone 10	PS4709	F	SAME	9.53E-08	6.68E-08	1.36E-07	From same selection series as PS4735.

Table S1f: Streptomycin→rifampicin→fosfomycin (SRF) exposure. (Refer to SRF in Figure 2a.)

Strain description	Lab strain #	Fluctuation assay batch ID	*Mutation rate (MR) measurements w/ 95% CIs*				Notes
			Qualitative MR relative to ancestor	Max likelihood MR	Lower 95% CI	Upper 95% CI	
MUTATOR ANCESTOR	PS2533	J	N/A	8.45E-08	6.25E-08	1.14E-07	
MUTATOR ANCESTOR	PS2533	B	N/A	7.34E-08	5.54E-08	9.72E-08	
MUTATOR ANCESTOR	PS2533	F	N/A	9.64E-08	7.36E-08	1.26E-07	
MUTATOR ANCESTOR	PS2533	C	N/A	6.74E-08	4.77E-08	9.51E-08	
SRF SERIES 1, clone 1	PS4530	J	SAME	1.47E-07	1.11E-07	1.93E-07	
SRF SERIES 1, clone 2	PS4531	J	SAME	5.22E-08	3.65E-08	7.47E-08	
SRF SERIES 2, clone 1	PS4533	B	SAME	7.73E-08	5.29E-08	1.13E-07	From the same selection series as PS4451-4456.
SRF SERIES 2, clone 2	PS4534	B	SAME	1.03E-07	6.84E-08	1.54E-07	From the same selection series as PS4451-4456.
SRF SERIES 2, clone 3	PS4535	B	LOWER	2.15E-08	1.07E-08	4.33E-08	From the same selection series as PS4451-4456.
SRF SERIES 2, clone 4	PS4536	B	SAME	7.55E-08	5.17E-08	1.10E-07	From the same selection series as PS4451-4456.

SRF SERIES 2, clone 5	PS4537	B	SAME	8.27E-08	6.11E-08	1.12E-07	From the same selection series as PS4451-4456.
SRF SERIES 2, clone 6	PS4538	B	SAME	5.05E-08	3.11E-08	8.19E-08	From the same selection series as PS4451-4456.
SRF SERIES 2, clone 7	PS4539	F	SAME	8.19E-08	5.77E-08	1.16E-07	From the same selection series as PS4451-4456.
SRF SERIES 2, clone 9	PS4541	B	SAME	7.50E-08	5.49E-08	1.02E-07	From the same selection series as PS4451-4456.
SRF SERIES 2, clone 10	PS4542	C	SAME	6.30E-08	4.48E-08	8.86E-08	From the same selection series as PS4451-4456.

Note for run J's ancestor in the SRF table above: the run J ancestor's population replicate plate #5 was an outlier in terms of colony counts. Therefore, five replicates were used for calculating run J's ancestral mutation rate.

Table S1g: Transformed clones after fosfomycin→rifampicin→streptomycin (FRS) exposure. (Refer to Figure 2b.)

Strain description	Lab strain #	Fluctuation assay batch ID	*Mutation rate (MR) measurements w/ 95% CIs*				Notes
			Qualitative MR relative to transformed ancestor	Max likelihood MR	Lower 95% CI	Upper 95% CI	
MUTATOR ANCESTOR TRANSFORMED W/ EMPTY PLASMID	PS4840	a	N/A	1.75E-07	1.26E-07	2.43E-07	The mutator ancestor is PS2533.
MUTATOR ANCESTOR TRANSFORMED W/ EMPTY PLASMID	PS4840	b	N/A	1.30E-07	8.44E-08	2.00E-07	The mutator ancestor is PS2533.
MUTATOR ANCESTOR TRANSFORMED W/ EMPTY PLASMID	PS4840	c	N/A	1.36E-07	9.18E-08	2.02E-07	The mutator ancestor is PS2533.
$FRS1_3$ transformed w/ *mutT*$^+$ plasmid	PS4835	c	SAME	6.74E-08	4.50E-08	1.01E-07	$FRS1_3$ is PS4148.
$FRS1_3$ transformed w/ empty plasmid	PS4836	c	HIGHER	5.03E-07	2.66E-07	9.49E-07	$FRS1_3$ is PS4148.
$FRS1_6$ transformed w/ *mutT*$^+$ plasmid	PS4837	b	SAME	1.52E-07	9.75E-08	2.37E-07	$FRS1_6$ is PS4151.
$FRS1_6$ transformed w/ empty plasmid	PS4838	b	HIGHER	7.22E-07	6.01E-07	8.67E-07	$FRS1_6$ is PS4151.
Mutator ancestor transformed w/ *mutT*$^+$ plasmid	PS4839	a	SAME	2.36E-07	1.57E-07	3.57E-07	The mutator ancestor is PS2533.

Table S2a.

Position	Mutation	Annotation	Description
111,230	$(C)_{6\rightarrow7}$	coding (187/390 nt)	dGTP preferring nucleoside triphosphate pyrophosphohydrolase
257,908	Δ776 bp		*[crl]*
1,134,811	A→G	N28S (AAC→AGC)	flagellar component of cell distal portion of basal body rod
1,785,924	T→G	T397P (ACC→CCC)	phosphoenolpyruvate synthase
2,661,746	A→C	L129R (CTG→CGG)	isc operon transcriptional repressor; suf operon transcriptional activator; oxidative stress and iron starvation inducible; autorepressor
3,201,898	T→G	I3S (ATT→AGT)	fused tRNA nucleotidyl transferase/2'3' cyclic phosphodiesterase/2'nucleotidase and phosphatase
3,253,604	T→C	D204G (GAT→GGT)	LysR family putative transcriptional regulator
3,336,866	A→C	intergenic (63/+97)	acid stress protein; putative BolA family transcriptional regulator/ABC transporter maintaining OM lipid asymmetry, cytoplasmic STAS component
3,474,425	T→G	K43T (AAA→ACA)	30S ribosomal subunit protein S12
3,573,091	T→C	T138A (ACG→GCG)	1,4 alpha glucan branching enzyme

The computational pipeline breseq's predictions for the mutations that were found in both $FRS1_3$ and $FRS1_6$, but not in the single-mutator ancestor.

Table S2b.

Position	Mutation	Annotation	Description
116,249	T→G	D284A (GAT→GCT)	T2SE secretion family protein; P loop ATPase superfamily protein
421,880	A→C	N299H (AAC→CAC)	proline specific permease
653,362	T→G	S376S (TCT→TCG)	sensory histidine kinase in two component regulatory system with CitB
1,195,544	T→C	V141A (GTT→GCT)	isocitrate dehydrogenase; e14 prophage attachment site; tellurite reductase
1,298,612	T→G	F5L (TTT→TTG)	UPF0056 family inner membrane protein
1,482,839	T→G	K8Q (AAA→CAA)	NADH azoreductase, FMN dependent
1,593,880	A→G	pseudogene (246/1461 nt)	pseudogene, AidA homolog
2,057,676	$(C)_{7\rightarrow 8}$	intergenic (+102/+351)	UPF0082 family protein/tRNA Asn
2,130,949	A→C	F374L (TTT→TTG)	putative glycosyl transferase
2,533,869	A→C	T36P (ACT→CCT)	phosphohistidinoprotein hexose phosphotransferase component of PTS system (Hpr)
2,972,258	A→C	E221A (GAA→GCA)	putative NADP(H) dependent aldo keto reductase
3,240,568	T→G	F209V (TTC→GTC)	sodium:serine/threonine symporter
3,455,411	T→G	intergenic (13/ 167)	general secretory pathway component, cryptic/general secretory pathway component, cryptic
3,846,412	A→C	V252V (GTT→GTG)	hexose phosphate transporter
3,891,348	T→G	A371A (GCT→GCG)	tryptophan transporter of low affinity
4,182,778	T→C	S512P (TCT→CCT)	RNA polymerase, beta subunit
4,374,740	A→C	I306M (ATT→ATG)	EF P Lys34 lysylation protein; weak lysine 2,3 aminomutase

The computational pipeline breseq's predictions for the mutations that were found in $FRS1_3$, but not in $FRS1_6$ or the single-mutator ancestor.

Table S2c.

Position	Mutation	Annotation	Description
647,883	A→C	I51M (ATT→ATG)	apo citrate lyase phosphoribosyl dephospho CoA transferase
1,066,992	A→G	Y34H (TAC→CAC)	uncharacterized protein
1,175,903	T→C	G159G (GGT→GGC)	lipoprotein releasing system transmembrane protein
1,256,370	T→G	intergenic (418/+351)	putative adhesin/catalase inhibitor protein; ATPase, K+ dependent, ribosome associated
1,555,425	T→C	D82G (GAC→GGC)	malate dehydrogenase, decarboxylating, NAD requiring; malic enzyme
1,605,296	T→C	N82N (AAT→AAC)	autoinducer 2 binding protein
1,660,739	A→C	N62H (AAC→CAC)	S and N oxide reductase, A subunit, periplasmic
1,970,661	A→C	T108T (ACT→ACG)	methyl accepting protein IV
1,975,465	A→G	V264A (GTA→GCA)	protein that enables flagellar motor rotation
2,244,571	A→C	D221A (GAT→GCT)	S formylglutathione hydrolase
2,401,092	A→C	L155R (CTG→CGG)	NADH:ubiquinone oxidoreductase, chain F
2,456,208	A→C	intergenic (562/+119)	putative fimbrial like adhesin protein/phosphohistidine phosphatase
2,651,740	T→G	E183A (GAA→GCA)	bacterial alpha2 macroglobulin colonization factor ECAM; anti host protease defense factor; periplasmic inner membrane anchored lipoprotein
2,817,916	T→G	intergenic (56/+143)	tRNA Arg/tRNA Arg
2,818,145	T→C	intergenic (10/+53)	
3,194,362	$(G)_{6\rightarrow7}$	coding (1499/1662 nt)	PHB family membrane protein, function unknown
3,560,424	A→C	pseudogene (30/606 nt)	pseudogene, DNA binding transcriptional repressor; regulator; Energy metabolism, carbon: Anaerobic respiration; repressor of the glp operon
3,676,498	A→C	I618S (ATT→AGT)	putative inner membrane anchored periplasmic AsmA family protein
3,850,366	A→C	Y121D (TAT→GAT)	response regulator in two component regulatory system wtih UhpB
3,936,152	A→C	N289H (AAT→CAT)	D ribose ABC transporter permease
4,182,782	A→C	Q513P (CAG→CCG)	RNA polymerase, beta subunit

4,272,989	T→G	I295L (ATC→CTC)	ATPase and DNA damage recognition protein of nucleotide excision repair excinuclease UvrABC

The computational pipeline breseq's predictions for the mutations that were found in $FRS1_6$, but not in $FRS1_3$ or the single-mutator ancestor.

Table S3.

	Media	Total # cells at start of this stage	# resistant cells at start of this stage	Total # cells at end of this stage	# resistant cells at end of this stage	Total # cells transferred to next stage	# resistant cells transferred to next stage
Day 1:	Flask of 10 mL plain LB	15 million	0	10 billion	1	1 thousand	1
Days 2-4:	Flask of 30 mL plain LB	1 thousand	1	30 billion	30 million	100 million	100 thousand
Days 5-6:	Nalidixic acid LB plate	100 million	100 thousand	100 thousand colonies	100 thousand colonies	N/A	N/A

The high numbers of mutant colonies in the fluctuation assay are a true indication of high mutation rates. Table S3 illustrates a hypothetical scenario where a nalidixic acid-resistant cell is seeded into a fluctuation assay flask at the start of Day 2. If this scenario took place, the nalidixic acid (nal) plate would contain ~100,000 colonies, which would appear as an uncountable lawn to the human eye. In the actual fluctuation assays, no lawns were observed on any nal plates. Hence, it is likely that the nal-resistant cells arose during the incubation from Day 2 - Day 4, not beforehand. These calculations exclude the existence of cells that are expected to spontaneously acquire nal resistance after day 2. The calculations thus suggest that a single nal-resistant mutant seeded into the culture by day 2 would be sufficient to cause a lawn to grow on nal, even in the absence of new nal-resistant mutants that normally appear on the nal plates of mutator fluctuation assays. The unlikeliness of the scenario illustrated in this table is compounded by the fact that even if a nal-resistant mutant cell emerged in the day 1 culture, the odds of it being transferred to the day 2 culture are low (0.00001%).

Table: Estimated fitnesses relative to ancestor (w) of 30 evolved populations after 900 generations of propagation under soft selection

w	Soft selection population, rank-ordered by w value
1.14	Ara+3
1.15	Ara+13
1.18	Ara+8
1.19	Ara+7
1.21	Ara+5
1.21	Ara-7
1.24	Ara-10
1.24	Ara+6
1.24	Ara+4
1.25	Ara+1
1.25	Ara+9
1.25	Ara+2
1.26	Ara+10
1.26	Ara-9
1.26	Ara-1
1.26	Ara-8
1.27	Ara-6
1.30	Ara+15
1.32	Ara-12
1.32	Ara-5
1.33	Ara-3
1.34	Ara-11
1.36	Ara+11
1.36	Ara-13
1.38	Ara+12
1.39	Ara-2
1.39	Ara-15
1.39	Ara-4
1.39	Ara-14
1.41	Ara+14

Table: Estimated selection coefficients (s) of 155 evolved clones after 900 generations of soft selection

Strain ID	Population & clone number	Mean selection coefficient (s) across 3 replicates	Standard deviation	Replicate 1	Replicate 2	Replicate 3
PS3838	Ara-1, c1	0.31	0.02	0.29	0.31	0.32
PS3878	Ara-1, c2	0.27	0.02	0.25	0.27	0.29
PS3879	Ara-1, c3	0.23	0.02	0.21	0.22	0.25
PS3880	Ara-1, c4	0.14	0.01	0.15	0.13	0.14
PS3881	Ara-1, c5	0.23	0.04	0.27	0.2	0.22
PS3839	Ara-2, c1	0.28	0.04	0.23	0.3	0.3
PS3882	Ara-2, c2	0.25	0.03	0.23	0.28	0.24
PS3883	Ara-2, c3	0.32	0.03	0.3	0.3	0.35
PS3884	Ara-2, c4	0.23	0.03	0.21	0.27	0.21
PS3885	Ara-2, c5	0.26	0.02	0.27	0.24	0.28
PS3840	Ara-3, c1	0.32	0.04	0.29	0.37	0.3
PS3886	Ara-3, c2	0.38	0.04	0.43	0.36	0.36
PS3887	Ara-3, c3	0.29	0.04	0.33	0.27	0.27
PS3888	Ara-3, c4	0.19	0.02	0.20	0.17	0.21
PS3889	Ara-3, c5	0.32	0.02	0.33	0.31	0.31
PS3841	Ara-4, c1	0.26	0.05	0.30	0.21	0.26
PS3890	Ara-4, c2	0.25	0.00	0.25	0.25	0.25
PS3891	Ara-4, c3	0.21	0.00	0.21	0.21	0.21
PS3892	Ara-4, c4	0.26	0.02	0.26	0.24	0.28
PS3893	Ara-4, c5	0.36	0.04	0.33	0.34	0.40
PS3842	Ara-5, c1	0.18	0.03	0.20	0.18	0.15
PS3894	Ara-5, c2	0.27	0.04	0.31	0.23	0.28
PS3895	Ara-5, c3	0.26	0.04	0.24	0.30	0.23
PS3896	Ara-5, c4	0.27	0.01	0.27	0.28	0.27
PS3897	Ara-5, c5	0.30	0.05	0.35	0.29	0.25
PS3843	Ara-6, c1	0.32	0.04	0.29	0.30	0.37
PS3898	Ara-6, c2	0.24	0.01	0.23	0.25	0.25
PS3899	Ara-6, c3	0.28	0.04	0.23	0.29	0.31
PS3900	Ara-6, c4	0.21	0.03	0.24	0.19	0.21
PS3901	Ara-6, c5	0.20	0.01	0.19	0.20	0.20
PS3844	Ara-7, c1	0.31	0.06	0.38	0.27	0.28

PS3902	Ara-7, c2	0.12	0.02	0.14	0.10	0.12
PS3903	Ara-7, c3	0.24	0.03	0.27	0.21	0.24
PS3904	Ara-7, c4	0.29	0.03	0.27	0.32	0.29
PS3905	Ara-7, c5	0.22	0.03	0.20	0.26	0.21
PS3845	Ara-8, c1	0.24	0.02	0.22	0.26	0.24
PS3906	Ara-8, c2	0.23	0.03	0.19	0.25	0.25
PS3907	Ara-8, c3	0.41	0.06	0.47	0.41	0.36
PS3908	Ara-8, c4	0.24	0.04	0.20	0.27	0.26
PS3909	Ara-8, c5	0.32	0.02	0.31	0.34	0.31
PS3846	Ara-9, c1	0.30	0.03	0.31	0.33	0.27
PS3910	Ara-9, c2	0.27	0.00	0.27	0.27	0.27
PS3911	Ara-9, c3	0.30	0.03	0.33	0.28	0.29
PS3912	Ara-9, c4	0.20	0.03	0.17	0.22	0.20
PS3913	Ara-9, c5	0.24	0.02	0.23	0.26	0.24
PS3847	Ara-10, c1	0.29	0.04	0.27	0.27	0.34
PS3914	Ara-10, c2	0.35	0.04	0.30	0.36	0.38
PS3915	Ara-10, c3	0.27	0.02	0.25	0.29	0.28
PS3916	Ara-10, c4	0.25	0.02	0.25	0.23	0.26
PS3917	Ara-10, c5	0.31	0.03	0.29	0.29	0.35
PS3848	Ara-11, c1	0.22	0.02	0.22	0.20	0.23
PS3928	Ara-11, c2	0.35	0.02	0.34	0.37	0.35
PS3929	Ara-11, c3	0.38	0.06	0.44	0.32	0.37
PS3930	Ara-11, c4	0.41	0.05	0.44	0.44	0.36
PS3931	Ara-11, c5	0.33	0.05	0.30	0.31	0.39
PS3849	Ara-12, c1	0.15	0.01	0.14	0.15	0.15
PS3932	Ara-12, c2	0.31	0.01	0.31	0.32	0.30
PS3933	Ara-12, c3	0.32	0.02	0.31	0.31	0.34
PS3934	Ara-12, c4	0.27	0.01	0.26	0.28	0.28
PS3935	Ara-12, c5	0.25	0.03	0.23	0.24	0.29
PS3850	Ara-13, c1	0.23	0.02	0.23	0.21	0.24
PS3936	Ara-13, c2	0.24	0.02	0.24	0.22	0.26
PS3937	Ara-13, c3	0.28	0.04	0.23	0.31	0.30
PS3938	Ara-13, c4	0.20	0.03	0.22	0.22	0.16
PS3939	Ara-13, c5	0.29	0.02	0.32	0.29	0.27
PS3851	Ara-14, c1	0.22	0.02	0.20	0.22	0.24
PS3940	Ara-14, c2	0.16	0.04	0.13	0.14	0.20
PS3941	Ara-14, c3	0.24	0.02	0.25	0.25	0.21
PS3942	Ara-14, c4	0.28	0.07	0.32	0.20	0.33

