Adaptation in *Escherichia coli*: ecological and genetic constraints on diversification

by

Alana Schick

B.Sc., The University of British Columbia, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Zoology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

January 2014

Abstract

There is growing evidence that disruptive selection generated by intraspecific resource competition may be a common mechanism for generating biological diversity. Adaptive dynamics models provide a framework describing how frequency dependent selection drives such diversification, but these models don't consider the complexities that arise as a result of gene interactions. Here, we explore the relative effects of ecological and genetic constraints on diversification using an experimental system of *Escherichia coli* in which diversification is driven by frequency dependence based on resource use. Diversified populations consist of ecotypes that consume glucose and acetate at different rates, and a mutation in the *arcA* gene has been identified that has a large effect on this phenotype. By isolating clones of each ecotype from a previously diversified population, we find that the effect of the *arcA* mutation on rediversification depends on both the ecotype and the genetic background. While some of these observations are consistent with predictions made by adaptive dynamics models, others cannot be explained without also accounting for epistasis and genetic constraints, highlighting the importance of considering both ecological and genetic factors when predicting diversification. Adaptation in this system also provides an example of an interaction between ecological and evolutionary processes, adding to a growing number of studies that exhibit a clear feedback between these two processes.

Preface

- Chapter 3 is based on work completed in UBC's Biodiversity Research Centre, in the lab of Professor Michael Doebeli. I designed the experiments, with the advice of Dr. Mickael Le Gac. I was responsible for carrying out the experiments described here, conducted and/or supervised all media preparation, bacterial transfers, and data collection.

- A version of Chapter 3 has been submitted for publication. Schick, A., and Doebeli, M. 2013. Adaptation in *Escherichia coli*: Intraspecific ecological interactions and genetic constraints determine diversification. I conducted all the experiments, statistical analysis and wrote the manuscript.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

Completing this thesis would not have been possible without the support of numerous people. First and foremost, I would like to thank my advisor, Michael Doebeli, for his encouragement and assistance over the years. Thanks to members past and present of the Doebeli and Otto lab groups for discussions and feedback. Thanks to Thomas Prasloski and Sarita Balabhadra for help collecting data. Thanks to Mickael Le Gac for help designing the experiments and providing the strains. Thanks to my committee members, Mike Whitlock and Leticia Aviles for guidance. Lastly, special thanks to Aleeza Gerstein and Sam Yeaman for helping make sense of it all (data and life included).

Chapter 1

General Introduction

"Nothing in evolution or ecology makes sense except in the light of the other."
[40]

As evolutionary biologists, our attempts to understand adaptive evolution are driven largely by the desire for foresight; to accurately predict the future of populations. For a single population in an isolated and controlled environment (like an evolutionary vacuum), this might be possible. In nature, however, there is no such thing. We are limited by complexity due to an essentially unlimited number of interdependent variables. For this reason, the frontier of evolutionary biology is fundamentally different than the frontier of other sciences. There are some evolutionary biologists who claim that all the interesting questions in evolution have been answered. In some ways, they have. Certainly many (if not most) of the mechanisms that drive evolution are well understood. The limitation of biology, however, is not a limitation in kind (like in many other sciences), but a limitation in scale. In other words, we may know what the mechanisms are, but our abilities to make predictions are still weak due to the infinite interactions between things. These limitations are being broken through. Data is pouring in at a rate faster than ever, especially at the molecular level, making this a most exciting time to be a biologist. In my opinion, the discoveries that advance evolutionary biology in the next hundred years will not be discoveries of unknown processes or mechanisms, but the prevalence and relative importance of specific mechanisms.

For a long time, evolutionary biologists have made their lives simpler by assuming that ecological and evolutionary processes occur on different time scales and can therefore be treated independently. This has led to a dichotomy in researchers of ecologists and evolutionary biologists. It is becoming increasingly clear, however, that these processes do occur on the same time scale and that one cannot be fully understood without the other. In a recent and frequently cited review of eco-evolutionary dynamics, Schoener (2011) states that evidence for ecological change affecting evolution is abundant, but evidence of the reverse is sparse. I suspect that this is due in

large part to the complexity and difficulty of mechanistically linking evolutionary and ecological dynamics. For example, it may be easier to see how species interactions drive changes in allele frequencies, but harder to see how changes in allele frequencies drive species interactions. In bacterial populations, however, ecology is much more simple, making it possible to link these processes mechanistically. The work presented in this thesis adds to a growing number of studies that show the feedback between ecological and evolutionary processes on the same time scale.