PS3943	Ara-14, c5	0.34	0.05	0.33	0.29	0.39
PS3852	Ara-15, c1	0.17	0.02	0.19	0.18	0.15
PS3944	Ara-15, c2	0.30	0.04	0.34	0.26	0.29
PS3945	Ara-15, c3	0.29	0.04	0.28	0.26	0.33
PS3946	Ara-15, c4	0.37	0.02	0.36	0.36	0.39
PS3947	Ara-15, c5	0.31	0.01	0.32	0.31	0.30
PS3853	Ara+1, c1	0.27	0.03	0.24	0.29	0.27
PS3948	Ara+1, c2	0.23	0.03	0.23	0.21	0.26
PS3949	Ara+1, c3	0.23	0.05	0.25	0.18	0.27
PS3950	Ara+1, c4	0.20	0.04	0.15	0.23	0.21
PS3951	Ara+1, c5	0.23	0.02	0.22	0.25	0.23
PS3854	Ara+2, c1	0.25	0.01	0.25	0.26	0.25
PS3952	Ara+2, c2	0.27	0.04	0.27	0.31	0.23
PS3953	Ara+2, c3	0.20	0.03	0.16	0.22	0.22
PS3954	Ara+2, c4	0.26	0.02	0.24	0.26	0.28
PS3955	Ara+2, c5	0.18	0.03	0.18	0.16	0.21
PS3855	Ara+3, c1	0.21	0.03	0.21	0.18	0.24
PS3956	Ara+3, c2	0.22	0.04	0.27	0.20	0.20
PS3957	Ara+3, c3	0.21	0.03	0.24	0.21	0.19
PS3958	Ara+3, c4	0.24	0.03	0.27	0.22	0.22
PS3959	Ara+3, c5	0.27	0.02	0.28	0.28	0.24
PS3856	Ara+4, c1	0.21	0.02	0.20	0.21	0.23
PS3960	Ara+4, c2	0.21	0.02	0.21	0.23	0.20
PS3961	Ara+4, c3	0.22	0.02	0.21	0.24	0.22
PS3962	Ara+4, c4	0.22	0.03	0.20	0.22	0.25
PS3963	Ara+4, c5	0.26	0.02	0.27	0.26	0.24
PS3857	Ara+5, c1	0.25	0.02	0.26	0.23	0.26
PS3964	Ara+5, c2	0.26	0.03	0.29	0.25	0.23
PS3965	Ara+5, c3	0.21	0.03	0.19	0.21	0.24
PS3966	Ara+5, c4	0.22	0.02	0.23	0.22	0.20
PS3967	Ara+5, c5	0.27	0.02	0.28	0.25	0.29
PS3858	Ara+6, c1	0.23	0.03	0.23	0.26	0.20
PS3978	Ara+6, c2	0.28	0.03	0.26	0.26	0.32
PS3979	Ara+6, c3	0.27	0.03	0.26	0.25	0.30
PS3980	Ara+6, c4	0.33	0.01	0.32	0.33	0.33
PS3981	Ara+6, c5	0.26	0.03	0.24	0.26	0.29
PS3859	Ara+7, c1	0.21	0.02	0.20	0.24	0.20
PS3982	Ara+7, c2	0.27	0.02	0.29	0.25	0.26

PS3983	Ara+7, c3	0.19	0.02	0.18	0.21	0.17
PS3984	Ara+7, c4	0.14	0.05	0.10	0.19	0.13
PS3985	Ara+7, c5	0.18	0.03	0.16	0.22	0.16
PS3860	Ara+8, c1	0.23	0.03	0.25	0.25	0.19
PS3986	Ara+8, c2	0.23	0.04	0.20	0.22	0.27
PS3987	Ara+8, c3	0.20	0.01	0.19	0.20	0.21
PS3988	Ara+8, c4	0.17	0.02	0.17	0.15	0.19
PS3989	Ara+8, c5	0.31	0.02	0.31	0.32	0.29
PS3861	Ara+9, c1	0.20	0.04	0.24	0.17	0.20
PS3990	Ara+9, c2	0.28	0.01	0.27	0.28	0.28
PS3991	Ara+9, c3	0.19	0.01	0.19	0.19	0.20
PS3992	Ara+9, c4	0.21	0.01	0.22	0.22	0.20
PS3993	Ara+9, c5	0.20	0.03	0.20	0.23	0.17
PS3862	Ara+10, c1	0.21	0.03	0.23	0.21	0.18
PS3994	Ara+10, c2	0.19	0.02	0.17	0.20	0.21
PS3995	Ara+10, c3	0.22	0.04	0.18	0.25	0.22
PS3996	Ara+10, c4	0.26	0.02	0.25	0.28	0.25
PS3997	Ara+10, c5	0.21	0.01	0.20	0.22	0.22
PS3863	Ara+11, c1	0.17	0.03	0.19	0.14	0.17
PS3998	Ara+11, c2	0.20	0.03	0.22	0.17	0.21
PS3999	Ara+11, c3	0.21	0.04	0.24	0.21	0.17
PS4000	Ara+11, c4	0.19	0.04	0.23	0.16	0.18
PS4001	Ara+11, c5	0.09	0.04	0.06	0.08	0.13
PS3864	Ara+12, c1	0.20	0.03	0.17	0.22	0.22
PS4002	Ara+12, c2	0.18	0.01	0.19	0.19	0.17
PS4003	Ara+12, c3	0.23	0.02	0.21	0.25	0.24
PS4004	Ara+12, c4	0.18	0.02	0.18	0.17	0.20
PS4005	Ara+12, c5	0.16	0.02	0.13	0.17	0.17
PS3865	Ara+13, c1	0.20	0.01	0.19	0.19	0.21
PS4006	Ara+13, c2	0.22	0.01	0.22	0.22	0.21
PS4007	Ara+13, c3	0.16	0.04	0.15	0.13	0.20
PS4008	Ara+13, c4	0.17	0.03	0.15	0.21	0.16
PS4009	Ara+13, c5	0.26	0.02	0.25	0.29	0.25
PS3866	Ara+14, c1	0.15	0.02	0.16	0.16	0.13
PS4010	Ara+14, c2	0.30	0.02	0.29	0.32	0.28
PS4011	Ara+14, c3	0.20	0.02	0.18	0.20	0.22
PS4012	Ara+14, c4	0.18	0.02	0.18	0.16	0.19
PS4013	Ara+14, c5	0.25	0.01	0.24	0.25	0.25

PS3867	Ara+15, c1	0.20	0.02	0.20	0.18	0.22
PS4014	Ara+15, c2	0.26	0.02	0.27	0.24	0.26
PS4015	Ara+15, c3	0.21	0.01	0.21	0.22	0.21
PS4016	Ara+15, c4	0.23	0.02	0.22	0.25	0.23
PS4017	Ara+15, c5	0.09	0.02	0.09	0.11	0.08

Table: Estimated mutation rates for 155 clones after 900 generations of soft selection, as measured by fluctuation assays

	Clone #	Max likelihood:	Lower 95% CI	Upper 95% CI
Ara-1	1	3.16E-08	2.12E-08	4.74E-08
	2	2.21E-08	1.51E-08	3.22E-08
	3	1.84E-08	1.24E-08	2.73E-08
	4	1.31E-08	8.27E-09	2.08E-08
	5	2.04E-08	1.37E-08	3.05E-08
Ara-2	1	1.58E-08	9.55E-09	2.40E-08
	2	1.32E-08	7.07E-09	2.46E-08
	3	7.12E-09	4.01E-09	1.27E-08
	4	9.86E-09	5.62E-09	1.73E-08
	5	2.01E-08	1.26E-08	3.19E-08
Ara-3	1	5.62E-08	3.39E-08	7.59E-08
	2	5.04E-08	3.70E-08	6.87E-08
	3	2.13E-08	1.39E-08	3.26E-08
	4	3.33E-07	2.60E-07	4.27E-07
	5	8.28E-10	1.17E-10	5.88E-09
	6	9.47E-09	5.07E-09	1.77E-08
	7	5.18E-08	3.69E-08	7.28E-08
	8	1.13E-08	6.52E-09	1.95E-08
	9	4.96E-08	3.32E-08	7.42E-08
	10	5.96E-10	8.40E-11	4.23E-09
Ara-4	1	2.51E-09	1.20E-09	1.51E-08
	2	3.70E-08	2.40E-08	5.68E-08
	3	4.76E-08	3.44E-08	6.57E-08
	4	2.14E-08	1.39E-08	3.28E-08
	5	6.43E-09	3.13E-09	1.32E-08
Ara-5	1	6.31E-08	3.76E-08	9.46E-08
	2	1.73E-08	1.15E-08	2.61E-08
	3	3.18E-08	2.28E-08	4.44E-08
	4	1.43E-08	8.05E-09	2.54E-08
	5	2.13E-08	1.38E-08	3.27E-08
Ara-6	1	7.08E-09	2.99E-09	1.50E-08
	2	1.18E-08	6.65E-09	2.09E-08
	3	9.52E-09	5.39E-09	1.68E-08
	4	3.59E-08	2.30E-08	5.61E-08
	5	7.13E-09	3.44E-09	1.48E-08

Ara-7	1	5.01E-08	2.99E-08	6.69E-08
	2	2.99E-08	2.02E-08	4.41E-08
	3	3.62E-08	2.62E-08	5.00E-08
	4	1.92E-08	1.24E-08	2.96E-08
	5	1.44E-08	8.47E-09	2.46E-08
Ara-8	1	1.58E-08	7.51E-09	2.67E-08
	2	2.40E-08	1.45E-08	3.97E-08
	3	2.11E-08	1.21E-08	3.69E-08
	4	4.44E-08	2.88E-08	6.84E-08
	5	3.21E-08	2.02E-08	5.08E-08
Ara-9	1	6.31E-09	2.69E-09	1.51E-08
	2	1.93E-08	9.97E-09	3.73E-08
	3	3.51E-08	2.09E-08	5.87E-08
	4	1.78E-08	1.05E-08	3.02E-08
	5	UNKNOWN	UNKNOWN	UNKNOWN
Ara-10	1	5.62E-08	3.80E-08	8.51E-08
	2	1.35E-08	7.65E-09	2.39E-08
	3	9.92E-09	5.53E-09	1.78E-08
	4	2.22E-08	1.46E-08	3.37E-08
	5	6.06E-08	4.51E-08	8.15E-08
Ara-11	1	4.47E-09	1.91E-09	1.07E-08
	2	6.85E-08	5.09E-08	9.22E-08
	3	1.77E-08	1.04E-08	3.00E-08
	4	1.52E-09	2.14E-10	1.09E-08
	5	1.88E-08	1.21E-08	2.91E-08
Ara-12	1	1.78E-08	8.43E-09	2.99E-08
	2	6.12E-08	4.47E-08	8.39E-08
	3	2.19E-08	1.28E-08	3.76E-08
	4	5.53E-08	3.58E-08	8.55E-08
	5	7.97E-09	3.60E-09	1.76E-08
Ara-13	1	1.26E-07	7.51E-08	1.68E-07
	2	6.43E-09	2.76E-09	1.49E-08
	3	4.36E-08	2.95E-08	6.42E-08
	4	2.02E-08	9.94E-09	4.11E-08
	5			
Ara-14	1	1.12E-07	7.51E-08	1.68E-07
	2	1.44E-08	7.15E-09	2.89E-08
	3	8.68E-08	6.32E-08	1.19E-07
	4	2.44E-08	1.67E-08	3.56E-08
	5	6.67E-08	5.10E-08	8.72E-08

Ara-15	1	1.78E-08	1.06E-08	2.99E-08
	2	1.81E-09	6.53E-10	5.02E-09
	3	3.06E-08	2.00E-08	4.68E-08
	4	1.41E-08	8.88E-09	2.25E-08
	5	1.24E-08	7.60E-09	2.03E-08
Ara+1	1	1.34E-08	7.33E-09	2.44E-08
	2	3.83E-08	2.45E-08	6.00E-08
	3	4.80E-08	3.35E-08	6.86E-08
	4	1.92E-08	1.24E-08	2.97E-08
	5	1.61E-08	9.78E-09	2.64E-08
Ara+2	1	1.48E-09	2.94E-10	7.41E-09
	2	6.29E-08	3.97E-08	9.95E-08
	3	1.67E-08	1.08E-08	2.59E-08
	4	6.69E-10	9.43E-11	4.75E-09
	5	3.35E-08	2.23E-08	5.05E-08
Ara+3	1	9.77E-10	3.12E-10	3.06E-09
	2	7.40E-10	1.04E-10	5.26E-09
	3	1.50E-09	3.14E-10	7.14E-09
	4	7.10E-10	1.00E-10	5.04E-09
	5	9.68E-10	2.03E-10	4.62E-09
Ara+4	1	2.16E-07	1.62E-07	2.87E-07
	2	6.51E-08	4.57E-08	9.29E-08
	3	8.50E-08	6.31E-08	1.14E-07
	4	6.75E-08	4.63E-08	9.84E-08
	5	1.21E-07	9.48E-08	1.55E-07
Ara+5	1	1.90E-08	1.06E-08	3.38E-08
	2	1.72E-09	4.25E-10	6.98E-09
	3	1.16E-08	5.57E-09	2.40E-08
	4	9.24E-09	4.21E-09	2.03E-08
	5	1.44E-08	7.47E-09	2.79E-08
Ara+6	1	5.04E-08	3.12E-08	8.14E-08
	2	1.02E-08	5.63E-09	1.84E-08
	3	2.63E-08	1.47E-08	4.71E-08
	4	4.22E-08	2.56E-08	6.95E-08
	5	2.96E-08	1.83E-08	4.79E-08
Ara+7	1	2.09E-08	1.28E-08	3.43E-08
	2	1.88E-08	1.02E-08	3.46E-08
	3	1.62E-08	8.33E-09	3.14E-08
	4	1.39E-08	7.42E-09	2.59E-08
	5	1.86E-08	1.13E-08	3.05E-08

Ara+8	1	1.78E-08	1.06E-08	2.67E-08
	2	3.16E-08	2.08E-08	4.81E-08
	3	4.37E-08	2.93E-08	6.51E-08
	4	1.55E-08	8.13E-09	2.95E-08
	5	2.44E-08	1.26E-08	4.72E-08
Ara+9	1	1.78E-08	9.55E-09	3.02E-08
	2	1.70E-08	1.03E-08	2.80E-08
	3	2.17E-08	1.36E-08	3.47E-08
	4	3.10E-08	2.08E-08	4.62E-08
	5	2.95E-08	1.89E-08	4.61E-08
Ara+10	1	1.41E-08	7.59E-09	2.40E-08
	2	3.13E-08	2.20E-08	4.44E-08
	3	5.52E-08	3.40E-08	8.95E-08
	4	2.58E-08	1.68E-08	3.97E-08
	5	1.97E-08	1.21E-08	3.19E-08
Ara+11	1	1.00E-09	1.07E-09	4.27E-09
	2	1.29E-08	4.66E-09	3.59E-08
	3	1.21E-08	6.83E-09	2.16E-08
	4	7.17E-08	4.76E-08	1.08E-07
	5	2.34E-08	1.42E-08	3.86E-08
Ara+12	1	8.91E-09	2.99E-09	2.12E-08
	2	5.07E-08	3.58E-08	7.17E-08
	3	7.72E-09	4.60E-09	1.29E-08
	4	2.73E-08	1.64E-08	4.56E-08
	5	2.14E-08	1.39E-08	3.30E-08
Ara+13	1	2.00E-08	1.06E-08	2.99E-08
	2	1.33E-08	8.15E-09	2.16E-08
	3	4.41E-08	3.03E-08	6.41E-08
	4	1.35E-08	8.09E-09	2.26E-08
	5	2.81E-08	1.85E-08	4.27E-08
Ara+14	1	1.26E-08	5.97E-09	2.12E-08
	2	3.09E-08	2.15E-08	4.45E-08
	3	2.74E-08	1.85E-08	4.06E-08
	4	2.10E-08	1.41E-08	3.14E-08
	5	3.31E-08	2.14E-08	5.12E-08
Ara+15	1	1.26E-09	1.06E-09	4.74E-09
	2	1.75E-08	1.02E-08	2.99E-08
	3	1.20E-09	2.39E-10	6.03E-09
	4	9.95E-09	5.70E-09	1.74E-08
	5	2.76E-08	1.87E-08	4.08E-08

		95% lower:	95% upper:
The median Ara$^-$ ancestor used is:	3.55E-08	2.38E-08	5.32E-08
The median Ara$^+$ ancestor used is:	2.36E-08	1.72E-08	3.25E-08

List of candidate loci for whole-genome sequencing analysis (Chapter III)

ada, ahpC-F, alkA-B, crfC, dam, dcm, diaA, dinB-G, dnaA-B-C-D-E-G-N-Q-X, dnaJ, dnaK, dnaT, dps, exo, fnr, fur, glyQ-S, gspB, gyrA, gyrB, hda, hns, holA, holB, holC, holD, holE, katE, katG, lexA, ligA, ligB, mfd, miaA, mukB, mukE, mukF, mutH-L-M-S-T-Y, ndh, nei, nfi, nfo, nrdA, nrdB, nrdD, nrdE, nrdF, nth, ogt, oxyR, phrB, pnp, polA-B, priA, priB, priC, prlC, ratA-C, recA-B-C-D-F-G-J-N-O-R-T, recQ, rep, rpoS, ruvA-B-C, sbcB-C-D, sodB, ssb, topA, topB, tus, umuC-D, ung, uvrA-B-C, uvrD, vsr, xthA, ycdX, yciV, ydaV

Table: Breseq output for PS3840
54
mutations Ancestor: PS174

Ara-3, c1

Mutations

position	Mutation	Annotation	Description
12,325	A→G	K55K (AAA→AAG)	molecular chaperone DnaK
58,259	A→G	G213G (GGT→GGC)	chaperone SurA
122,549	A→G	G155G (GGA→GGG)	protein AmpE
149,268	T→C	H77R (CAC→CGC)	aspartate 1-decarboxylase
198,498	A→G	T251A (ACA→GCA)	UDP pyrophosphate synthase
198,588	C→T	P24S (CCG→TCG)	phosphatidate cytidylyltransferase
204,036	T→C	V75A (GTA→GCA)	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase
215,658	A→G	E162E (GAA→GAG)	tRNA lysidine(34) synthetase TilS
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
518,532	G→T	intergenic (-3/-314)	ureidoglycolate dehydrogenase/acyl-CoA synthetase FdrA
550,485	(G)7→8	intergenic (-123/+45)	hypothetical protein/hypothetical protein
688,067	(T)9→8	intergenic (+82/-121)	PTS N-acetylglucosamine transporter subunit IIABC/glutamine--tRNA ligase
739,553	C→T	N19N (AAC→AAT)	2-oxoglutarate dehydrogenase subunit E1
765,319	A→G	L45P (CTG→CCG)	zinc transporter ZitB
835,029	C→T	Q25* (CAG→TAG)	23S rRNA (adenine(1618)-N(6))-methyltransferase
850,497	C→T	A373T (GCC→ACC)	DUF1479 domain-containing protein
996,741	G→A	E1108E (GAG→GAA)	chromosome partition protein MukB
1,062,292	(C)7→6	coding (603/1140 nt)	O-antigen capsule outer membrane auxillary protein export channel
1,112,148	(G)7→8	noncoding (73/88 nt)	tRNA-Ser
1,185,088	(G)5→6	intergenic (-122/+25)	peptidoglycan-binding protein LysM/transcription-repair coupling factor

1,329,736	T→C	I106T (ATC→ACC)	DNA topoisomerase 1
1,348,131	C→T	D116N (GAT→AAT)	hypothetical protein
1,615,565	(A)6→5	coding (30/300 nt)	transposase
1,651,850	(C)6→7	coding (743/1035 nt)	AI-2 transporter TqsA
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,073,972	G→A	Q283* (CAG→TAG)	outer membrane assembly protein AsmA
2,136,455	T→C	E152E (GAA→GAG)	transcriptional regulator
2,195,257	(C)9→8	coding (110/837 nt)	S-formylglutathione hydrolase YeiG
2,273,508	C→T	P234S (CCG→TCG)	acetyl-CoA acetyltransferase
2,623,106	G→A	A39V (GCG→GTG)	pyridoxine 5'-phosphate synthase
2,657,487	G→A	intergenic (-68/-67)	ribosomal large subunit pseudouridine synthase D/outer membrane protein assembly factor BamD
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,752,924	C→T	R324R (CGC→CGT)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,855,987	T→C	N354S (AAC→AGC)	exonuclease V subunit gamma
3,095,115	A→G	R68R (CGT→CGC)	YgiQ family radical SAM protein
3,136,426	A→G	R174R (CGT→CGC)	undecaprenyl-diphosphatase
3,183,328	G→A	R73H (CGC→CAC)	hypothetical protein
3,338,605	A→G	A145A (GCA→GCG)	tRNA dihydrouridine synthase DusB
3,372,913	G→A	V74V (GTC→GTT)	30S ribosomal protein S5
3,598,150	G→A	P177L (CCG→CTG)	cytochrome-c peroxidase
3,607,791	A→G	L18L (TTG→CTG)	hypothetical protein
3,928,944	C→T	A243V (GCG→GTG)	transcription termination factor Rho
3,973,951	Δ1 bp	coding (355/1023 nt)	lysophospholipase L2
4,057,563	A→G	E42E (GAA→GAG)	hypothetical protein
4,101,092	C→T	E47K (GAA→AAA)	HslU--HslV peptidase ATPase subunit

4,435,818	A→G	intergenic (+8/-6)	sugar ABC transporter ATP-binding protein/sugar ABC transporter permease
4,557,387	(G)5→6	intergenic (-105/+57)	ATPase AAA/hypothetical protein
4,616,454	A→G	D309G (GAC→GGC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3886
43
mutations Ara-3, c2

Ancestor: PS174

Mutations

position	Mutation	Annotation	Description
58,259	A→G	G213G (GGT→GGC)	chaperone SurA
122,549	A→G	G155G (GGA→GGG)	protein AmpE
163,715	(C)6→7	coding (623/705 nt)	sugar fermentation stimulation protein SfsA
198,498	A→G	T251A (ACA→GCA)	UDP pyrophosphate synthase
198,588	C→T	P24S (CCG→TCG)	phosphatidate cytidylyltransferase
204,036	T→C	V75A (GTA→GCA)	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase
253,670	C→T	D331N (GAC→AAC)	flagellar biosynthesis protein FlhA
275,134	(G)6→5	coding (254/711 nt)	hypothetical protein
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
708,695	A→G	V146A (GTC→GCC)	K+-transporting ATPase subunit B
750,844	C→T	P662L (CCG→CTG)	mannosylglycerate hydrolase
850,497	C→T	A373T (GCC→ACC)	DUF1479 domain-containing protein
996,741	G→A	E1108E (GAG→GAA)	chromosome partition protein MukB
1,134,240	C→T	A306T (GCT→ACT)	N-methyl-L-tryptophan oxidase
1,185,088	(G)5→6	intergenic (-122/+25)	peptidoglycan-binding protein LysM/transcription-repair coupling factor
1,348,131	C→T	D116N (GAT→AAT)	hypothetical protein
1,498,555	A→G	T158A (ACT→GCT)	type IV secretion protein Rhs
1,615,565	(A)6→5	coding (30/300 nt)	transposase
1,651,850	(C)6→7	coding (743/1035 nt)	AI-2 transporter TqsA
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase

1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,031,439	G→A	N453N (AAC→AAT)	phosphomannomutase
2,073,972	G→A	Q283* (CAG→TAG)	outer membrane assembly protein AsmA
2,195,257	(C)9→8	coding (110/837 nt)	S-formylglutathione hydrolase YeiG
2,273,508	C→T	P234S (CCG→TCG)	acetyl-CoA acetyltransferase
2,640,604	T→C	V126A (GTA→GCA)	thiol disulfide reductase thioredoxin
2,657,487	G→A	intergenic (-68/-67)	ribosomal large subunit pseudouridine synthase D/outer membrane protein assembly factor BamD
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,944,264	T→C	intergenic (-103/+153)	D-3-phosphoglycerate dehydrogenase/ribose-5-phosphate isomerase
3,095,115	A→G	R68R (CGT→CGC)	YgiQ family radical SAM protein
3,183,328	G→A	R73H (CGC→CAC)	hypothetical protein
3,372,913	G→A	V74V (GTC→GTT)	30S ribosomal protein S5
3,924,321	C→T	R615C (CGC→TGC)	ATP-dependent DNA helicase Rep
3,928,944	C→T	A243V (GCG→GTG)	transcription termination factor Rho
4,101,092	C→T	E47K (GAA→AAA)	HslU--HslV peptidase ATPase subunit
4,435,818	A→G	intergenic (+8/-6)	sugar ABC transporter ATP-binding protein/sugar ABC transporter permease
4,557,387	(G)5→6	intergenic (-105/+57)	ATPase AAA/hypothetical protein

4,616,454	A→G	D309G (GAC→GGC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3887

49 mutations
Ancestor:
PS174 Ara-3, c3

Mutations position	Mutation	Annotation	Description
122,549	A→G	G155G (GGA→GGG)	protein AmpE
198,498	A→G	T251A (ACA→GCA)	UDP pyrophosphate synthase
242,203	(G)8→7	intergenic (+194/-610)	transposase/hypothetical protein
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
850,497	C→T	A373T (GCC→ACC)	DUF1479 domain-containing protein
949,624	T→C	intergenic (-516/-29)	thioredoxin reductase/leucine-responsive regulatory protein
996,741	G→A	E1108E (GAG→GAA)	chromosome partition protein MukB
1,046,822	Δ1 bp	coding (239/1191 nt)	ribosomal RNA large subunit methyltransferase I
1,185,088	(G)5→6	intergenic (-122/+25)	peptidoglycan-binding protein LysM/transcription-repair coupling factor
1,348,131	C→T	D116N (GAT→AAT)	hypothetical protein
1,349,565	(G)7→8	intergenic (-157/+211)	enoyl-ACP reductase/hypothetical protein
1,567,814	A→G	G310G (GGT→GGC)	serine/threonine protein kinase
1,615,565	(A)6→5	coding (30/300 nt)	transposase
1,651,850	(C)6→7	coding (743/1035 nt)	AI-2 transporter TqsA
1,735,768	C→T	V75V (GTG→GTA)	transpeptidase
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,860,456	T→C	intergenic (-69/-34)	LysR family transcriptional regulator/tartrate dehydrogenase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
1,977,320	G→A	V122V (GTG→GTA)	5-hydroxyisourate hydrolase
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,073,972	G→A	Q283* (CAG→TAG)	outer membrane assembly protein AsmA
2,195,257	(C)9→8	coding (110/837 nt)	S-formylglutathione hydrolase YeiG

2,247,103	(C)7→8	coding (525/618 nt)	heme ABC exporter ATP-binding protein CcmA
2,273,508	C→T	P234S (CCG→TCG)	acetyl-CoA acetyltransferase
2,406,553	(C)7→6	intergenic (+243/-123)	long-chain fatty acid transporter/hypothetical protein
2,654,923	G→A	A220V (GCG→GTG)	chaperone protein ClpB
2,657,487	G→A	intergenic (-68/-67)	ribosomal large subunit pseudouridine synthase D/outer membrane protein assembly factor BamD
2,715,819	C→T	R351H (CGC→CAC)	alanine--tRNA ligase
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,887,059	(G)7→8	coding (999/2298 nt)	xanthine dehydrogenase molybdenum-binding subunit XdhA
3,061,132	C→T	S256L (TCG→TTG)	8-amino-7-oxononanoate synthase
3,095,115	A→G	R68R (CGT→CGC)	YgiQ family radical SAM protein
3,183,328	G→A	R73H (CGC→CAC)	hypothetical protein
3,215,347	(A)5→6	coding (54/435 nt)	PTS N-acetylgalactosamine transporter subunit IIA
3,372,913	G→A	V74V (GTC→GTT)	30S ribosomal protein S5
3,625,457	(C)7→8	coding (21/753 nt)	cell division protein
3,893,551	+G	intergenic (+6/-50)	low affinity potassium transport system protein kup/transposase
3,928,944	C→T	A243V (GCG→GTG)	transcription termination factor Rho
3,993,750	C→T	G18D (GGT→GAT)	2-dehydro-3-deoxy-6-phosphogalactonate aldolase
4,101,092	C→T	E47K (GAA→AAA)	HslU--HslV peptidase ATPase subunit
4,196,522	T→C	*534Q (TAA→CAA)	malate synthase A
4,221,003	C→T	A62T (GCC→ACC)	D-xylose-proton symporter
4,435,818	A→G	intergenic (+8/-6)	sugar ABC transporter ATP-binding protein/sugar ABC transporter permease
4,557,387	(G)5→6	intergenic (-105/+57)	ATPase AAA/hypothetical protein

4,616,454	A→G	D309G (GAC→GGC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3888

63 mutations
Ancestor:
PS174 Ara-3, c4

Mutations position	Mutation	Annotation	Description
61,862	C→T	R102W (CGG→TGG)	molecular chaperone DjlA
116,733	T→C	G162G (GGT→GGC)	guanosine monophosphate reductase
122,549	A→G	G155G (GGA→GGG)	protein AmpE
198,498	A→G	T251A (ACA→GCA)	UDP pyrophosphate synthase
275,004	T→C	R128R (AGA→AGG)	hypothetical protein
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
724,436	C→T	P20L (CCG→CTG)	hypothetical protein
850,497	C→T	A373T (GCC→ACC)	DUF1479 domain-containing protein
996,741	G→A	E1108E (GAG→GAA)	chromosome partition protein MukB
1,019,315	(G)7→8	coding (418/1071 nt)	fimbrial-like adhesin protein
1,021,251	(G)7→8	coding (214/711 nt)	periplasmic pilin chaperone
1,166,098	T→C	intergenic (+125/-86)	beta-ketoacyl-ACP reductase/acyl carrier protein
1,185,088	(G)5→6	intergenic (-122/+25)	peptidoglycan-binding protein LysM/transcription-repair coupling factor
1,348,131	C→T	D116N (GAT→AAT)	hypothetical protein
1,369,915	G→A	A444A (GCG→GCA)	sucrose phosphorylase
1,615,565	(A)6→5	coding (30/300 nt)	transposase
1,651,850	(C)6→7	coding (743/1035 nt)	AI-2 transporter TqsA
1,729,731	+C	coding (1367/2103 nt)	oxidoreductase subunit
1,776,766	T→C	K84R (AAA→AGA)	50S ribosomal protein L20
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,073,972	G→A	Q283* (CAG→TAG)	outer membrane assembly protein AsmA
2,103,918	(CCAG)7→8	intergenic (-1227/-3)	tail sheath protein/hypothetical protein
2,160,564	G→A	R90Q (CGA→CAA)	hypothetical protein
2,162,254	A→G	L653L (TTA→TTG)	hypothetical protein
2,177,900	C→T	E203E (GAG→GAA)	oxidoreductase
2,195,257	(C)9→8	coding (110/837 nt)	S-formylglutathione hydrolase YeiG

2,273,508	C→T	P234S (CCG→TCG)	acetyl-CoA acetyltransferase
2,349,304	T→C	D310G (GAT→GGT)	transcriptional regulator
2,416,132	Δ1 bp	coding (272/1539 nt)	multidrug resistance protein EmrY
2,445,097	(C)7→8	coding (1243/1257 nt)	ion channel protein
2,595,028	A→G	F120F (TTT→TTC)	mutarotase superfamily protein, YphB family
2,657,487	G→A	intergenic (-68/-67)	ribosomal large subunit pseudouridine synthase D/outer membrane protein assembly factor BamD
2,716,283	G→C	G196G (GGC→GGG)	alanine--tRNA ligase
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,887,059	(G)7→8	coding (999/2298 nt)	xanthine dehydrogenase molybdenum-binding subunit XdhA
2,899,326	T→C	L223L (TTA→TTG)	XdhC-CoxI family protein with NAD(P)-binding Rossman fold
2,942,080	(C)5→6	coding (143/603 nt)	5-formyltetrahydrofolate cyclo-ligase
3,036,213	G→A	A241V (GCG→GTG)	type II secretion system protein GspE
3,048,371	T→C	intergenic (-200/+155)	glycolate permease GlcA/malate synthase G
3,086,643	G→A	G79D (GGC→GAC)	cystathionine beta-lyase
3,095,115	A→G	R68R (CGT→CGC)	YgiQ family radical SAM protein
3,183,328	G→A	R73H (CGC→CAC)	hypothetical protein
3,372,913	G→A	V74V (GTC→GTT)	30S ribosomal protein S5
3,501,010	G→A	S588S (AGC→AGT)	glycogen debranching enzyme
3,549,346	G→A	G24D (GGT→GAT)	nickel-responsive regulator
3,570,568	(T)8→7	intergenic (-159/-232)	universal stress protein B/universal stress protein A
3,648,282	T→C	I172T (ATC→ACC)	bifunctional glyoxylate/hydroxypyruvate reductase B
3,928,944	C→T	A243V (GCG→GTG)	transcription termination factor Rho
4,101,092	C→T	E47K (GAA→AAA)	HslU--HslV peptidase ATPase subunit
4,155,151	T→C	noncoding (69/85 nt)	tRNA-Tyr

4,278,641	A→G	R404R (CGT→CGC)	multidrug resistance outer membrane protein MdtP
4,335,416	A→G	V641A (GTA→GCA)	lysine decarboxylase inducible
4,377,858	A→G	E151G (GAG→GGG)	tRNA (adenosine(37)-N6)-dimethylallyltransferase MiaA
4,435,818	A→G	intergenic (+8/-6)	sugar ABC transporter ATP-binding protein/sugar ABC transporter permease
4,530,139	(G)7→6	coding (418/903 nt)	fimbrial protein FimH
4,557,387	(G)5→6	intergenic (-105/+57)	ATPase AAA/hypothetical protein
4,616,454	A→G	D309G (GAC→GGC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3889

60 mutations

Ancestor: PS174 Ara-3, c5

Mutations position	Mutation	Annotation	Description
205,940	A→G	D180G (GAC→GGC)	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
242,203	(G)8→7	intergenic (+194/-610)	transposase/hypothetical protein
249,293	(A)5→4	coding (396/768 nt)	class II glutamine amidotransferase
352,377	(G)8→7	intergenic (-139/+47)	transcriptional regulator/membrane protein
473,893	T→C	E471E (GAA→GAG)	hypothetical protein
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
557,763	C→T	R139C (CGC→TGC)	hypothetical protein
643,414	A→G	V183A (GTT→GCT)	LysR family transcriptional regulator
726,271	A→G	M108V (ATG→GTG)	LamB/YcsF family protein
754,540	T→C	V206A (GTG→GCG)	cytochrome bd-I ubiquinol oxidase subunit 2
777,364	T→C	V61A (GTA→GCA)	molybdenum import ATP-binding protein ModC
829,377	G→A	A134A (GCG→GCA)	dehydrogenase
866,524	C→T	Q244* (CAA→TAA)	cyclic di-GMP phosphodiesterase
867,078	T→C	C428C (TGT→TGC)	cyclic di-GMP phosphodiesterase
934,453	Δ1 bp	coding (1275/1659 nt)	OLD family ATP-dependent endonuclease; DUF2813 family protein
935,831	T→C	intergenic (-5/-110)	hypothetical protein/MacA family efflux pump subunit
1,116,531	A→G	Y56H (TAT→CAT)	curli production assembly/transport component CsgF
1,146,686	C→T	S87S (AGC→AGT)	flagellar basal body rod modification protein
1,157,644	C→T	A431A (GCG→GCA)	ribonuclease E
1,167,318	A→G	T271A (ACG→GCG)	beta-ketoacyl-[acyl-carrier-protein] synthase II
1,245,579	C→T	H90H (CAC→CAT)	hypothetical protein
1,425,781	A→G	V17A (GTA→GCA)	cold-shock protein
1,434,043	C→T	G136D (GGT→GAT)	lactate dehydrogenase
1,567,864	C→T	D294N (GAT→AAT)	serine/threonine protein kinase
1,592,123	(T)8→7	coding (141/393 nt)	TIGR00156 family protein
1,733,850	G→A	G296S (GGT→AGT)	pyruvate kinase I
1,738,206	G→A	R285C (CGC→TGC)	FeS cluster assembly protein SufD

1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,832,844	T→C	E119E (GAA→GAG)	oxidoreductase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,116,047	(C)7→8	coding (351/1224 nt)	protein SERAC1
2,211,437	T→C	T388A (ACT→GCT)	PTS fructose transporter subunit EIIBC
2,547,178	G→A	A415A (GCG→GCA)	exopolyphosphatase
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,768,317	T→C	E179E (GAA→GAG)	adenylyl-sulfate kinase
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,789,067	(T)8→7	intergenic (-294/-25)	oxidoreductase/hypothetical protein
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,834,573	T→C	L196P (CTA→CCA)	transcriptional regulator
2,855,110	A→T	D646E (GAT→GAA)	exonuclease V subunit gamma
2,877,995	+T	coding (751/837 nt)	5-keto-4-deoxyuronate isomerase
2,971,929	A→G	intergenic (-127/-9)	hypothetical protein/hypothetical protein
3,110,964	T→C	R14R (AGA→AGG)	phosphodiesterase
3,313,750	(T)8→9	intergenic (+52/+4)	membrane protein/hypothetical protein
3,320,151	C→T	intergenic (-232/+198)	metalloprotease TldD/membrane protein
3,343,219	A→T	D99V (GAT→GTT)	multidrug efflux RND transporter permease subunit
3,541,662	C→T	A254T (GCT→ACT)	arabinose efflux transporter
3,736,409	T→C	V296A (GTT→GCT)	ADP-heptose--LPS heptosyltransferase
3,810,685	(A)9→8	intergenic (-27/+48)	DNA-binding response regulator/acetolactate synthase isozyme 1 small subunit
3,869,974	G→A	G264G (GGC→GGT)	phosphate ABC transporter permease
3,873,549	C→T	M162I (ATG→ATA)	glutamine--fructose-6-phosphate aminotransferase
3,899,393	C→T	A98V (GCG→GTG)	ribokinase
3,918,533	C→T	S468S (AGC→AGT)	PLP-dependent threonine dehydratase
4,114,480	Δ1 bp	coding (539/2181 nt)	catalase/peroxidase HPI
4,389,947	A→G	E164G (GAG→GGG)	hypothetical protein
4,502,023	T→C	T243A (ACG→GCG)	protein FecR
4,615,902	G→A	R125H (CGC→CAC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS4326

41 mutations Ara-3, c6

Ancestor: PS174

Mutations position	Mutation	Annotation	Description
86,722	(T)5→6	intergenic (-210/-450)	leu operon leader peptide/transcriptional regulator
95,273	A→G	T353A (ACC→GCC)	peptidoglycan synthase FtsI
95,611	C→T	R465R (CGC→CGT)	peptidoglycan synthase FtsI
165,940	C→T	P344S (CCG→TCG)	ATP-dependent helicase
404,303	C→T	P83L (CCG→CTG)	thiamine-monophosphate kinase
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
550,485	(G)7→6	intergenic (-123/+45)	hypothetical protein/hypothetical protein
1,004,229	T→C	intergenic (-158/+447)	phosphoporin protein E/asparagine--tRNA ligase
1,027,997	T→C	S315P (TCC→CCC)	ABC transporter ATP-binding protein
1,120,321	C→T	intergenic (+27/-68)	hypothetical protein/RNase III inhibitor
1,167,186	C→T	R227C (CGT→TGT)	beta-ketoacyl-[acyl-carrier-protein] synthase II
1,282,766	C→T	A918V (GCC→GTC)	nitrate reductase subunit alpha
1,330,510	C→T	A364V (GCG→GTG)	DNA topoisomerase 1
1,346,338	T→C	Q315R (CAG→CGG)	exoribonuclease II
1,424,981	(C)6→7	coding (300/591 nt)	DNA invertase
1,445,883	A→G	D469G (GAC→GGC)	hypothetical protein
1,452,077	T→C	T39A (ACC→GCC)	azoreductase
1,485,138	G→A	A89A (GCG→GCA)	ABC transporter permease
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,833,442	T→C	D239G (GAC→GGC)	sugar kinase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,010,659	T→C	T126A (ACG→GCG)	hypothetical protein
2,306,975	(C)5→4	coding (953/1206 nt)	L-rhamnonate dehydratase
2,406,553	(C)7→8	intergenic (+243/-123)	long-chain fatty acid transporter/hypothetical protein
2,613,244	T→C	H1006R (CAC→CGC)	phosphoribosylformylglycinamidine synthase
2,722,614	G→A	V7I (GTC→ATC)	sorbitol 6-phosphate dehydrogenase