Though advances in genome sequencing technologies are drastically increasing the information known about the mutations involved in adaptation, this information is meaningless without ecological context. In other words, the effect of a given mutation depends on the genetic background in which it arises, but it also depends on the genetic makeup of the population in which it arises like the frequencies of alleles involved in resource consumption, for example. Many studies have looked at the effects of mutations on phenotype and fitness, sometimes including how this effect depends on environment and genetic background, but few studies have investigated the effect of an adaptive mutation on subsequent adaptation or diversification. Here, I investigate the role of a single mutation affecting resource consumption on subsequent diversification. Since this mutation alters the rate at which resources are consumed in an environment in which individuals compete for resources, there is a direct link between the frequency of this allele and the selective environment. In other words, the resources available dictate the selective pressure on the types present, and in turn, the types present determine which resources are available.

Diversification in resource exploitation is common in microbes. In this work, *Escherichia coli* adapts to a two-resource environment by diversifying in metabolic strategy. Interestingly, diversification is also observed in environments supplemented with only one resource ([46]), often due to cross-feeding mechanisms. In these cases, at least one of the other types evolves to consume a metabolic byproduct of the other type. In this situation, it is easy to see that these types would be maintained by negative frequency-dependent selection. Because of how common this type of selection seems to be, insights from the experiment presented here could be applied to a variety of other organisms.

This thesis contains two core chapters. The next chapter provides the background for the main body of the thesis. This part begins with a summary of the contributions of microbial evolution experiments to understanding adaptation, followed by a brief description of the experiments that preceded the current experiment, which is described in the last chapter. One

main finding of this work is that a mutation with no effect on phenotype has a profound effect on subsequent adaptation, which has important consequences for predicting evolutionary trajectories.

Chapter 2

Background

2.1 Generalizations from evolution experiments

Since there already exist many detailed reviews of the merits and limitations of laboratory experimental evolution (including [15], [19], and [27] for example), I'm not going to provide another one here. Instead, I will briefly summarize the light that these experiments have shed on our general understanding and predicting of adaptation. Though evolution experiments have informed a wide array of topics, many key results fall into three very broad categories: repeatability, epistasis, and the rate of adaptation. In short, we have learned that adaptation is somewhat repeatable, epistasis is rampant, and it seems like the rate of adaptation decreases over time after introduction to a new environment, but perhaps not.

2.1.1 Repeatability

One question that evolutionary biologists often hear is 'how repeatable is adaptation?' Any understanding of this question informs the ultimate question in evolutionary biology: can we predict evolutionary change? In other words, how much of the adaptation we observe is constrained to follow a certain path and how much is due to stochasticity? This question can be (and has been) informed by studies of the parallelism of adaptation, in which multiple populations are allowed to evolve and adapt to a given environment or a set of environments and the outcomes are compared. Due to the ease of growing microbial populations in a controlled laboratory environment, many of the studies that address this question have been conducted with microbes (including [32], [51], [47], and [39]). When adaptation is compared across populations, this question of parallelism can be applied to either genotypes or phenotypes, as well as the mapping between them. Decades of observations of natural populations have demonstrated that parallel phenotypic evolution is common (including well known examples such as anolis lizard ecomorhps ([33]) and threespine stickleback traits ([48]), but it is only recently that evolutionary biologists have been able to address

whether the genetic changes underlying that phenotypic parallelism is also parallel. Microbial evolution experiments are especially well suited to answer this question because the 'history of evolution' can be controlled by initializing populations that are identical genetically, removing any contingency that adaptation might have on previous adaptation. So far, microbial studies that compare the mutations that arise during adaptation between replicate populations (i.e. a measure of genetic parallelism) have revealed that there is a mix of shared and unique mutations. This includes studies of adaptation to novel nutrient environments ([25], [37], and reviewed in more detail in [11]), antibiotics ([17]), and even mutations underlying cancer progression ([23]). In many of these studies, it is unclear which of the mutations are adaptive and which are neutral due to the arduous task of investigating the effects of single mutations on phenotypes. Given what we know about mapping genotypes to phenotypes (in particular that many genotypes result in the same phenotype), these results should not be surprising. What remains to be seen is what factors determine how parallel the mutations underlying adaptation are, and with the ever-increasing ease of acquiring sequence data, these factors will begin to be understood.