2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,928,105	T→C	intergenic (+59/+137)	tRNA-modifying protein YgfZ/hemolysin III family protein
3,000,408	G→A	G282D (GGC→GAC)	hypothetical protein
3,049,414	T→C	E428E (GAA→GAG)	malate synthase G
3,098,271	C→T	P625P (CCG→CCA)	DNA topoisomerase IV subunit A
3,258,862	G→A	A148V (GCC→GTC)	phosphoglucosamine mutase
3,707,155	C→T	P61L (CCG→CTG)	PTS mannitol transporter subunit IIABC
3,944,484	(C)8→7	noncoding (79/87 nt)	tRNA-Leu
4,271,209	C→T	R238R (CGC→CGT)	heme lyase NrfEFG subunit NrfE
4,595,450	A→G	noncoding (71/87 nt)	tRNA-Leu

Table: Breseq output for PS4322

49 mutations Ancestor: PS174

Ara-3, c7

Mutations position	Mutation	Annotation	Description
122,549	A→G	G155G (GGA→GGG)	protein AmpE
198,498	A→G	T251A (ACA→GCA)	UDP pyrophosphate synthase
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
640,131	(T)9→8	intergenic (+49/+6)	cold-shock protein CspE/camphor resistance protein CrcB
706,457	G→T	A18D (GCC→GAC)	two-component sensor histidine kinase
850,497	C→T	A373T (GCC→ACC)	DUF1479 domain-containing protein
899,485	A→G	T180A (ACA→GCA)	baseplate assembly protein
920,580	C→T	intergenic (-139/+79)	arginine ABC transporter ATP-binding protein ArtP/lipoprotein
996,741	G→A	E1108E (GAG→GAA)	chromosome partition protein MukB
1,028,684	(A)7→8	coding (1630/1908 nt)	ABC transporter ATP-binding protein
1,133,042	(T)7→8	coding (4/576 nt)	hypothetical protein
1,185,088	(G)5→6	intergenic (-122/+25)	peptidoglycan-binding protein LysM/transcription-repair coupling factor
1,329,747	A→G	T110A (ACC→GCC)	DNA topoisomerase 1
1,348,131	C→T	D116N (GAT→AAT)	hypothetical protein
1,615,565	(A)6→5	coding (30/300 nt)	transposase
1,651,850	(C)6→7	coding (743/1035 nt)	AI-2 transporter TqsA
1,718,984	G→A	R102H (CGT→CAT)	cyclopropane-fatty-acyl-phospholipid synthase
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,910,251	T→C	G265G (GGT→GGC)	phosphoribosylglycinamide formyltransferase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,073,972	G→A	Q283* (CAG→TAG)	outer membrane assembly protein AsmA
2,195,257	(C)9→8	coding (110/837 nt)	S-formylglutathione hydrolase YeiG
2,240,779	G→A	A97A (GCG→GCA)	DNA-binding response regulator
2,273,508	C→T	P234S (CCG→TCG)	acetyl-CoA acetyltransferase

2,657,487	G→A	intergenic (-68/-67)	ribosomal large subunit pseudouridine synthase D/outer membrane protein assembly factor BamD
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,856,077	T→C	D324G (GAC→GGC)	exonuclease V subunit gamma
3,030,162	G→A	D359D (GAC→GAT)	type II secretion system protein GspL
3,095,115	A→G	R68R (CGT→CGC)	YgiQ family radical SAM protein
3,183,328	G→A	R73H (CGC→CAC)	hypothetical protein
3,212,761	T→C	L307P (CTG→CCG)	tagatose-bisphosphate aldolase
3,311,652	G→A	P172S (CCG→TCG)	malate dehydrogenase
3,372,913	G→A	V74V (GTC→GTT)	30S ribosomal protein S5
3,525,425	A→G	W244R (TGG→CGG)	branched-chain amino acid ABC transporter permease
3,570,568	(T)8→7	intergenic (-159/-232)	universal stress protein B/universal stress protein A
3,729,365	A→G	T167A (ACT→GCT)	hypothetical protein
3,773,470	(G)6→7	coding (110/2319 nt)	alpha-xylosidase
3,928,944	C→T	A243V (GCG→GTG)	transcription termination factor Rho
4,101,092	C→T	E47K (GAA→AAA)	HslU--HslV peptidase ATPase subunit
4,435,818	A→G	intergenic (+8/-6)	sugar ABC transporter ATP-binding protein/sugar ABC transporter permease
4,493,255	A→G	Y128H (TAC→CAC)	transposase
4,557,387	(G)5→6	intergenic (-105/+57)	ATPase AAA/hypothetical protein
4,572,277	G→A	H202Y (CAT→TAT)	4-hydroxyphenylacetate catabolism regulatory protein HpaA
4,616,454	A→G	D309G (GAC→GGC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS4323

42 mutations Ancestor: PS174

Ara-3, c8

Mutations

position	Mutation	Annotation	Description
242,203	(G)8→7	intergenic (+194/-610)	transposase/hypothetical protein
404,303	C→T	P83L (CCG→CTG)	thiamine-monophosphate kinase
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
550,485	(G)7→6	intergenic (-123/+45)	hypothetical protein/hypothetical protein
1,004,229	T→C	intergenic (-158/+447)	phosphoporin protein E/asparagine--tRNA ligase
1,167,186	C→T	R227C (CGT→TGT)	beta-ketoacyl-[acyl-carrier-protein] synthase II
1,282,766	C→T	A918V (GCC→GTC)	nitrate reductase subunit alpha
1,330,510	C→T	A364V (GCG→GTG)	DNA topoisomerase 1
1,346,338	T→C	Q315R (CAG→CGG)	exoribonuclease II
1,377,490	A→G	E413E (GAA→GAG)	glycosyl hydrolase family 65
1,424,981	(C)6→7	coding (300/591 nt)	DNA invertase
1,445,883	A→G	D469G (GAC→GGC)	hypothetical protein
1,485,138	G→A	A89A (GCG→GCA)	ABC transporter permease
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,833,442	T→C	D239G (GAC→GGC)	sugar kinase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
1,981,661	G→A	W4* (TGG→TGA)	DgsA anti-repressor MtfA
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,058,383	(T)5→6	coding (569/1224 nt)	colanic acid biosynthesis glycosyltransferase WcaI
2,112,692	(A)7→6	intergenic (+302/+14)	Presumed portal vertex protein/exonuclease SbcC
2,127,175	G→A	A155V (GCG→GTG)	galactitol permease IIC component
2,292,699	T→C	Y379H (TAT→CAT)	ribonucleoside-diphosphate reductase 1 subunit alpha
2,306,975	(C)5→4	coding (953/1206 nt)	L-rhamnonate dehydratase
2,362,744	T→C	T9A (ACT→GCT)	NUDIX hydrolase
2,406,553	(C)7→8	intergenic (+243/-123)	long-chain fatty acid transporter/hypothetical protein
2,602,856	(C)7→8	coding (589/3282 nt)	hypothetical protein
2,613,244	T→C	H1006R (CAC→CGC)	phosphoribosylformylglycinamidine synthase
2,722,614	G→A	V7I (GTC→ATC)	sorbitol 6-phosphate dehydrogenase

2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
3,000,408	G→A	G282D (GGC→GAC)	hypothetical protein
3,049,414	T→C	E428E (GAA→GAG)	malate synthase G
3,098,271	C→T	P625P (CCG→CCA)	DNA topoisomerase IV subunit A
3,258,862	G→A	A148V (GCC→GTC)	phosphoglucosamine mutase
3,374,025	T→C	A7A (GCA→GCG)	50S ribosomal protein L6
4,271,209	C→T	R238R (CGC→CGT)	heme lyase NrfEFG subunit NrfE
4,271,998	Δ1 bp	coding (1503/1608 nt)	heme lyase NrfEFG subunit NrfE
4,571,185	C→T	G185G (GGG→GGA)	4-hydroxyphenylacetate 3-monooxygenase, oxygenase component
4,582,033	A→G	N381S (AAC→AGC)	methyl-accepting chemotaxis protein

Table: Breseq output for PS4324

53 mutations

Ancestor: PS174

Ara-3, c9

Mutations

position	Mutation	Annotation	Description
122,549	A→G	G155G (GGA→GGG)	protein AmpE
198,498	A→G	T251A (ACA→GCA)	UDP pyrophosphate synthase
286,155	C→T	A540V (GCC→GTC)	intimin-like adhesin FdeC
355,304	C→T	R64C (CGT→TGT)	taurine ABC transporter ATP-binding protein
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
512,677	T→C	intergenic (+3/-54)	cyclic amidohydrolase/purine permease
649,583	A→T	W407R (TGG→AGG)	penicillin-binding protein 2
714,264	C→G	Q101E (CAA→GAA)	hypothetical protein
850,497	C→T	A373T (GCC→ACC)	DUF1479 domain-containing protein
996,741	G→A	E1108E (GAG→GAA)	chromosome partition protein MukB
1,185,088	(G)5→6	intergenic (-122/+25)	peptidoglycan-binding protein LysM/transcription-repair coupling factor
1,188,940	C→T	G274D (GGC→GAC)	membrane protein
1,348,131	C→T	D116N (GAT→AAT)	hypothetical protein
1,387,611	C→T	Q105* (CAA→TAA)	dipeptide epimerase
1,427,728	A→G	Y258H (TAT→CAT)	porin
1,592,123	(T)8→7	coding (141/393 nt)	TIGR00156 family protein
1,615,565	(A)6→5	coding (30/300 nt)	transposase
1,651,850	(C)6→7	coding (743/1035 nt)	AI-2 transporter TqsA
1,661,205	G→A	D377N (GAT→AAT)	two-component sensor histidine kinase
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,073,972	G→A	Q283* (CAG→TAG)	outer membrane assembly protein AsmA
2,103,918	(CCAG)7→8	intergenic (-1227/-3)	tail sheath protein/hypothetical protein
2,112,704	T→C	intergenic (+314/+2)	Presumed portal vertex protein/exonuclease SbcC
2,195,257	(C)9→8	coding (110/837 nt)	S-formylglutathione hydrolase YeiG
2,273,508	C→T	P234S (CCG→TCG)	acetyl-CoA acetyltransferase

2,463,928	T→C	R379R (CGT→CGC)	phosphoenolpyruvate--protein phosphotransferase
2,513,076	A→G	L214L (TTA→TTG)	succinyl-diaminopimelate desuccinylase
2,607,671	G→A	V380M (GTG→ATG)	flavohemoprotein
2,657,487	G→A	intergenic (-68/-67)	ribosomal large subunit pseudouridine synthase D/outer membrane protein assembly factor BamD
2,684,948	A→G	E96G (GAA→GGA)	hydroxyglutarate oxidase
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
3,021,896	(T)9→8	coding (789/1140 nt)	hypothetical protein
3,095,115	A→G	R68R (CGT→CGC)	YgiQ family radical SAM protein
3,183,328	G→A	R73H (CGC→CAC)	hypothetical protein
3,193,869	T→C	A528A (GCA→GCG)	PFL-like enzyme TdcE
3,333,916	C→T	R37C (CGC→TGC)	acetyl-CoA carboxylase biotin carboxylase subunit
3,372,913	G→A	V74V (GTC→GTT)	30S ribosomal protein S5
3,398,080	G→A	G385G (GGC→GGT)	translation elongation factor EF-Tu 1
3,625,457	(C)7→8	coding (21/753 nt)	cell division protein
3,839,098	C→T	M461I (ATG→ATA)	DNA gyrase subunit B
3,928,944	C→T	A243V (GCG→GTG)	transcription termination factor Rho
4,101,092	C→T	E47K (GAA→AAA)	HslU--HslV peptidase ATPase subunit
4,349,853	T→C	R36R (CGT→CGC)	molecular chaperone GroEL
4,435,818	A→G	intergenic (+8/-6)	sugar ABC transporter ATP-binding protein/sugar ABC transporter permease
4,481,275	(G)6→7	intergenic (+107/+294)	integrase/phosphoethanolamine transferase YjgX
4,557,387	(G)5→6	intergenic (-105/+57)	ATPase AAA/hypothetical protein
4,616,454	A→G	D309G (GAC→GGC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS4325

58 mutations Ancestor: PS174

Ara-3, c10

Mutations position	Mutation	Annotation	Description
189,343	C→T	R717H (CGC→CAC)	bifunctional uridylyltransferase/uridylyl-removing protein
198,147	A→G	T134A (ACC→GCC)	UDP pyrophosphate synthase
205,940	A→G	D180G (GAC→GGC)	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
242,203	(G)8→7	intergenic (+194/-610)	transposase/hypothetical protein
249,293	(A)5→4	coding (396/768 nt)	class II glutamine amidotransferase
276,931	C→T	A23A (GCG→GCA)	hypothetical protein
452,782	T→C	intergenic (-7/+539)	Hha toxicity modulator TomB/multidrug efflux RND transporter permease subunit
460,375	(A)6→7	coding (1773/3363 nt)	hypothetical protein
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
754,540	T→C	V206A (GTG→GCG)	cytochrome bd-I ubiquinol oxidase subunit 2
777,364	T→C	V61A (GTA→GCA)	molybdenum import ATP-binding protein ModC
914,865	A→G	H57R (CAT→CGT)	hypothetical protein
924,347	A→G	Y215H (TAT→CAT)	NAD(P)-dependent oxidoreductase
935,831	T→C	intergenic (-5/-110)	hypothetical protein/MacA family efflux pump subunit
961,087	A→G	D194G (GAT→GGT)	dimethyl sulfoxide reductase subunit B
1,053,954	G→A	A39T (GCC→ACC)	hydrogenase-1 operon protein HyaF
1,167,318	A→G	T271A (ACG→GCG)	beta-ketoacyl-[acyl-carrier-protein] synthase II
1,349,565	(G)7→6	intergenic (-157/+211)	enoyl-ACP reductase/hypothetical protein
1,368,412	(T)9→8	intergenic (+41/-172)	thiosulfate sulfurtransferase PspE/sucrose phosphorylase
1,393,316	A→G	F326L (TTC→CTC)	low conductance mechanosensitive channel YnaI
1,405,924	Δ1 bp	coding (705/984 nt)	zinc transporter ZntB
1,457,175	C→T	P204S (CCA→TCA)	hypothetical protein
1,569,540	C→T	G365R (GGA→AGA)	outer membrane autotransporter barrel domain-containing protein
1,637,143	G→A	R627H (CGT→CAT)	dimethyl sulfoxide reductase subunit A
1,638,277	T→C	V167A (GTC→GCC)	dimethyl sulfoxide reductase subunit A

1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,276,886	G→A	T1376I (ACC→ATC)	hypothetical protein
2,486,605	G→A	R337C (CGC→TGC)	ethanolamine utilization protein EutA
2,537,246	C→T	G111G (GGC→GGT)	beta-barrel assembly-enhancing protease
2,716,428	C→T	G148E (GGG→GAG)	alanine--tRNA ligase
2,734,097	G→A	R10C (CGC→TGC)	transcriptional regulator
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,758,213	T→A	L340Q (CTG→CAG)	membrane protein
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,789,067	(T)8→7	intergenic (-294/-25)	oxidoreductase/hypothetical protein
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,823,301	T→C	pseudogene (94/450 nt)	fuculose phosphate aldolase
2,834,573	T→C	L196P (CTA→CCA)	transcriptional regulator
2,847,176	T→C	E1149E (GAA→GAG)	exodeoxyribonuclease V subunit beta
2,855,110	A→T	D646E (GAT→GAA)	exonuclease V subunit gamma
2,877,995	+T	coding (751/837 nt)	5-keto-4-deoxyuronate isomerase
2,902,480	G→A	V223I (GTC→ATC)	putative selenate reductase subunit YgfK
3,001,341	G→A	G593D (GGC→GAC)	hypothetical protein
3,137,030	C→T	intergenic (-83/+8)	undecaprenyl-diphosphatase/dihydroneopterin aldolase
3,169,099	G→A	R108H (CGC→CAC)	hypothetical protein
3,488,740	G→A	T156M (ACG→ATG)	transcriptional regulator
3,559,034	T→C	N398S (AAC→AGC)	ribosome-associated ATPase
3,736,409	T→C	V296A (GTT→GCT)	ADP-heptose--LPS heptosyltransferase
3,737,707	T→C	A40A (GCA→GCG)	ligase
3,897,827	G→A	A245T (GCA→ACA)	ribose ABC transporter permease
4,138,146	(A)8→9	intergenic (+62/-205)	argininosuccinate lyase/hydrogen peroxide-inducible genes activator
4,353,770	A→G	V200A (GTG→GCG)	EF-P beta-lysylation protein EpmB
4,484,920	A→G	V135A (GTG→GCG)	transposase
4,546,084	A→G	R227R (CGT→CGC)	2-hydroxyglutaryl-CoA dehydratase

Table: Breseq output for PS3895

65 mutations Ancestor: PS174

Ara-5, c3

Mutations position	Mutation	Annotation	Description
94,683	T→C	L156P (CTG→CCG)	peptidoglycan synthase FtsI
159,330	A→G	G135G (GGT→GGC)	fimbrial protein
354,824	A→G	I229V (ATC→GTC)	taurine ABC transporter substrate-binding protein
357,742	T→C	L298L (TTA→TTG)	delta-aminolevulinic acid dehydratase
426,979	C→T	P381L (CCG→CTG)	trigger factor
450,038	A→G	I216T (ATC→ACC)	hypothetical protein
458,368	G→A	A181T (GCC→ACC)	transcriptional regulator
468,684	A→G	S500G (AGC→GGC)	molecular chaperone HtpG
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
558,275	C→T	A71V (GCA→GTA)	hypothetical protein
588,320	C→T	S12S (AGC→AGT)	hypothetical protein
938,293	C→T	P414L (CCC→CTC)	macrolide ABC transporter permease/ATP-binding protein MacB
939,433	A→G	intergenic (-137/-186)	cold-shock protein CspD/ATP-dependent Clp protease adaptor ClpS
1,113,754	C→T	P103S (CCG→TCG)	phosphatase
1,160,518	C→T	intergenic (+50/+62)	ribosomal large subunit pseudouridine synthase C/m(7)GTP pyrophosphatase
1,281,213	A→G	K400K (AAA→AAG)	nitrate reductase subunit alpha
1,350,323	G→A	G227G (GGC→GGT)	peptide transport system ATP-binding protein SapF
1,395,345	A→G	W150R (TGG→CGG)	universal stress protein E
1,469,638	A→G	G307G (GGT→GGC)	acetyltransferase
1,562,056	T→C	A64A (GCA→GCG)	fimbrial-like adhesin protein
1,611,931	T→C	S107S (TCA→TCG)	lysozyme
1,730,390	C→T	V236V (GTG→GTA)	oxidoreductase subunit
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
1,972,464	A→G	L79S (TTG→TCG)	excinuclease ABC subunit C
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,068,716	T→C	Y48C (TAC→TGC)	tyrosine kinase