2.1.2 Epistasis

The second major obstacle in predicting adaptation is the prevalence of epistasis. Because of epistasis, evolutionary biologists can not simply use a distribution of fitness effects of specific mutations to predict which ones will increase in frequency over the course of adaptation. An interaction between gene products means that many of the predicted fitness effects only apply to a single mutational step. There exists much empirical evidence showing that the fitness effect of a mutation (during adaptation) is strongly dependent on the genetic background in which is arises (such as [45], [9], [31], [35], and [41], for example). Due to complex interactions between gene products, predicting the effects of any mutation on phenotype becomes next to impossible. Microbial evolution experiments have begun to investigate how different combinations of adaptive mutations affect fitness (including [7], [28], and [61]), showing that understanding epistasis is a crucial component predicting evolutionary trajectories. Since many of these studies have investigated the interactions between mutations that rose to high frequencies under strong selection, they may be missing important interactions between mutations that arose but were selected against because of negative epistasis. This issue is beginning to be addressed with mutation accumulation experiments, which remove clonal interference from the population by reducing

5

the population size to one individual at regular intervals. The idea behind this is to minimize selection so that spontaneous mutations can be studied without a bias towards those that are beneficial and against deleterious ones (reviewed in [24]). So far, these mutation accumulation experiments and studies of the fitness effects of different combinations of mutations have shown that epistasis is rampant and can dramatically alter evolutionary trajectories, making it very important for predicting adaptation. How exactly epistasis impacts adaptation (for example by accelerating or slowing adaptation) has been informed by these evolution experiments in microbes, but the larger-scale experiments that are now becoming possible will provide a much more general understanding of the role of epistasis in predicting evolution.

2.1.3   Rate of adaptation

Due to many convenient properties of microbes, evolutionary biologists have been able to answer questions about the rate of adaptation using experimental evolution. In particular, because populations can be frozen and stored for long periods of time, direct comparisons can be made between ancestral and evolved populations or individuals. Most studies that seek to characterize how quickly populations adapt to novel environments find that the largest gains in fitness occur upon introduction to the novel environment, and decrease over time ([5], [3]). This is consistent with the theory that during an 'adaptive walk' up a fitness peak, it becomes increasingly difficult to climb as the population ascends the peak, because mutations in the direction of the optimum become less common ([43]). One of the major problems with using comparisons between ancestral and evolved populations to quantify the rate of adaptation is in defining and measuring fitness. Often, these studies use growth rate relative to the ancestor when grown together as the metric for fitness, but that does not necessarily simulate the environment in which mutations arose that increased fitness. For example, the observations of decreasing fitness gains may only be measuring one aspect of fitness and other positive fitness effects may go unnoticed. Furthermore, fitness could be frequency dependent, which isn't usually captured in typical fitness assays. In some cases, certain types have been found to arise and invade ancestral populations repeatedly without any fitness benefit as typically measured ([22] for example). Fitness aside, rate of adaptation has also been described by the rate of new mutations being substituted in the population. Sequencing has revealed that the substitution rate of new mutations usually decreases over the course of adaptation (reviewed in [11]). One notable exception to this has been that the rate of genomic evolution in the long-term lines of

6

the Lenski lab has been relatively constant over about 20,000 generations ([3]). As with repeatability and epistasis, the rate of adaptation in terms of substitution rate is being explored to new depths with ever-advancing sequencing technology.

Evolution experiments have informed the impact that each of these topics has on predicting adaptation. As genomic data for these experiments begins to pile up (reviewed in [2], [8], [11]), it will be interesting to see how much of the stochasticity in evolutionary trajectories can be understood. In regards to the model system used in the work presented here, we know that some of the genetic changes underlying adaptation are parallel, and some are not ([25]). Interestingly, the degree of parallelism is higher in one ecotype than the other. Epistasis is evident; the effects of at least one mutation depend on the genetic background in which it arises ([31]). Lastly, the substitution rate was found to be higher in some lines than others ([25]), but it's unclear if it decreases over the course of the experiment.

2.2 Model system

2.2.1 Motivation for sympatric speciation experiments

For many decades following the modern synthesis, it was thought that speciation occurred most commonly in allopatry; through reproductive isolation between subpopulations acquired in geographic isolation ([13]). Theoretically, it was much easier to show how populations that become physically subdivided could diverge. On the other hand, theory supporting speciation was largely non-existent at the time, leading some scientists ([36] for example) to remain unconvinced that speciation was even possible without geographical isolation. Near the turn of the millennium, however, a solid theoretical framework for predicting sympatric speciation was laid down by [21], [38] and [12]. This framework helped to alleviate doubts of sympatric speciation as an observable process in nature, spurring many researchers to consider the possibility. Some theoreticians remained skeptical that sympatric speciation was common, arguing that the conditions under which it was predicted to occur theoretically were too strict to be found in nature ([10]), though some researchers did find that these specific conditions were found in natural populations ([59]). Following that, some reviews concluded that sympatric speciation is theoretically plausible and has been observed in multiple instances, but that it is still unknown how common it is ([4]). The experiment presented in this work is based on a near-decade long series

of experiments that initially sought to understand and explore sympatric diversification of microbes.

2.2.2 Previous experiments

In the initial experiment, conducted by Friesen *et al.* (2004), a strain of *E. coli* was evolved in 12 replicate populations by growing in a liquid medium containing glucose and acetate as carbon sources and diluting by 1:100 every 24 hours ([18]). In this experiment, 12 out of the 12 populations were found to contain two distinct types of colonies after 1000 generations. These two different types of colonies differed mainly in the size of the colony after about 12 hours of growth on a solid agar medium. This initial study also found that the two types differed in their diauxic growth patterns; one type grew faster during the first phase of growth (glucose), and the second type had a shorter lag time between switching over to the second phase of growth (acetate). These two types were named FS (for fast-switcher) and SS (for slow-switcher). Furthermore, they found that these types were maintained by negative frequency-dependent selection.

Next, competition experiments were performed between clones isolated from different replicate lines and different evolutionary environments ([57]). In the initial experiment, populations were propagated in single resource environments (glucose only and acetate only) as well as the mixed resource environment. FS and SS clones were isolated from either the same or different environments and competed against each other to determine if the competitive relationship was dependent on the environment in which the clone had evolved. Clones that had evolved in the same environment (a FS from glucose+acetate vs. a SS from a different glucose+acetate line) demonstrated the same competitive relationships (maintained by negative frequency dependence), showing that diversification was parallel. Clones that had evolved in different environments (a FS from glucose only vs. a SS from glucose+acetate, for example) demonstrated varied competitive relationships with no evidence of stable intermediate frequencies, showing that diversification was not parallel.

Measurements of glucose and acetate concentration in the growth medium over a 24-hr period showed that acetate concentration increased during the first phase of growth of ancestral and SS strains (since acetate is a byproduct of glucose metabolism), while acetate did not increase in the first phase of growth of an FS strain ([53]). This indicated that acetate consumption was not repressed in the FS strain. To investigate the genetic mechanism behind this, expression levels of genes known to be involved in acetate metabolism

were measured. An increase in expression of the *aceB* gene was found to be associated with an insertion mutation in the *iclR* gene. PCR screening of the *iclR* gene showed that this insertion mutation was present in 8 of 9 FS clones isolated from the same line, but not present in any of the SS clones, and not present in any FS clones isolated from different lines, demonstrating there are other genetic changes underly the FS phenotype.

To further investigate the genetic differences between FS and SS ecotypes, microarrays were used to profile global transcription of an FS clone, an SS clone, and the ancestral clone ([30]). The FS and SS clones used in this study were isolated from the 1000 generation time point of a single population. This study found many differentially expressed genes in common when comparing FS to the ancestor and SS to the ancestor. These changes were thought to be generally adaptive to growth in serial batch culture and included genes involved with translational efficiency, glucose uptake capacity, and survival during stationary phase. Genes differentially expressed between the FS and SS clone were associated with upregulation of the TCA cycle and acetate consumption in FS and with upregulation of genes involved in acetate excretion in SS. Most importantly, this study strongly supports the argument that diversification in this system is driven by competition for carbon sources, as shown by the metabolic differences between FS and SS clones.

To test the theory that populations first undergo a phase of directional selection before diversification becomes adaptive, the next study isolated SS clones from three different timepoints (before the branching point) and allowed them to re-evolve ([54]). This was to determine if the likelihood of diversification changed over time. The re-evolution period was about 140 generations and this experiment found that the clones isolated from 400 generations were much more likely to diversify in that time than clones isolated from earlier, in line with evolutionary branching models predicting that populations evolve towards a branching point before diversifying. This study also describes the diverse ecotypes in terms of specialists and generalists. Interestingly, instead of populations evolving towards a glucose/acetate generalist before diversifying into a glucose specialist type and an acetate specialist type, they evolve towards a glucose specialist before diversifying into an ever more specialized glucose specialist and a glucose/acetate generalist.

Since the previously described experiment only investigated the likelihood of diversification before the branching point, and only in a single population, another 'rediversification' experiment was conducted using multiple FS and SS clones isolated from the endpoint of the original evolution

experiment ([56]). This study, lasting for about 200 generations, found that only SS-initiated populations diversified into populations containing both FS and SS types. In contrast, none of the FS-initiated populations diversified, suggesting that a much higher proportion of mutations cause a phenotypic change from an SS to an FS strain as opposed to the reverse direction. This study highlights the importance of mutational bias in predicting evolutionary outcomes.