2,074,105	+C	coding (714/1854 nt)	outer membrane assembly protein AsmA
2,075,282	T→C	H47R (CAC→CGC)	deoxycytidine triphosphate deaminase
2,091,451	G→A	L1027L (TTG→TTA)	multidrug resistance protein MdtB
2,105,116	(A)8→7	intergenic (+77/+135)	hypothetical protein/hypothetical protein
2,143,706	G→A	intergenic (-124/+99)	transposase/hypothetical protein
2,169,976	C→T	D235N (GAT→AAT)	ABC transporter substrate-binding protein
2,195,257	(C)9→10	coding (110/837 nt)	S-formylglutathione hydrolase YeiG
2,209,306	T→C	L83L (TTA→TTG)	pseudouridine-5'-phosphate glycosidase
2,405,204	C→T	P79S (CCG→TCG)	long-chain fatty acid transporter
2,603,167	A→G	V93A (GTC→GCC)	hypothetical protein
2,674,546	T→C	V486A (GTG→GCG)	DNA repair protein RecN
2,748,760	T→C	V216A (GTG→GCG)	hydrogenase expression/formation protein HypE
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
3,294,992	+T	coding (560/1419 nt)	glutamate synthase subunit beta
3,377,568	C→T	G74D (GGT→GAT)	30S ribosomal protein S3
3,384,708	(C)6→5	coding (427/1953 nt)	type II secretion system protein GspD
3,482,859	T→C	L392P (CTG→CCG)	transcriptional regulator MalT
3,625,339	A→G	S47P (TCA→CCA)	cell division protein
3,711,693	G→A	M112I (ATG→ATA)	hypothetical protein
3,877,265	G→A	N159N (AAC→AAT)	ATP synthase subunit beta
3,931,945	A→G	T5A (ACT→GCT)	UDP-N-acetyl glucosamine 2-epimerase
3,999,040	G→A	G115D (GGC→GAC)	ubiquinone biosynthesis protein UbiB
4,106,206	A→G	C30R (TGT→CGT)	primosomal protein N'
4,165,678	A→G	T240A (ACC→GCC)	DNA-directed RNA polymerase subunit beta'
4,272,039	(G)7→8	coding (1544/1608 nt)	heme lyase NrfEFG subunit NrfE
4,286,424	A→G	G66G (GGT→GGC)	D-allose kinase
4,322,943	A→G	P287P (CCA→CCG)	melibiose/sodium symporter
4,503,832	T→C	*75Q (TAA→CAA)	transposase

4,519,211	A→G	G87G (GGT→GGC)	porin
4,560,609	C→T	A139T (GCC→ACC)	SAM-dependent methyltransferase
4,589,209	A→G	M71T (ATG→ACG)	hypothetical protein
4,615,999	G→A	W157* (TGG→TGA)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR
4,620,854	A→G	S637G (AGC→GGC)	murein transglycosylase
4,625,831	A→G	T326A (ACT→GCT)	two-component sensor histidine kinase

Table: Breseq output for PS3850

56 mutations

Ancestor: PS174

Ara-13, c1

Mutations

position	Mutation	Annotation	Description
173,845	T→C	D143D (GAT→GAC)	iron(3+)-hydroxamate-binding protein FhuD
291,745	T→C	I302M (ATA→ATG)	pyridine nucleotide-disulfide oxidoreductase
442,701	C→T	V588V (GTC→GTT)	multidrug ABC transporter permease/ATP-binding protein
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
524,732	C→T	V159V (GTG→GTA)	5-(carboxyamino)imidazole ribonucleotide mutase
705,093	A→G	S473P (TCT→CCT)	two-component sensor histidine kinase
718,637	A→G	N214S (AAT→AGT)	hypothetical protein
722,540	C→T	G268R (GGG→AGG)	dipeptide permease D
761,482	T→C	noncoding (47/76 nt)	tRNA-Lys
794,720	(G)7→8	pseudogene (339/519 nt)	carbohydrate kinase
801,659	C→T	M56I (ATG→ATA)	adenosylmethionine--8-amino-7-oxononanoate aminotransferase BioA
917,552	(C)6→7	intergenic (-270/+21)	arginine/ornithine ABC transporter substrate-binding protein/arginine ABC transporter permease ArtM
1,018,114	(A)8→7	coding (1808/2601 nt)	outer membrane usher protein
1,049,720	G→A	G164D (GGC→GAC)	hydrogenase-1 small chain
1,056,281	T→C	L484P (CTG→CCG)	cytochrome bd-II ubiquinol oxidase subunit 1
1,063,579	C→T	A486A (GCG→GCA)	membrane protein
1,380,846	T→C	G177G (GGT→GGC)	outer membrane protein G
1,601,691	(G)6→5	coding (305/591 nt)	Rac prophage; site-specific recombinase
1,768,598	T→C	pseudogene (690/713 nt)	cyclic di-GMP regulator CdgR
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
1,998,083	G→A	Q278* (CAA→TAA)	transcriptional regulator
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,074,105	+C	coding (714/1854 nt)	outer membrane assembly protein AsmA

2,287,204	C→T	V1222V (GTG→GTA)	adhesin
2,340,717	C→T	G106D (GGT→GAT)	NADH-quinone oxidoreductase subunit H
2,363,628	T→C	E150E (GAA→GAG)	glutathione S-transferase
2,403,690	G→A	N148N (AAC→AAT)	3-ketoacyl-CoA thiolase
2,489,912	G→A	S143S (AGC→AGT)	ethanolamine utilization protein EutG
2,686,651	T→C	G233G (GGT→GGC)	NAD-dependent succinate-semialdehyde dehydrogenase
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,843,209	C→T	D141N (GAC→AAC)	N-acetylmuramoyl-L-alanine amidase
2,997,463	(A)9→8	coding (146/918 nt)	hypothetical protein
3,000,933	T→C	L457P (CTC→CCC)	hypothetical protein
3,056,915	T→C	T55A (ACG→GCG)	DNA-binding protein
3,064,385	C→T	P254P (CCG→CCA)	hypothetical protein
3,108,719	G→A	R284C (CGT→TGT)	DNA topoisomerase IV subunit B
3,115,964	G→A	P206S (CCT→TCT)	dioxygenase
3,128,670	G→A	I306I (ATC→ATT)	bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase
3,267,410	G→A	S187L (TCA→TTA)	transporter
3,323,094	T→C	I352M (ATA→ATG)	membrane protein
3,370,793	T→C	T404A (ACC→GCC)	protein translocase subunit SecY
3,420,937	A→G	L155L (CTA→CTG)	transporter
3,604,897	T→C	V207A (GTG→GCG)	transporter
3,635,421	(C)5→4	coding (828/1020 nt)	peptide ABC transporter permease
3,886,842	C→T	intergenic (-26/+64)	FMN-binding protein MioC/AsnC family transcriptional regulator
3,895,270	IS*1* (–) +9 bp	coding (113-121/420 nt)	D-ribose pyranase
4,201,885	C→T	A225A (GCG→GCA)	acetate operon repressor
4,216,720	A→G	E52G (GAG→GGG)	membrane protein
4,380,869	C→T	P226L (CCG→CTG)	protease modulator HflK
4,500,121	A→G	V530A (GTG→GCG)	fe(3+) dicitrate transporter fecA
4,579,623	T→C	N145D (AAC→GAC)	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase

Table: Breseq output for PS3939

52 mutations

Ancestor: PS174 Ara-13, c5

Mutations position	Mutation	Annotation	Description
1,927	C→T	P531L (CCG→CTG)	bifunctional aspartate kinase/homoserine dehydrogenase I
77,432	C→T	A285T (GCA→ACA)	thiamine-binding periplasmic protein
191,822	C→T	G176D (GGC→GAC)	methionine aminopeptidase
205,400	A→G	intergenic (+2/-2)	beta-hydroxyacyl-ACP dehydratase/acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
213,493	C→T	P325S (CCG→TCG)	lysine decarboxylase constitutive
267,508	A→G	intergenic (+70/+57)	tRNA-Thr/hypothetical protein
273,276	(G)6→7	coding (103/615 nt)	hypothetical protein
440,596	C→T	I49I (ATC→ATT)	transcriptional regulator
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
643,953	A→G	S3S (AGT→AGC)	LysR family transcriptional regulator
915,380	A→G	E52E (GAA→GAG)	23S rRNA (uracil(747)-C(5))-methyltransferase RlmC
1,011,660	C→T	G88G (GGG→GGA)	aliphatic sulfonate ABC transporter permease
1,076,261	A→G	N239S (AAC→AGC)	trimethylamine N-oxide reductase I catalytic subunit
1,241,884	C→T	W113* (TGG→TAG)	K+/H+ antiporter NhaP2
1,331,256	A→G	T613A (ACC→GCC)	DNA topoisomerase 1
1,434,271	A→G	V60A (GTG→GCG)	lactate dehydrogenase
1,584,236	A→T	I56F (ATT→TTT)	sugar efflux transporter
1,722,106	A→G	intergenic (+9/+32)	multidrug resistance protein MdtK/hypothetical protein
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,928,843	A→G	R150R (CGT→CGC)	aspartate--tRNA ligase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein

2,204,273	C→T	A354T (GCG→ACG)	nucleoside permease
2,280,238	A→G	W259R (TGG→CGG)	hypothetical protein
2,516,024	C→T	V222I (GTC→ATC)	tRNA cytosine(34) acetyltransferase TmcA
2,630,492	A→G	F136L (TTT→CTT)	anti-sigma-E factor RseA
2,636,776	A→G	V13A (GTG→GCG)	transcriptional regulator
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,783,123	G→A	G143D (GGC→GAC)	hypothetical protein
2,794,824	A→G	L62L (TTA→TTG)	TPM domain protein phosphatase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,856,191	A→G	L286P (CTC→CCC)	exonuclease V subunit gamma
2,978,755	T→C	V23A (GTG→GCG)	hypothetical protein
3,081,661	T→C	V84V (GTA→GTG)	dienelactone hydrolase
3,138,439	T→C	Q212R (CAG→CGG)	transcriptional activator TtdR
3,182,737	G→A	A7A (GCG→GCA)	hypothetical protein
3,191,276	(A)9→8	intergenic (-233/+43)	transporter/L-serine dehydratase TdcG
3,278,852	G→A	E80K (GAA→AAA)	sugar ABC transporter substrate-binding protein
3,532,876	A→G	V186A (GTT→GCT)	cell division ATP-binding protein FtsE
3,664,483	T→C	V103A (GTC→GCC)	xylose ABC transporter permease
3,665,197	T→C	V341A (GTG→GCG)	xylose ABC transporter permease
3,679,161	G→A	G43G (GGG→GGA)	3-keto-L-gulonate-6-phosphate decarboxylase
3,769,736	G→A	A80T (GCC→ACC)	AsmA family protein
3,810,685	(A)9→8	intergenic (27/+48)	DNA binding response regulator/acetolactate synthase isozyme 1 small subunit
3,882,872	A→G	intergenic (66/+551)	ATP synthase I/16S rRNA methyltransferase
4,085,443	G→A	G48S (GGC→AGC)	repressor CpxP
4,163,165	T→C	C770C (TGT→TGC)	DNA directed RNA polymerase subunit beta
4,455,350	G→A	P108L (CCG→CTG)	aspartate carbamoyltransferase catalytic subunit
4,615,818	T→C	I97T (ATT→ACT)	trifunctional nicotinamide nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3941

134 mutations Ancestor: PS174

Ara-14, c3

Mutations position	Mutation	Annotation	Description
25,559	C→T	G27G (GGC→GGT)	bifunctional riboflavin kinase/FMN adenylyltransferase
47,894	A→G	Q171R (CAG→CGG)	protein FixB
51,205	A→G	T402A (ACG→GCG)	transporter
101,830	G→A	T208T (ACG→ACA)	lipid II flippase FtsW
209,520	C→T	A518V (GCG→GTG)	DNA polymerase III subunit alpha
222,460	C→T	G126S (GGC→AGC)	hypothetical protein
339,821	C→T	A41A (GCG→GCA)	lac repressor
359,153	G→A	intergenic (-518/-5)	delta-aminolevulinic acid dehydratase/outer membrane autotransporter barrel domain-containing protein
387,844	A→G	intergenic (+237/-170)	PAS domain-containing sensor histidine kinase/branched-chain amino acid transport system 2 carrier protein
390,721	T→C	V438A (GTG→GCG)	proline-specific permease ProY
467,950	A→G	H255R (CAC→CGC)	molecular chaperone HtpG
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
479,796	A→G	W118R (TGG→CGG)	TraB family protein
501,669	IS*1* (+) +8 bp	intergenic (+602/+135)	hypothetical protein/hypothetical protein
502,189	T→C	intergenic (-7/+109)	hypothetical protein/tRNA 2-selenouridine(34) synthase MnmH
510,265	T→C	L156P (CTA→CCA)	allantoin transporter
560,206	T→C	Y260C (TAC→TGC)	protease 7
599,383	A→G	D840G (GAC→GGC)	enterobactin synthase subunit F
631,673	A→G	V290A (GTG→GCG)	citrate lyase subunit beta
757,637	A→G	A138A (GCA→GCG)	cell envelope integrity/translocation protein TolA
794,883	A→G	pseudogene (502/519 nt)	carbohydrate kinase
859,941	C→T	S274S (AGC→AGT)	beta-aspartyl-peptidase
876,641	T→C	C139C (TGT→TGC)	multidrug transporter MdfA
952,683	A→G	Q801R (CAG→CGG)	DNA translocase FtsK
1,022,518	G→A	P220P (CCG→CCA)	dihydroorotate dehydrogenase 2
1,029,530	A→G	M147V (ATG→GTG)	paraquat-inducible protein A

1,142,649	C→T	T80M (ACG→ATG)	lipid II flippase MurJ
1,153,352	A→G	K136E (AAA→GAA)	flagellar hook-associated protein 1
1,166,588	T→C	A27A (GCT→GCC)	beta-ketoacyl-[acyl-carrier-protein] synthase II
1,189,978	C→T	intergenic (-218/-44)	membrane protein/lipoprotein-releasing system protein LolC
1,217,410	T→C	intergenic (+253/-79)	Blue light, low temperature and stress induced protein/hypothetical protein
1,283,493	C→A	T1160T (ACC→ACA)	nitrate reductase subunit alpha
1,291,150	G→A	intergenic (+101/-101)	regulator of RpoS/UTP--glucose-1-phosphate uridylyltransferase
1,334,901	A→G	A233A (GCA→GCG)	aconitate hydratase 1
1,339,870	C→T	T314I (ACC→ATC)	LPS assembly protein B
1,340,285	T→C	intergenic (+186/-7)	LPS assembly protein B/orotidine 5'-phosphate decarboxylase
1,368,412	(T)9→8	intergenic (+41/-172)	thiosulfate sulfurtransferase PspE/sucrose phosphorylase
1,404,781	A→G	L62P (CTC→CCC)	diguanylate cyclase
1,436,500	A→G	H615R (CAC→CGC)	hypothetical protein
1,493,094	T→C	T28A (ACC→GCC)	TonB-dependent receptor
1,507,870	C→T	V45M (GTG→ATG)	hypothetical protein
1,633,551	G→A	R58Q (CGG→CAG)	spermidine N1-acetyltransferase
1,702,708	T→C	E40G (GAG→GGG)	oxidoreductase
1,707,922	A→G	T548A (ACC→GCC)	ATP-dependent helicase
1,733,535	G→A	A191T (GCT→ACT)	pyruvate kinase I
1,763,253	G→A	V376V (GTC→GTT)	phosphoenolpyruvate synthase
1,818,400	T→C	T78A (ACC→GCC)	hypothetical protein
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,844,962	T→C	L233L (TTA→CTA)	PrkA family serine protein kinase
1,863,040	C→T	P420P (CCC→CCT)	BCCT family transporter
1,922,608	G→A	A207T (GCC→ACC)	zinc transporter
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,009,189	G→A	intergenic (-24/-3)	transposase/hypothetical protein
2,074,620	G→A	R67* (CGA→TGA)	outer membrane assembly protein AsmA
2,075,220	C→T	V68M (GTG→ATG)	deoxycytidine triphosphate deaminase
2,075,522	G→A	F211F (TTC→TTT)	uridine kinase
2,104,468	T→C	V183A (GTT→GCT)	hypothetical protein
2,119,644	T→C	E46G (GAG→GGG)	hypothetical protein

2,129,328	T→C	Y160Y (TAT→TAC)	transposase
2,175,936	T→C	H58R (CAC→CGC)	D-alanyl-D-alanine endopeptidase
2,182,679	C→T	G141G (GGC→GGT)	CidB/LrgB family autolysis modulator
2,194,090	T→C	Y39C (TAC→TGC)	membrane protein
2,226,430	T→C	V355A (GTT→GCT)	ABC transporter permease
2,284,778	T→C	Y448C (TAC→TGC)	DNA gyrase subunit A
2,304,907	T→A	Q151L (CAG→CTG)	2-keto-3-deoxy-L-rhamnonate aldolase
2,368,514	A→G	C104C (TGT→TGC)	histidine ABC transporter permease
2,405,357	G→A	A130T (GCT→ACT)	long-chain fatty acid transporter
2,432,062	T→C	E153G (GAA→GGA)	glutamate--pyruvate aminotransferase AlaC
2,450,360	T→C	T117A (ACC→GCC)	hypothetical protein
2,451,570	G→A	L76L (TTG→TTA)	DUF1323 family DNA-binding protein
2,476,777	G→A	A455T (GCA→ACA)	PTS N-acetylmuramic acid transporter subunits IIBC
2,502,774	T→C	T278A (ACG→GCG)	DUF1176 domain-containing protein
2,516,979	T→C	N196S (AAC→AGC)	hypothetical protein
2,528,925	G→A	A395T (GCA→ACA)	hydrogenase-4 component F
2,545,562	G→A	C567Y (TGT→TAT)	polyphosphate kinase
2,616,302	(G)7→6	intergenic (-42/-216)	phosphoribosylformylglycinamidine synthase/lytic transglycosylase F
2,659,728	T→C	L193L (TTG→CTG)	P-protein
2,703,482	T→C	G105G (GGT→GGC)	transporter
2,749,220	A→T	D8V (GAT→GTT)	formate hydrogenlyase transcriptional activator
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,811,628	T→C	S96G (AGT→GGT)	flavodoxin
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,825,488	G→A	T228M (ACG→ATG)	sugar kinase
2,850,647	T→C	T953A (ACA→GCA)	protease 3
2,860,897	C→T	E720K (GAA→AAA)	phosphoenolpyruvate--protein phosphotransferase PtsP
2,869,251	T→C	T462A (ACC→GCC)	bifunctional 2-acylglycerophosphoethanolamine acyltransferase/acyl-ACP synthetase
2,958,441	A→G	A334A (GCT→GCC)	D-erythrose-4-phosphate dehydrogenase
3,005,198	A→G	A25A (GCA→GCG)	hypothetical protein
3,009,849	T→C	F62L (TTT→CTT)	hypothetical protein

3,031,037	G→A	P68S (CCC→TCC)	type II secretion system protein GspL
3,036,978	G→A	P673S (CCG→TCG)	type II secretion system protein GspD
3,042,998	G→A	V1002V (GTC→GTT)	hypothetical protein
3,056,473	G→A	T202I (ACC→ATC)	DNA-binding protein
3,114,873	A→G	Q94R (CAG→CGG)	ATP-Grasp family ATPase
3,176,862	T→C	N305S (AAC→AGC)	uronate isomerase
3,221,647	G→A	A31T (GCC→ACC)	galactosamine-6-phosphate isomerase
3,224,694	G→A	G216D (GGC→GAC)	membrane protein
3,307,653	T→C	G81G (GGA→GGG)	AFG1 family ATPase
3,313,750	(T)8→9	intergenic (+52/+4)	membrane protein/hypothetical protein
3,337,818	C→T	C286C (TGC→TGT)	ribosomal protein L11 methyltransferase
3,376,646	G→A	R7C (CGT→TGT)	50S ribosomal protein L29
3,419,270	T→C	E33E (GAA→GAG)	cell filamentation protein Fic
3,452,520	T→C	L323P (CTG→CCG)	carboxypeptidase/penicillin-binding protein 1A
3,465,485	G→A	intergenic (-217/-11)	DNA-binding response regulator/transcription elongation factor GreB
3,482,933	T→C	W417R (TGG→CGG)	transcriptional regulator MalT
3,489,144	A→G	S21S (AGT→AGC)	transcriptional regulator
3,492,559	G→A	R126C (CGC→TGC)	hypothetical protein
3,813,609	C→T	intergenic (+36/-244)	type I toxin-antitoxin system toxin TisB/multidrug resistance protein D
3,903,378	T→C	intergenic (-244/-235)	transcriptional regulator/16S ribosomal RNA
3,936,533	A→G	Q89Q (CAA→CAG)	dTDP-fucosamine acetyltransferase
3,955,702	G→A	A63V (GCA→GTA)	iron donor protein CyaY
4,011,576	C→T	T20I (ACC→ATC)	potassium transporter
4,052,456	G→A	V191M (GTG→ATG)	ribokinase
4,151,616	T→C	intergenic (+83/-52)	5S ribosomal RNA/UDP-N-acetylenolpyruvoylglucosamine reductase
4,155,943	A→G	Y130C (TAC→TGC)	translation elongation factor EF-Tu 2
4,157,878	C→T	F175F (TTC→TTT)	transcription termination/antitermination protein NusG
4,179,083	A→G	intergenic (+10/-33)	endonuclease V/hypothetical protein
4,232,731	C→T	R262C (CGT→TGT)	4-hydroxybenzoate octaprenyltransferase
4,316,408	A→G	V79A (GTA→GCA)	transcriptional regulator

4,343,459	A→G	L102P (CTT→CCT)	protein-disulfide reductase DsbD
4,367,482	(C)6→7	coding (44/3324 nt)	miniconductance mechanosensitive channel MscM
4,368,584	A→G	intergenic (-69/+28)	phosphatidylserine decarboxylase proenzyme/ribosome small subunit-dependent GTPase
4,395,663	A→G	D186G (GAC→GGC)	esterase
4,401,864	C→T	L4L (CTA→TTA)	L-ribulose-5-phosphate 4-epimerase
4,419,857	C→T	R568H (CGC→CAC)	3'-nucleotidase
4,470,490	A→G	D349G (GAC→GGC)	LPS export ABC transporter permease LptF
4,475,811	(G)7→8	intergenic (-59/+4)	fructuronate transporter/gluconate 5-dehydrogenase
4,537,683	(T)9→8	intergenic (-358/-124)	hypothetical protein/hypothetical protein
4,549,286	C→T	A50T (GCC→ACC)	multidrug resistance protein MdtM
4,610,370	G→A	L77L (CTG→CTA)	purine-nucleoside phosphorylase

Table: Breseq output for PS3944

56 mutations — Ancestor: PS174 Ara-15, c2

Mutations position	Mutation	Annotation	Description
205,940	A→G	D180G (GAC→GGC)	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
410,633	T→C	R221R (CGT→CGC)	tRNA 4-thiouridine(8) synthase ThiI
478,245	A→G	N422S (AAC→AGC)	bifunctional UDP-sugar hydrolase/5'-nucleotidase
614,832	A→G	intergenic (+135/-47)	carbon starvation protein A/hypothetical protein
725,450	C→T	G141G (GGC→GGT)	allophanate hydrolase
860,901	C→T	A277V (GCG→GTG)	glutathione import ATP-binding protein GsiA
860,974	(C)8→7	coding (903/1872 nt)	glutathione import ATP-binding protein GsiA
907,990	G→A	A34T (GCC→ACC)	NADPH-dependent oxidoreductase
914,926	(T)5→6	coding (231/489 nt)	hypothetical protein
959,996	A→G	K649E (AAA→GAA)	dimethyl sulfoxide reductase subunit A
1,048,242	G→A	L90F (CTC→TTC)	BAX inhibitor protein
1,112,099	(A)8→7	pseudogene (70/100 nt)	hypothetical protein
1,228,809	A→G	T190T (ACA→ACG)	isomerase/hydrolase
1,260,778	T→C	Q50R (CAG→CGG)	sodium-independent anion transporter
1,321,467	A→G	E218E (GAA→GAG)	phosphatase
1,339,302	C→T	R125C (CGC→TGC)	LPS assembly protein B
1,534,197	C→T	R224H (CGC→CAC)	peptide ABC transporter permease
1,621,688	G→A	L20L (CTC→CTT)	hypothetical protein
1,691,520	(T)9→8	intergenic (+55/-51)	dipeptide and tripeptide permease A/glutathione S-transferase
1,797,787	G→A	S231S (AGC→AGT)	N,N'-diacetylchitobiose permease IIC component
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,895,820	T→C	V329A (GTG→GCG)	hypothetical protein
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,007,567	(C)7→6	coding (372/408 nt)	hypothetical protein
2,159,146	A→G	Q752Q (CAA→CAG)	hypothetical protein
2,331,092	G→A	intergenic (+655/-160)	deubiquitinase/hypothetical protein

2,336,650	T→C	intergenic (-85/+146)	NADH-quinone oxidoreductase subunit M/NADH-quinone oxidoreductase subunit L
2,418,947	G→A	G46D (GGT→GAT)	two-component system sensor histidine kinase EvgS
2,543,881	T→C	Y7H (TAC→CAC)	polyphosphate kinase
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,771,282	G→A	intergenic (-92/-160)	sulfate adenylyltransferase subunit 2/Zn-dependent exopeptidase M28
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
3,008,653	C→T	pseudogene (510/644 nt)	antitoxin of toxin-antitoxin stability system
3,025,354	T→C	pseudogene (55/696 nt)	pyrophosphorylase
3,153,677	T→C	V456A (GTC→GCC)	putrescine aminotransferase
3,224,741	A→G	T232A (ACA→GCA)	membrane protein
3,326,788	G→A	A361V (GCG→GTG)	cell shape-determining protein MreC
3,464,118	T→C	V145V (GTA→GTG)	two-component sensor histidine kinase
3,645,349	A→G	Y484H (TAT→CAT)	biotin sulfoxide reductase
3,762,825	T→C	V690A (GTG→GCG)	bifunctional (p)ppGpp synthetase II/ guanosine-3',5'-bis pyrophosphate 3'-pyrophosphohydrolase
3,788,128	T→C	intergenic (+166/-113)	transposase/autotransporter
3,799,272	G→A	S218S (AGC→AGT)	methionine ABC transporter substrate-binding protein
3,828,914	C→T	E242K (GAA→AAA)	hypothetical protein
3,837,541	G→A	Q96* (CAA→TAA)	DUF937 domain-containing protein
3,866,258	T→C	D226G (GAT→GGT)	transcription antitermination protein BlgG
3,890,185	C→T	D424N (GAT→AAT)	ATPase RavA
3,928,436	G→A	A74T (GCC→ACC)	transcription termination factor Rho
4,010,303	A→G	D257G (GAC→GGC)	Xaa-Pro dipeptidase
4,108,445	C→T	E20K (GAA→AAA)	met repressor
4,174,149	G→A	A521V (GCC→GTC)	phosphomethylpyrimidine synthase ThiC
4,227,058	T→C	G52G (GGT→GGC)	maltoporin
4,616,055	A→G	Y176C (TAC→TGC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

4,616,260	Δ1 bp	coding (732/1233 nt)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3855

64 mutations

Ancestor: PS2717

Ara+3, c1

Mutations

position	Mutation	Annotation	Description
40,939	C→T	P360P (CCG→CCA)	ATP-dependent acyl-CoA ligase
70,867	T→C	D92G (GAC→GGC)	L-arabinose isomerase
205,940	A→G	D180G (GAC→GGC)	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
278,454	(C)5→6	coding (1061/2526 nt)	hypothetical protein
294,213	(G)7→8	intergenic (+483/-44)	transcriptional regulator/cysteine-rich LutA family protein; electron transport chain YkgEFG component
358,875	G→A	intergenic (-240/-283)	delta-aminolevulinic acid dehydratase/outer membrane autotransporter barrel domain-containing protein
414,481	C→T	G203D (GGC→GAC)	membrane protein
440,899	C→T	P150P (CCC→CCT)	transcriptional regulator
442,945	A→G	L81L (CTA→CTG)	multidrug ABC transporter ATP-binding protein
568,795	C→T	intergenic (-266/+5)	type II secretion system protein E/hypothetical protein
709,215	A→G	V538A (GTT→GCT)	potassium-transporting ATPase A chain
787,879	Δ12,090 bp		***16 genes***
843,554	A→G	V160V (GTA→GTG)	outer membrane protein X
864,982	C→T	R181C (CGC→TGC)	glutathione ABC transporter permease
927,084	Δ1 bp	coding (673/1719 nt)	pyruvate dehydrogenase [ubiquinone]
998,942	(G)6→7	coding (804/1848 nt)	transpeptidase
1,066,453	(A)7→6	intergenic (-30/+77)	hypothetical protein/threonine-rich inner membrane protein GfcA
1,259,789	T→C	N380D (AAC→GAC)	sodium-independent anion transporter
1,368,412	(T)9→8	intergenic (+41/-172)	thiosulfate sulfurtransferase PspE/sucrose phosphorylase
1,377,723	A→G	N491S (AAC→AGC)	glycosyl hydrolase family 65
1,418,812	(T)7→6	coding (151/1458 nt)	potassium transporter TrkG
1,503,957	G→A	R1227R (CGG→CGA)	RHS element protein RhsA

1,551,024	A→T	L183M (TTG→ATG)	hypothetical protein
1,556,848	T→C	intergenic (-373/+29)	sulfatase/AraC family transcriptional regulator
1,559,052	(C)6→7	coding (1386/2280 nt)	hypothetical protein
1,571,344	C→T	A186A (GCC→GCT)	histidine kinase
1,719,068	T→C	M130T (ATG→ACG)	cyclopropane-fatty-acyl-phospholipid synthase
1,813,663	G→A	L121L (CTG→CTA)	ABC transporter substrate-binding protein
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,892,985	G→A	P137S (CCG→TCG)	tail-specific protease
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,151,650	(A)8→7	coding (176/3633 nt)	hypothetical protein
2,195,257	(C)9→10	coding (110/837 nt)	S-formylglutathione hydrolase YeiG
2,197,332	T→C	T226A (ACG→GCG)	colicin I receptor
2,295,712	G→A	R157C (CGC→TGC)	protein InaA
2,491,130	G→A	Q13* (CAA→TAA)	ethanolamine utilization protein EutJ
2,578,128	C→T	G503S (GGT→AGT)	molecular chaperone HscA
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,818,291	C→T	G378G (GGC→GGT)	serine transporter
2,823,412	A→G	intergenic (-18/+115)	fuculose phosphate aldolase/LacI family transcriptional regulator
2,930,675	G→A	G419S (GGC→AGC)	6-phospho-beta-glucosidase
3,013,384	Δ1 bp	intergenic (+48/-24)	arabinose-5-phosphate isomerase/capsule polysaccharide export inner-membrane protein KpsE
3,091,884	(C)6→7	coding (63/927 nt)	DUF3828 domain-containing protein
3,095,865	G→A	T326I (ACC→ATC)	cell division protein FtsP
3,206,732	G→A	P369S (CCG→TCG)	galactarate transporter
3,252,797	G→A	P148P (CCC→CCT)	ribosome maturation factor
3,254,001	A→G	N44S (AAC→AGC)	argininosuccinate synthase
3,464,440	G→A	A38V (GCG→GTG)	two-component sensor histidine kinase
3,470,427	G→A	A452A (GCG→GCA)	ferrous iron transporter B
3,560,737	C→T	R185H (CGC→CAC)	membrane protein

3,875,773	T→C	intergenic (-207/+146)	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase/glucosamine-1-phosphate acetyltransferase/ATP synthase epsilon chain
4,041,902	C→T	E218K (GAG→AAG)	porin
4,161,472	C→T	A206V (GCG→GTG)	DNA-directed RNA polymerase subunit beta
4,179,767	(C)6→7	intergenic (+61/-126)	hypothetical protein/DNA-binding protein HU-alpha
4,291,009	C→T	A269T (GCC→ACC)	RpiR family transcriptional regulator
4,311,510	(C)5→4	coding (477/1092 nt)	two-component sensor histidine kinase
4,454,460	A→G	Y89H (TAT→CAT)	aspartate carbamoyltransferase regulatory subunit
4,541,538	A→G	I19I (ATT→ATC)	nucleoside recognition pore and gate family inner membrane transporter
4,599,986	G→A	D447N (GAT→AAT)	peptide chain release factor 3
4,615,610	A→G	T28A (ACC→GCC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3956

46 mutations Ancestor: PS2717

Ara+3, c2

Mutations

position	Mutation	Annotation	Description
40,939	C→T	P360P (CCG→CCA)	ATP-dependent acyl-CoA ligase
70,867	T→C	D92G (GAC→GGC)	L-arabinose isomerase
205,940	A→G	D180G (GAC→GGC)	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
271,979	C→T	G31G (GGG→GGA)	xanthine dehydrogenase
278,454	(C)5→6	coding (1061/2526 nt)	hypothetical protein
407,151	G→A	G483G (GGC→GGT)	1-deoxy-D-xylulose-5-phosphate synthase
444,528	T→C	intergenic (+44/-137)	multidrug ABC transporter ATP-binding protein/nitrogen regulatory protein P-II 2
550,485	(G)7→6	intergenic (-123/+45)	hypothetical protein/hypothetical protein
568,795	C→T	intergenic (-266/+5)	type II secretion system protein E/hypothetical protein
689,146	A→G	D320G (GAC→GGC)	glutamine--tRNA ligase
770,901	C→T	S253N (AGC→AAC)	galactose-1-phosphate uridylyltransferase
927,084	Δ1 bp	coding (673/1719 nt)	pyruvate dehydrogenase [ubiquinone]
968,857	G→A	D597D (GAC→GAT)	formate acetyltransferase 1
1,201,601	A→G	T130A (ACC→GCC)	peptidase T
1,259,789	T→C	N380D (AAC→GAC)	sodium-independent anion transporter
1,551,024	A→T	L183M (TTG→ATG)	hypothetical protein
1,556,848	T→C	intergenic (-373/+29)	sulfatase/AraC family transcriptional regulator
1,559,052	(C)6→7	coding (1386/2280 nt)	hypothetical protein
1,681,755	T→C	L4P (CTT→CCT)	divisome-associated membrane protein Blr
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,842,180	G→A	A229V (GCG→GTG)	aldo/keto reductase
1,884,080	T→C	V19A (GTA→GCA)	hypothetical protein
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,103,918	(CCAG)7→8	intergenic (-1227/-3)	tail sheath protein/hypothetical protein
2,295,712	G→A	R157C (CGC→TGC)	protein InaA

2,590,693	G→A	T321T (ACG→ACA)	3-phenylpropionate/cinnamic acid dioxygenase subunit alpha
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,818,291	C→T	G378G (GGC→GGT)	serine transporter
2,965,839	C→T	L516L (CTG→CTA)	transketolase
3,095,865	G→A	T326I (ACC→ATC)	cell division protein FtsP
3,140,815	(C)6→7	intergenic (+22/-26)	L+-tartrate dehydratase subunit beta/antiporter
3,206,732	G→A	P369S (CCG→TCG)	galactarate transporter
3,252,797	G→A	P148P (CCC→CCT)	ribosome maturation factor
3,699,773	G→A	R966Q (CGG→CAG)	RHS element protein RhsA
3,875,773	T→C	intergenic (-207/+146)	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase/glucosamine-1-phosphate acetyltransferase/ATP synthase epsilon chain
4,041,902	C→T	E218K (GAG→AAG)	porin
4,100,651	C→T	E194K (GAG→AAG)	HslU--HslV peptidase ATPase subunit
4,179,767	(C)6→7	intergenic (+61/-126)	hypothetical protein/DNA-binding protein HU-alpha
4,482,386	T→C	pseudogene (298/1115 nt)	phosphoethanolamine transferase YjgX
4,530,139	(G)7→8	coding (418/903 nt)	fimbrial protein FimH
4,541,538	A→G	I19I (ATT→ATC)	nucleoside recognition pore and gate family inner membrane transporter
4,615,610	A→G	T28A (ACC→GCC)	trifunctional nicotinamide-nucleotide adenylyltransferase/ribosylnicotinamide kinase/transcriptional regulator NadR

Table: Breseq output for PS3856

45 mutations Ancestor: PS2717

Ara+4, c1

Mutations

position	Mutation	Annotation	Description
70,867	T→C	D92G (GAC→GGC)	L-arabinose isomerase
71,498	A→G	V452A (GTC→GCC)	ribulokinase
255,260	A→G	E214G (GAG→GGG)	hypothetical protein
278,454	(C)5→6	coding (1061/2526 nt)	hypothetical protein
380,182	C→T	M241I (ATG→ATA)	MFS transporter AraJ
403,036	(C)6→7	intergenic (+56/-33)	bifunctional diaminohydroxyphosphoribosylaminopyrimidine deaminase/5-amino-6-(5-phosphoribosylamino)uracil reductase/6,7-dimethyl-8-ribityllumazine synthase
422,600	(G)7→6	intergenic (-287/+175)	cytochrome ubiquinol oxidase subunit II/AmpG family muropeptide MFS transporter
865,524	T→C	pseudogene (72/199 nt)	hypothetical protein
879,507	C→T	A129T (GCT→ACT)	hypothetical protein
908,704	A→G	T11A (ACG→GCG)	ribosomal protein S6--L-glutamate ligase
913,058	C→T	A73V (GCT→GTT)	putrescine ABC transporter permease
934,155	A→G	E326G (GAG→GGG)	OLD family ATP-dependent endonuclease; DUF2813 family protein
1,081,210	(T)8→7	coding (885/1260 nt)	hypothetical protein
1,551,024	A→T	L183M (TTG→ATG)	hypothetical protein
1,559,052	(C)6→7	coding (1386/2280 nt)	hypothetical protein
1,570,698	(C)7→8	intergenic (-66/-89)	outer membrane autotransporter barrel domain-containing protein/histidine kinase
1,733,175	G→A	E71K (GAA→AAA)	pyruvate kinase I
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,073,879	G→A	Q314* (CAG→TAG)	outer membrane assembly protein AsmA
2,131,794	G→A	S300S (AGC→AGT)	fructose-bisphosphate aldolase
2,208,359	G→A	A55V (GCA→GTA)	nucleoside permease

2,223,905	T→C	A118A (GCT→GCC)	peptide ABC transporter substrate-binding protein
2,269,616	T→C	S351P (TCA→CCA)	acetoacetate metabolism regulatory protein AtoC
2,295,712	G→A	R157C (CGC→TGC)	protein InaA
2,321,534	A→T	D314E (GAT→GAA)	o-succinylbenzoate synthase
2,329,538	(C)6→7	coding (310/1209 nt)	deubiquitinase
2,685,147	C→T	G162G (GGC→GGT)	hydroxyglutarate oxidase
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,818,291	C→T	G378G (GGC→GGT)	serine transporter
2,913,467	C→T	A198V (GCG→GTG)	guanine/hypoxanthine permease GhxQ
2,955,940	T→C	E352E (GAA→GAG)	class II fructose-bisphosphate aldolase
3,039,519	T→C	E155E (GAA→GAG)	type II secretion system protein GspC
3,095,865	G→A	T326I (ACC→ATC)	cell division protein FtsP
3,360,899	(C)6→5	coding (567/1125 nt)	protein smf
3,511,594	A→G	I222T (ATC→ACC)	oxidoreductase
3,515,583	C→T	A50V (GCC→GTC)	heat-shock protein
3,762,048	A→G	Y431C (TAC→TGC)	bifunctional (p)ppGpp synthetase II/ guanosine-3',5'-bis pyrophosphate 3'-pyrophosphohydrolase
3,875,773	T→C	intergenic (-207/+146)	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase/glucosamine-1-phosphate acetyltransferase/ATP synthase epsilon chain
3,891,544	(G)7→10	intergenic (-90/-133)	ATPase RavA/low affinity potassium transport system protein kup
4,308,608	(G)6→7	coding (535/879 nt)	hypothetical protein

Table: Breseq output for PS3960

48 mutations Ancestor: PS2717

Ara+4, c2

Mutations

position	Mutation	Annotation	Description
70,867	T→C	D92G (GAC→GGC)	L-arabinose isomerase
245,953	A→G	V641A (GTG→GCG)	acyl-CoA dehydrogenase
278,454	(C)5→6	coding (1061/2526 nt)	hypothetical protein
279,790	A→G	I140T (ATT→ACT)	fimbrial chaperone EcpB
292,747	C→T	intergenic (-97/-129)	pyridine nucleotide-disulfide oxidoreductase/transcriptional regulator
403,036	(C)6→7	intergenic (+56/-33)	bifunctional diaminohydroxyphosphoribosylaminopyrimidine deaminase / 5-amino-6-(5-phosphoribosylamino)uracil reductase/6,7-dimethyl-8-ribityllumazine synthase
492,355	A→G	Q152R (CAG→CGG)	ABC transporter ATP-binding protein
511,687	C→T	G125G (GGC→GGT)	cyclic amidohydrolase
837,880	(G)5→6	coding (226/2226 nt)	moderate conductance mechanosensitive channel YbiO
934,155	A→G	E326G (GAG→GGG)	OLD family ATP-dependent endonuclease; DUF2813 family protein
1,016,506	(G)5→4	coding (200/2601 nt)	outer membrane usher protein
1,057,373	T→C	S329S (AGT→AGC)	cytochrome bd-II ubiquinol oxidase subunit 2
1,202,723	C→T	A296A (GCG→GCA)	50S ribosomal protein L16 arginine hydroxylase
1,551,024	A→T	L183M (TTG→ATG)	hypothetical protein
1,559,052	(C)6→7	coding (1386/2280 nt)	hypothetical protein
1,592,123	(T)8→7	coding (141/393 nt)	TIGR00156 family protein
1,696,666	A→G	L44P (CTG→CCG)	anhydro-N-acetylmuramic acid kinase
1,733,175	G→A	E71K (GAA→AAA)	pyruvate kinase I
1,778,398	G→A	C480C (TGC→TGT)	threonine--tRNA ligase
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase / nicotinamidase
1,928,421	T→C	E291G (GAA→GGA)	aspartate--tRNA ligase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein

2,007,567	(C)7→8	coding (372/408 nt)	hypothetical protein
2,073,879	G→A	Q314* (CAG→TAG)	outer membrane assembly protein AsmA
2,295,712	G→A	R157C (CGC→TGC)	protein InaA
2,447,323	G→A	L19L (CTG→CTA)	nucleoside permease NupC
2,643,534	T→C	F697L (TTT→CTT)	protein lysine acetyltransferase
2,674,488	G→A	G467S (GGT→AGT)	DNA repair protein RecN
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,818,291	C→T	G378G (GGC→GGT)	serine transporter
2,913,467	C→T	A198V (GCG→GTG)	guanine/hypoxanthine permease GhxQ
2,955,940	T→C	E352E (GAA→GAG)	class II fructose-bisphosphate aldolase
2,968,295	C→A	A211D (GCC→GAC)	metalloprotease LoiP
3,095,865	G→A	T326I (ACC→ATC)	cell division protein FtsP
3,225,840	G→A	R598H (CGT→CAT)	membrane protein
3,360,899	(C)6→5	coding (567/1125 nt)	protein smf
3,762,048	A→G	Y431C (TAC→TGC)	bifunctional (p)ppGpp synthetase II / guanosine-3',5'-bis pyrophosphate 3'-pyrophosphohydrolase
3,875,773	T→C	intergenic (-207/+146)	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase / glucosamine-1-phosphate acetyltransferase/ATP synthase epsilon chain
3,891,544	(G)7→8	intergenic (-90/-133)	ATPase RavA/low affinity potassium transport system protein kup
4,308,608	(G)6→7	coding (535/879 nt)	hypothetical protein
4,316,831	(A)8→9	intergenic (-188/+137)	transcriptional regulator/arginine decarboxylase
4,454,981	A→G	V231A (GTG→GCG)	aspartate carbamoyltransferase catalytic subunit
4,456,467	Δ1 bp	coding (20/396 nt)	hypothetical protein
4,481,674	A→G	pseudogene (1010/1115 nt)	phosphoethanolamine transferase YjgX

Table: Breseq output for PS3964

59 mutations Ancestor: PS2717

Ara+5, c2

Mutations position	Mutation	Annotation	Description
8,196	T→C	intergenic (-239/-40)	sodium:alanine symporter/transaldolase
70,867	T→C	D92G (GAC→GGC)	L-arabinose isomerase
96,814	A→G	H282R (CAC→CGC)	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2, 6-diaminopimelate ligase
114,959	G→A	Q149* (CAG→TAG)	cell division protein ZapD
242,168	C→T	intergenic (+159/-645)	transposase/hypothetical protein
278,454	(C)5→6	coding (1061/2526 nt)	hypothetical protein
303,153	C→T	R348C (CGC→TGC)	choline transporter
362,281	A→G	T44A (ACT→GCT)	transcriptional regulator
447,657	C→T	T71I (ACA→ATA)	lipoprotein
500,339	(G)6→7	coding (224/711 nt)	hypothetical protein
629,166	A→G	L132P (CTC→CCC)	2-(5''-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase
1,160,626	A→G	L180P (CTG→CCG)	m(7)GTP pyrophosphatase
1,166,278	A→G	D32G (GAC→GGC)	acyl carrier protein
1,278,576	T→C	F157S (TTC→TCC)	nitrate/nitrite transporter NarK
1,342,003	G→A	H239Y (CAC→TAC)	DeoR family transcriptional regulator
1,383,134	G→A	V217I (GTC→ATC)	hypothetical protein
1,454,062	G→A	A557A (GCG→GCA)	ATP-dependent helicase
1,551,024	A→T	L183M (TTG→ATG)	hypothetical protein
1,559,052	(C)6→7	coding (1386/2280 nt)	hypothetical protein
1,654,391	T→C	K236K (AAA→AAG)	NAD(P) transhydrogenase subunit alpha
1,701,047	G→A	C578Y (TGT→TAT)	fusaric acid resistance protein
1,737,579	T→C	N71S (AAC→AGC)	cysteine desulfurase
1,751,877	T→C	F275L (TTC→CTC)	quinate/shikimate dehydrogenase
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,873,224	G→A	S29S (AGC→AGT)	hypothetical protein
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
1,937,938	(C)6→7	coding (99/747 nt)	copper homeostasis protein CutC
1,964,675	(A)5→4	intergenic (-553/-244)	L-arabinose-binding periplasmic protein/non-heme ferritin

2,295,712	G→A	R157C (CGC→TGC)	protein InaA
2,384,364	T→C	V278V (GTA→GTG)	3-oxoacyl-ACP synthase I
2,405,937	G→A	W323* (TGG→TAG)	long-chain fatty acid transporter
2,478,993	G→A	P85S (CCT→TCT)	membrane protein
2,566,577	G→A	A523V (GCG→GTG)	penicillin-binding protein 1C
2,578,205	G→A	A477V (GCC→GTC)	molecular chaperone HscA
2,696,146	A→G	Y16C (TAC→TGC)	ribonucleotide-diphosphate reductase subunit alpha
2,731,583	C→T	A235T (GCG→ACG)	carbamoyltransferase HypF
2,735,430	G→A	A349A (GCG→GCA)	PTS cellobiose/arbutin/salicin transporter subunit IIBC
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,818,291	C→T	G378G (GGC→GGT)	serine transporter
2,836,432	C→T	R252H (CGT→CAT)	glycine cleavage system transcriptional activator
2,950,951	A→G	A133A (GCA→GCG)	propionyl-CoA--succinate CoA transferase
3,095,865	G→A	T326I (ACC→ATC)	cell division protein FtsP
3,118,561	T→C	G26G (GGA→GGG)	3,4-dihydroxy-2-butanone-4-phosphate synthase
3,133,077	T→C	E241E (GAA→GAG)	inorganic triphosphatase
3,301,015	T→C	T154A (ACA→GCA)	N-acetylneuraminate lyase
3,336,303	T→C	V269A (GTA→GCA)	sodium:pantothenate symporter
3,426,297	G→A	A189T (GCG→ACG)	siroheme synthase
3,700,548	G→A	G1224G (GGG→GGA)	RHS element protein RhsA
3,875,773	T→C	intergenic (-207/+146)	bifunctional N-acetylglucosamine-1-phosphate uridyltransferase/glucosamine-1-phosphate acetyltransferase/ATP synthase epsilon chain
3,893,551	+G	intergenic (+6/-50)	low affinity potassium transport system protein kup/transposase
3,923,132	C→T	D218D (GAC→GAT)	ATP-dependent DNA helicase Rep
4,179,767	(C)6→7	intergenic (+61/-126)	hypothetical protein/DNA-binding protein HU-alpha
4,372,135	C→T	D30N (GAT→AAT)	tRNA epoxyqueuosine(34) reductase QueG
4,480,205	C→T	V76V (GTC→GTT)	integrase
4,614,357	(T)5→4	coding (232/1383 nt)	DNA repair protein RadA

Table: Breseq output for PS3870

20 mutations

Ancestor: PS174 Ara- ancestral clone

Mutations

position	Mutation	Annotation	Description
1,018,516	C→T	A737V (GCC→GTC)	outer membrane usher protein
1,380,772	T→C	L153L (TTG→CTG)	outer membrane protein G
1,808,815	T→C	T145A (ACT→GCT)	acetylornithine aminotransferase
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,914,652	T→C	D80G (GAC→GGC)	glucose-6-phosphate dehydrogenase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
2,123,452	A→G	T40A (ACC→GCC)	lipid kinase YegS
2,690,612	(A)5→4	coding (186/663 nt)	transcriptional regulator
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,818,291	C→T	G378G (GGC→GGT)	serine transporter
3,156,827	T→C	Y460H (TAC→CAC)	beta-galactosidase subunit alpha
3,281,332	T→C	Y461H (TAC→CAC)	RNA polymerase sigma-54 factor
3,283,514	(C)6→7	intergenic (+41/-173)	phosphocarrier protein NPr/hypothetical protein
3,284,281	G→A	E199K (GAG→AAG)	hypothetical protein
3,567,142	T→C	T320A (ACA→GCA)	membrane protein
3,937,495	A→G	T184A (ACG→GCG)	dTDP-4-amino-4,6-dideoxygalactose aminotransferase

Table: Breseq output for PS3869

21 mutations Ancestor: PS2717

Ara+ ancestral clone

Mutations position	Mutation	Annotation	Description
70,867	T→C	D92G (GAC→GGC)	L-arabinose isomerase
272,553	C→T	G69S (GGC→AGC)	aldehyde dehydrogenase iron-sulfur subunit
278,454	(C)5→6	coding (1061/2526 nt)	hypothetical protein
1,111,966	(G)7→8	intergenic (-244/-64)	transposase/hypothetical protein
1,349,565	(G)7→8	intergenic (-157/+211)	enoyl-ACP reductase/hypothetical protein
1,559,052	(C)6→7	coding (1386/2280 nt)	hypothetical protein
1,829,717	G→T	T201T (ACG→ACT)	bifunctional pyrazinamidase/nicotinamidase
1,930,637	T→G	V158G (GTG→GGG)	hypothetical protein
1,953,113	G→A	P326S (CCG→TCG)	chemotaxis protein CheA
2,285,026	A→G	R365R (CGT→CGC)	DNA gyrase subunit A
2,295,712	G→A	R157C (CGC→TGC)	protein InaA
2,752,473	+G	coding (521/2562 nt)	DNA mismatch repair protein MutS
2,776,775	C→T	P13P (CCG→CCA)	phosphoadenosine phosphosulfate reductase
2,801,595	G→A	P329S (CCG→TCG)	GTP pyrophosphokinase
2,807,931	(C)5→6	coding (384/1341 nt)	glucarate dehydratase
2,815,688	T→C	V151A (GTG→GCG)	LOG family protein YgdH
2,818,291	C→T	G378G (GGC→GGT)	serine transporter
2,840,684	C→T	intergenic (-223/+16)	tRNA threonylcarbamoyladenosine dehydratase/murein transglycosylase A
3,095,865	G→A	T326I (ACC→ATC)	cell division protein FtsP
3,141,891	G→A	A351T (GCC→ACC)	antiporter
3,368,349	T→C	T196A (ACC→GCC)	DNA-directed RNA polymerase subunit alpha

BIBLIOGRAPHY

Anderson J, Daifuku R, Loeb L. Viral Error Catastrophe by Mutagenic Nucleosides. Annual Review of Microbiology. 2004;58:183-205

Andersson DI and Hughes D. Muller's ratchet decreases fitness of a DNA-based microbe. Proc Natl Acad Sci USA. 1996;93(2):906-7.

André J-B, Godelle B. The evolution of mutation rate in finite asexual populations. Genetics. 2006;172(1):611–26. DOI:10.1534/genetics.105.046680

Baquero F, Negri MC, Morosini MI, Blázquez J. The antibiotic selective process: concentration-specific amplification of low-level resistant populations. Ciba Found Symp. 1997;207:93-111.

Baquero MR, Nilsson A, Turrientes M, Sandvang D, Galán JC, Martínez JL, et al. Polymorphic mutation frequencies in *Escherichia coli*: emergence of weak mutators in clinical isolates. J Bacteriol. 2004;186(16):5538-42.

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, et al. Genome evolution and adaptation in a long-term experiment with *Escherichia coli.* Nature. 2009;461:1243–7. DOI:10.1038/nature08480

Belshaw R, Gardner A, Rambaut A, Pybus OG. Pacing a small cage: mutation and RNA viruses. Trends Ecol Evol. 2007;23(4):188-93.

Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B. 1995;57(1):289-300.

Bernstein H, Byers GS, Michod RE. Evolution of sexual reproduction: importance of DNA repair, complementation, and variation. Am Nat. 1980;117:537-549

Bhatnagar S, Bessman M. Studies on the mutator gene, *mutT* of *Escherichia coli.* J Biol Chem. 1988;263(18):8953–7.

Billington OJ, McHugh TD, Gillespie SH. Physiological cost of rifampin resistance induced in vitro in *Mycobacterium tuberculosis.* Antimicrob Agents Chemother. 1999;43(8):1866-9.

Blount Z, Borland C, and Lenski R. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli.* PNAS. 2008;105(23):7899-7906. doi: 10.1073/pnas.0803151105

Bridges BA. Elevated Mutation Rate in *mutT* Bacteria during Starvation: Evidence for DNA Turnover? J Bacteriol. 1996;178(9):2709–2711

Bromham L, Hua X, Lanfear R, Cowman P. Exploring the Relationships between Mutation Rates, Life History, Genome Size, Environment, and Species Richness in Flowering Plants. Am Nat. 2014;185(4):507-24. doi: 10.1086/680052.

Bull JJ, WIlke CO. Lethal mutagenesis of bacteria. Genetics. 2008;180(2):1061-70. doi: 10.1534/genetics.108.091413.

Carlton BC and Brown BJ. Gene mutation. Manual of methods for general bacteriology. American Society for Microbiology, Washington, DC. 1981. 222-242.

Chang C-H, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. 2015. Cell. 2015;162(6):1229-1241.

Chao L and Cox EC. Competition between high and low mutating strains of *Escherichia coli*. Evolution. 1983;37:125–134

Conrad DF, Keebler JE, DePristo MA, Lindsay SJ, Zhang Y, Casals F, et al. Variation in genome-wide mutation rates within and between human families. Nat Genet. 2011;43(7):712-4.

Connolly DM and Winkler ME. Genetic and physiological relationships among the *miaA* gene, 2-methylthio-N6-(delta 2-isopentenyl)-adenosine tRNA modification, and spontaneous mutagenesis in *Escherichia coli* K-12. J Bacteriol. 1989;171(6):3233-46.

Crotty S, Cameron C, Andino R. RNA virus error catastrophe: Direct molecular test by using ribavirin. PNAS. 2001;98(12): 6895–6900.

Crotty S, Andino R. Implications of high RNA virus mutation rates: lethal mutagenesis and the antiviral drug ribavirin. Microbes Infect. 2002;4(13):1301-7.

Crow JF and Kimura M. An introduction to population genetics theory. 1970. New York, NY: Harper & Row.

Deatherage D, Barrick J. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol Biol. 2014;1151:165–88. DOI:10.1007/978-1- 4939-0554- 6_12

Denamur E, Bonacorsi S, Giraud A, Duriez P, Hilali F, Amorin C, et al. High frequency of mutator strains among human uropathogenic *Escherichia coli* isolates. J Bacteriol. 2002;184(2):605-9.

Denamur E, Matic I. Evolution of mutation rates in bacteria. Molecular Microbiology. 2006;60(4):820-827.

de Visser JA. The fate of microbial mutators. Microbiology. 2002;148(5):1247-52.

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

Drake JW. Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci USA. 1993;90(9):4171-5.

Drake J, Holland J. Mutation rates among RNA viruses. PNAS. 1999;96(24):13910–13913. DOI:10.1073/pnas.96.24.13910

Duchene D and Bromham L. Rates of molecular evolution and diversification in plants: chloroplast substitution rates correlate with species-richness in the Proteaceae. BMC Evol Biol. 2013;13:65. doi: 10.1186/1471-2148-13-65.

Eghbal MM, Sprouffske K, Dartey JO, Sniegowski PD. Mutator populations of *Escherichia coli* substitute additional, spontaneously originated mutator alleles under lethal selection. 2017. (In review.)

Eghbal MM, Sprouffske K, Dartey JO, Sniegowski PD. Data from: Mutator populations of *Escherichia coli* substitute additional, spontaneously originated mutator alleles under lethal selection. Dryad Digit Repos. 2017. DOI:10.5061/dryad.k76d4

Eigen M and Schuster P. The hypercycle: a principle of natural self-organization. Part A: Emergence of a hypercycle. Naturwissenschaften. 1997;64:541–565.

Eigen M. Error catastrophe and antiviral strategy. Proc Natl Acad Sci USA. 2002;99(21):133374-6.

Elena SF, Sanjuán R. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J Virol 2005;79(18):11555-8.

Feldman MW, Liberman U. An evolutionary reduction principle for genetic modifiers. Proc Natl Acad Sci U S A. 1986;83(13):4824-7.

Field D, Magnasco MO, Moxon ER, Metzgar D, Tanaka MM, Wills C, Thaler DS. Contingency Loci, Mutator Alleles, and Their Interactions: Synergistic Strategies for Microbial Evolution and Adaptation in Pathogenesis. Annals of the NY Academy of Sciences. 1999;870:378-381.

Fisher RA. The genetical theory of natural selection. Oxford: Clarendon Press; 1930. DOI:10.5962/bhl.title.27468

Fowler R, Schaaper R. The role of the *mutT* gene of *Escherichia coli* in maintaining replication fidelity. FEMS Microbiol Rev. 1997;21(1):43–54. DOI:10.1111/j.1574-6976.1997.tb00344.x

Fowler RG, White SJ, Koyama C, Moore SC, Dunn RL, Schaaper RM. Interactions among the *Escherichia coli mutT, mutM,* and *mutY* damage prevention pathways. DNA Repair. 2003;3(2):159-73.

Furió V, Moya A, Sanjuán R. The cost of replication fidelity in an RNA virus. Proc Natl Acad Sci USA. 2005;102(29):10233-7.

Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a stress response and the regulation of evolvability. Crit Rev Biochem Mol Biol. 2007;42(5):399-435.

Gallet R, Froissart R, Ravigné V. Hard and Soft Selection in the Laboratory: Dissecting the Mechanisms of Polymorphism Maintenance in Spatially Heterogeneous Environments. 2017. Preprint on bioRxiv.

Gentile CF, Yu S, Serrano SA, Gerrish PJ, Sniegowski PD. Competition between high- and higher-mutating strains of *Escherichia coli.* Biol Lett. 2011;7:422–4. DOI:10.1098/rsbl.2010.1036

Gerrish P. A simple formula for obtaining markedly improved mutation rate estimates. Genetics. 2008;180(3):1773–8. DOI:10.1534/genetics.108.091777

Gerrish PJ, Colato A, Perelson AS, Sniegowski PD. Complete genetic linkage can subvert natural selection. PNAS. 2007;104(15). DOI:10.1073/pnas.0607280104

Gerrish PJ, Colato A, Sniegowski PD. Genomic mutation rates that neutralize adaptive evolution and natural selection. J R Soc Interface. 2013;10(85).

Goldstein B. Resistance to rifampicin: a review. J Antibiot. 2014;67:625–30. DOI:10.1038/ja.2014.107

Gardo I, Campos P. Evolution of clonal populations approaching a fitness peak. Biol Lett. 2013;9(1).

Graci, J, Cameron C. Quasispecies, Error Catastrophe, and the Antiviral Activity of Ribavirin. Virology. 2002;298(2),175-180.

Graves C, Ros V, Stevenson B, Sniegowski P, Brisson D. Natural selection promotes antigenic evolvability. PLoS Pathog. 2013;9(11). DOI:10.1371/journal.ppat.1003766

Gross MD and Siegel EC. Incidence of mutator strains in *Escherichia coli* and coliforms in nature. Mutat. Res. 1981;91:107–110.

Hessen DO. Solar radiation and the evolution of life. Solar Radiation and Human Health. 2008;123-136.

Herr AJ, Kennedy SR, Knowels GM, Schultz EM, Preston BD. DNA Replication Error-Induced Extinction of Diploid Yeast. Genetics. 2014;196(3):677-691.

Holsinger K, Feldman M. Modifiers of mutation rate: Evolutionary optimum with complete selfing. PNAS. 1983;80:6732-6734.

Hughes D, Andersson DI. Selection of resistance at lethal and non-lethal antibiotic concentrations. Curr Opin Microbiol. 2012;15(5):555-60. doi: 10.1016/j.mib.2012.07.005.

Ishii K, Matsuda H, Iwasa Y, Sasaki A. Evolutionarily Stable Mutation Rate in a Periodically Changing Environment. Genetics. 1988;121:163-174.

Jeong H, Barbe V, Lee CH, Vallenet D, Yu DS, Choi S-H, et al. Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). J Mol Biol. 2009;394(4):644–52. DOI:10.1016/j.jmb.2009.09.052

Jiang X, Mu B, Huang Z, Zhang M, Wang X, Tao S. Impacts of mutation effects and population size on mutation rate in asexual populations: a simulation study. BMC Evo Bio. 2010;10:298.

Jin DJ, Gross C. Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. J Mol Biol. 1988;202(1):45–88. DOI:10.1016/0022-2836(88)90517-7

Jones S, Chen W-D, Parmigiani G, Diehl F, Beerenwinkel N, Antal T, et al. Comparative lesion sequencing provides insights into tumor evolution. PNAS. 2008;105:4283-4288.

Karam PA, Leslie SA. The evolution of the earth's background radiation level over geologic time. International Congress on Radiation Protection Proceedings. 1996;28(3).

Karran P. Appropriate Partners Make Good Matches. Science. 1995;268:1857-1858.

Kimura M. On the evolutionary adjustment of spontaneous mutation rates. Genetics. 1967;9(1):23-24.

Kinnersley M, Wenger J, Kroll E, Adams J, Sherlock G, Rosenzweig F. Ex uno plures: Clonal reinforcement drives evolution of a simple microbial community. PLoS Genet. 2014;10(6). DOI:10.1371/journal.pgen.1004430.g001

Kondrashov AS. Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? J Theor Biol. 1995;175(4):583-94.

Lanfear R, Ho SYW, Love D, Bromham L. Mutation rate is linked to diversification in birds. PNAS. 2010;107(47).

LeClerc JE, Li B, Payne W, Cebula T. High Mutation Frequencies Among *Escherichia coli* and *Salmonella* Pathogens. Science. 1996;274(5290):1208–11. DOI:10.1126/science.274.5290.1208

Lee H, Popodi E, Tang H, Foster PL. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. PNAS. 2012;109(41).

Lenski RR and Hattingh DE. Coexistence of two competitors on one resource and one inhibitor: a chemostat model based on bacteria and antibiotics. Journal of Theoretical Biology. 1986;122:83-93.

Lenski R, Travisano M. Dynamics of Adaptation and Diversification: A 10,000 Generation Experiment with Bacterial Populations. PNAS. 1994;91:6808–6814.

Lenski RE. The future of evolutionary biology. Ludus Vitalis. 2004;12:67-89.

Levinson G, Gutman GA. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol. 1987;4(3):203-21.

Lieberman TD, Michel J-B, Aingaran M, Potter-Bynoe G, Roux D, Davis MR, et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. Nature Genetics. 2011;43:1275-1280.

Li Z, Woo CJ, Iglesias-Ussel, Ronai D, Scharff MD. The generation of antibody diversity through somatic hypermutation and class switch recombination. Genes & Dev. 2004;18:1-11.

Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. Cancer Res. 1991;51(12):3075-9.

Loeb K, Loeb L. Significance of multiple mutations in cancer. Carcinogenesis. 2000;21(3):379–85. DOI:10.1093/carcin/21.3.379

Luria SE, Delbrück M. Mutations of bacteria from virus sensitivity to virus resistance.

Genetics. 1943;28(6):491–511.

Lynch M. The cellular, developmental, and population-genetic determinants of mutation-rate evolution. Genetics. 2008;180: 933-943.

Lynch M. Evolution of the Mutation Rate. Trends in Genetics. 2010;26:345-352.

Lynch M. The Lower Bound to the Evolution of Mutation Rates. Genome Biol Evol. 2011;3:1107-1118.

Lynch M, Ackerman M, Gout J-F, Long H, Sung W, Thomas WK, Foster PL. Genetic drift, selection, and the evolution of the mutation rate. Nature Reviews Genetics. 2016;17:704-714.

Mao EF, Lane L, Lee J, Miller JH. Proliferation of mutators in a cell population. J Bacteriol. 1997;179(2):417–22. DOI:10.1128/jb.179.2.417-422.1997

Maki H, Sekiguchi M. MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. Nature. 1992;355:273–5. DOI: 10.1038/355273a0

Martin A and Scharff MD. AID and mismatch repair in antibody diversification. Nature Reviews Immunology. 2002;2(8):605-614.

Maynard Smith J. The Evolution of Sex. 1978. Cambridge University Press.

McDonald MJ, Hsieh YY, Yu YH, Chang SL, Leu JY. The evolution of low mutation rates in experimental mutator populations of *Saccharomyces cerevisiae.* Curr Biol. 2012;22(13):1235-40.

Michaels ML, Cruz C, Miller JH. mutA and mutC: two mutator loci in Escherichia coli that stimulate transversions. PNAS. 1990;87(23):9211-5.

Miller JH. Experiments in molecular genetics. Cold Spring Harbor Laboratory. 1972.

Modrich P. Mismatch repair, genetic stability and tumour avoidance. Phil Trans Roy Soc Lond Ser B, 1996;347:89-95.

Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. Genetics 2000;156(1):297-304.

Nakamura S, Nakamura M, Kojima T, Yoshida H. *gyrA* and *gyrB* mutations in quinolone-resistant strains of *Escherichia coli*. Antimicrob Agents Chemother. 1989;33(2):254–5. DOI:10.1128/aac.33.2.254

Nguyen V, Wilson C, Hoemberger M, Stiller JB, Agafonov RV, Kutter S, et al.

Evolutionary drivers of thermoadaptation in enzyme catalysis. Science. 2017;355:289–294.

Nilsson AI, Berg OG, Aspevall O, Kahlmeter G, Andersson DI. Biological costs and mechanisms of fosfomycin resistance in *Escherichia coli*. Antimicrob Agents Chemother. 2003;47(9):2850–8. DOI:10.1128/AAC.47.9.2850-2858.2003

Nowell PC. The clonal evolution of tumor cell populations. Science. 1976;194:23–28.

Nucleotide. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; Accession No. NC_000913, *Escherichia coli* str. K-12 substr. MG1655, dGTP-preferring nucleoside triphosphate pyrophosphohydrolase. Available from: https://www.ncbi.nlm.nih.gov/nuccore/NC_000913.3?report=genbank&from=111044&to=111433

Oliver A and Mena A. Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. Clinical Microbiology and Infection. 2010;16(7):798-808.

Pal C, Maciá M, Oliver A, Schachar I, Buckling A. Coevolution with viruses drives the evolution of bacterial mutation rates. Nature. 2007;450:1079–81. DOI:10.1038/nature06350

Palmer M, Lipsitch M. The Influence of Hitchhiking and Deleterious Mutation Upon Asexual Mutation Rates. Genetics. 2006;173(1):461-472.

Papadopoulos D, Schneider D, Meier-Eiss J, Arber W, Lenski RE, Blot M. Genomic evolution during a 10,000-generation experiment with bacteria. Proceedings of the National Academy of Sciences. 1999;96(7):3807-3812.

Pope CF, O'Sullivan DM, McHugh TD, Gillespie SH. A practical guide to measuring mutation rates in antibiotic resistance. Antimicrob Agent Chemother. 2008;52(4):1209-14.

Raynes Y, Gazzara MR, Sniegowski PD. Mutator dynamics in sexual and asexual experimental populations of yeast. BMC Evolutionary Biology. 2011;11:158. DOI: 10.1186/1471-2148-11-158

Raynes Y, Halstead AL, Sniegowski PD. The effect of population bottlenecks on mutation rate evolution in asexual populations. J Evol Biol. 2014;27(1):161-9.

Raynes Y, Sniegowski PD. Experimental evolution and the dynamics of genomic mutation rate modifiers. Heredity. 2014;113:375–80. DOI:10.1038/hdy.2014.49

Ribeiro RM, Li H, Wang S, Stoddard MB, Learn GH, Korber BT, et al. Quantifying the Diversification of Hepatitis C Virus (HCV) during Primary Infection: Estimates of the In Vivo Mutation Rate. Plos Pathogens. 2012.

Rice WR. Analyzing Tables of Statistical Tests. Evolution. 1989;43(1):223-225.

Rocha EP, Matic I, Taddei F. Over-representation of repeats in stress response genes: a strategy to increase versatility under stressful conditions? Nucleic Acids Res. 2002;30(9):1886-94.

Rozen DE and Lenski RE. Long-term experimental evolution in *Escherichia coli.* VIII. Dynamics of a balanced polymorphism. American Naturalist. 2000;155:24-35.

Saenz Y, Zarazaga M, Brinas L, Ruiz-Larrea F, Torres C. Mutations in *gyrA* and *parC* genes in nalidixic acid-resistant *Escherichia coli* strains from food products, humans and animals. J Antimicrob Chemother. 2003;51(4):1001–5. DOI:10.1093/jac/dkg168

Sandegren L. Selection of antibiotic resistance at very low antibiotic concentrations. Ups J Med Sci. 2014;119(2):103–107.

Sawabe TS, Koizumi S, Fukui Y, Nakagawa S, Ivanova EP, Kita-Tsukamoto K, et al. Mutation is the main driving force in the diversification of the *Vibrio splendidus* clade. Microbes Environ. 2009;24(4):281-285.

Schaaff F, Reipert A, Bierbaum G. An elevated mutation frequency favors development of vancomycin resistance in Staphylococcus aureus. Antimicrob Agents Chemother. 2002;46(11):3540-8.

Schaaper RM, Bond BI, Fowler RG. A.T----C.G transversions and their prevention by the *Escherichia coli mutT* and *mutHLS* pathways. Mol Gen Genet. 1989;219(1-2):256-62.

Shaver AC, Dombrowski PG, Sweeney JY, Treis T, Zappala RM, Sniegowski PD. Fitness evolution and the rise of mutator alleles in experimental *Escherichia coli* populations. Genetics. 2002;162:557–566.

Shaver A, Sniegowski P. Spontaneously arising *mutL* mutators in evolving *Escherichia coli* populations are the result of changes in repeat length. J Bacteriol. 2003;185(20):6076–6082. DOI: 10.1128/JB.185.20.6076-6082.2003

Shibata DK, Lieber MR. Is There Any Genetic Instability in Human Cancer? DNA Repair. 2010;9(8):858–860.

Singh T, Hyun M, Sniegowski P. Evolution of mutation rates in hypermutable populations of *Escherichia coli* propagated at very small effective population size. Biol Lett. 2017;13(3).

Šmarda P, Bureša P, Horová L, Leitch IJ, Mucina L, Pacini E, et al. Ecological and evolutionary significance of genomic GC content diversity in monocots. PNAS. 2014;111(39).

Sniegowski PD, Gerrish PJ, Lenski RE. Evolution of high mutation rates in experimental populations of Escherichia coli. Nature. 1997;387:703-705.

Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. The evolution of mutation rates: separating causes from consequences. BioEssays. 2000;22(12):1057-1066.

Sniegowski P, Raynes Y. Mutation rates: how low can you go? Curr Biol. 2013;23(4).

Söderberg RJ and Berg OG. Kick-starting the ratchet: the fate of mutators in an asexual population. Genetics. 2011;187(4);1128-37.

Solé R, Deisboeck T. An error catastrophe in cancer? J Theor Biol. 2004;228(1):47–54. DOI:10.1016/j.jtbi.2003.08.018

Springer B, Kidan Y, Prammananan T, Ellrott K, Böttger E, Sander P. Mechanisms of streptomycin resistance: selection of mutations in the 16S rRNA gene conferring resistance. Antimicrob Agents Chemother. 2001;45(10):2877–2884. DOI:10.1128/AAC.45.10.2877-2884.2001

Sprouffske K, Merlo LMF, Gerrish PJ, Maley CC, Sniegowski PD. Cancer in light of experimental evolution. Curr Biol. 2012;22:762–71. DOI:10.1016/j.cub.2012.06.065

Sturtevant A. Essays on evolution. I. On the effects of selection on mutation rate. Q Rev Biol. 1937;12:464–7. DOI:10.1086/394543

Sung W, Ackerman M, Miller S, Doak T, Lynch M. Drift-barrier hypothesis and mutation-rate evolution. PNAS. 2012;109(45):18488–18492. doi: 10.1073/pnas.1216223109

Tadesse S, Graumann PL. DprA/Smf protein localizes at the DNA uptake machinery in competent *Bacillus subtilis* cells. BMC Microbiol. 2007;7:105.

Tan K, Zhou Q, Cheng B, Zhang Z, Joachimiak A, Tse-Dinh YC. Structural basis for suppression of hypernegative DNA supercoiling by *E. coli* topoisomerase I. Nucleic Acids Res. 2015;43(22):11031-46.

Tejero H, Montero F, Nuño JC. Theories of Lethal Mutagenesis: From Error Catastrophe to Lethal Defection. Curr Top Microbiol Immunol. 2016;392:161-79. doi: 10.1007/82_2015_463.

Tenaillon O, Toupance B, Le Nagard H, Taddei F, Godelle B. Mutators, Population Size, Adaptive Landscape and the Adaptation of Asexual Populations of Bacteria. Genetics. 1999;152(2):485-493.

Tenaillon O, Le Nagard H, Godelle B, Taddei F. Mutators and sex in bacteria: Conflict between adaptive strategies. 2000;97(19):10465-10470.

Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, et al. Tempo and mode of genome evolution in a 50,000-generation experiment. Nature. 2016;536:165-170.

Thi T Do, López E, Rodríguez-Rojas A, Rodríguez-Beltrán J, Couce A, Guelfo J, et al. Effect of *recA* inactivation on mutagenesis of *Escherichia coli* exposed to sublethal concentrations of antimicrobials. J Antimicrob Chemother. 2011;66(3):531–8. DOI:10.1093/jac/dkq496

Timms A, Steingrimsdottir H, Lehmann A, Bridges B. Mutant sequences in the *rpsL* gene of *Escherichia coli* B/r: Mechanistic implications for spontaneous and ultraviolet light mutagenesis. Mol Genet Genomics. 1992;232(1):89–96. DOI:10.1007/BF00299141

Tomlinson IP, Novelli MR, Bodmer WF. The mutation rate and cancer. PNAS. 1996;93(25).

Torres-Barceló C, Cabot G, Oliver A, Buckling A, Maclean RC. A trade-off between oxidative stress resistance and DNA repair plays a role in the evolution of elevated mutation rates in bacteria. Proc Biol Sci. 2013;280(1757).

Tröbner W, Piechocki R. Selection against hypermutability in Escherichia coli during long term evolution. Mol Gen Genet. 1984;198(2):177-8.

Tsui HC, Feng G, Winkler ME. Transcription of the *mutL* repair, *miaA* tRNA modification, *hfq* pleiotropic regulator, and *hflA* region protease genes of *Escherichia coli* K-12 from clustered Esigma32-specific promoters during heat shock. J Bacteriol. 1996;178(19):5719-31.

Turrientes MC, Baquero F, Levin BR, Martínez JL, Ripoll A, González-Alba JM, et al. Normal mutation rate variants arise in a mutator (*mutS*$^-$) *Escherichia coli* population. PLoS One. 2013;8(9).

Umeno M, Sasaki M, Anai M, and Takagi Y. Purification and Subunit Structure of recBC DNase from *Escherichia coli* Harboring a *recB* and *recC* Genes-Inserted Plasmid. J Biochem. 1985;98(3):681-685.

Vacher J, Grosjean H, Houssier C, Buckingham RH. The effect of point mutations affecting *Escherichia coli* tryptophan tRNA on anticodon-anticodon interactions and on UGA suppression. J Mol Biol. 1984;177(2):329-42.

Ventola CL. The antibiotic resistance crisis: Part 1. Pharmacy and Therapeutics. 2015;40(4):277-283.

Vignuzzi M, Stone JK, Andino R. Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. Virus Res. 2005. 107(2):173-81.

Viguera E, Canceill D, Ehrlich SD. Replication slippage involves DNA polymerase pausing and dissociation. EMBO J. 2001;20(10):2587-2595.

Wang Y, Arenas CD, Stoebel DM, Cooper TF. Genetic background affects epistatic interactions between two beneficial mutations. Biology Letters. 2012;9(1).

Waygood EB, Reiche B, Hengstenberg W, Lee J. Characterization of mutant histidine-containing proteins of the phosphoenolpyruvate:sugar phosphotransferase system of *Escherichia coli* and *Salmonella typhimurium.* J Bacteriol. 1987;169(6):2810–8. DOI:10.1128/jb.169.6.2810-2818.1987

Whitlock MC, Davis B. Genetic load. eLS. 2011. DOI: 10.1002/9780470015902.a0001787.pub2

Wielgoss S, Barrick JE, Tenaillon O, Wiser MJ, Dittmar WJ, Cruveiller S, et al. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. PNAS. 2013;110(1):222-7.

Xue Y, Wang Q, Long Q, Ng BL, Swerdlow H, Burton J, et al. Human Y Chromosome Base-Substitution Mutation Rate Measured by Direct Sequencing in a Deep-Rooting Pedigree. Current Biology. 2009;19(17):1453-1457.

Zheng Q. Methods for comparing mutation rates using fluctuation assay data. Mutat Res. 2015;777:20–2. DOI:10.1016/j.mrfmmm.2015.04.002

Zhu YO, Siegal ML, Hall DW, Petrov DA. Precise estimates of mutation rate and spectrum in yeast. PNAS. 2014;111(22):E2310–E2318. doi: 10.1073/pnas.1323011111