Following the experiment that identified differentially expressed genes between FS and SS clones, one specific mutation was chosen to investigate its phenotypic and fitness effects ([31]). Since many of the differentially expressed genes are regulated by the global transcription factor *arcA*, a point mutation in this gene was thought to underly the phenotypic differences between FS and SS. Using allelic replacement techniques, this mutation was inserted into the genomes of both FS and SS clones isolated from many different timepoints. Interestingly, the *arcA*$^-$ mutation had a large effect in the SS background, but not in the FS background. In particular, SS*arcA*$^-$ clones from every timepoint showed a dramatically reduced lag time between switching from glucose consumption to acetate consumption. Invasion experiments showed that the mutation was adaptive in SS backgrounds (SS*arcA*$^-$ outcompetes SS*arcA*$^+$) but not adaptive in FS backgrounds (FS*arcA*$^-$ does not outcompete FS*arcA*$^+$). Furthermore, the mutation was no longer adaptive in the presence of the FS ecotype; the mutation is only adaptive in SS when the FS type is not present, demonstrating an environment-dependent fitness effect.

The work presented in this thesis combines these last two studies to investigate the effect of the *arcA*$^-$ mutation on the likelihood of diversification. Simply, this "rediversification" experiment found that the mutation was associated with reduced diversification in the SS background and increased diversification in the FS background.

Chapter 3

Adaptation in *E.coli*: Intraspecific Ecological Interactions and Genetic Constraints Determine Diversification

3.1   Introduction

Diversity is abundant, and there are many mechanisms that can generate diversity. Some of these mechanisms have been shown to generate intraspecific diversity in sympatry, resulting in the formation of new species, or in the case of microbes, divergent ecotypes ([16], [1], and [29] for example). One such mechanism is disruptive selection that is generated by intraspecific competition for resources, resulting in diversification ([12]). With this mechanism, the combination of negative frequency-dependent ecological interactions and phenotypic evolution can transform an undifferentiated population into a diversified one ([14]). For example, *E. coli* propagated in a well-mixed resource-limited environment repeatedly diversifies into two distinct types ([18]). While this process has been observed several times, it remains unclear to what degree this divergence is deterministic, and what factors affect the likelihood of diversification in populations under this type of selection. Understanding the potential for these populations to diversify is critical to predicting evolutionary responses to changing conditions.

In populations that are initially isogenic, like a bacterium infecting a novel host, for example, variation for adaptation is supplied by changes in DNA sequence. Because of this, we often use mutation rate as a proxy for evolutionary potential, predicting that organisms with higher mutation rates will be more successful at adapting to novel environments. Though there does exist experimental evidence to support this relationship (such as

[6] and [52]), there are other factors influencing the ability of a population to adapt. We can group these factors into two broad categories: ecological factors that shape selective pressures, and genetic factors that provide adaptive mutations. Specifically, the evolutionary potential of a population to diversify depends on whether or not selection is disruptive due to the phenotypic composition of the population (ecological factors) and how likely it is that genetic mutations result in a phenotypic change that is relevant to resource consumption (genetic factors).

If we temporarily ignore underlying genetic constraints and assume that the probability of diversification depends only on ecology, we can use an adaptive landscape to predict evolutionary trajectories. A monomorphic population occupying a fitness valley between two peaks would have the highest chance of diversifying, while a population occupying a fitness peak or ridge would have a lower chance ([62], [20]). If the landscape is static, when a population that occupies a valley diversifies, the two diverged populations would each have a decreased likelihood of further diversifying as they ascend their respective peaks. Observations in diversifying populations of *Pseudomonas fluorescens* are consistent with this expectation, as they show a reduced propensity to diversify after the initial adaptive split ([44], [5]). When the shape of the fitness landscape is also determined by the phenotypic composition of the population (i.e., frequency-dependent), evolutionary change can result in a change in the shape of the landscape. In this case, a dynamic landscape can make predicting diversification much more difficult.

Adaptive dynamics models account for changing landscapes, describing the fitness of a given phenotype and subsequent evolution as a function of the composition of the population ([21]). While these types of models incorporate the context dependence of phenotypic fitness, which are typically ignored in classical population genetic models, they ignore the genetic details that determine which phenotypes are and are not accessible. Microbial evolution experiments have previously been interpreted using the adaptive dynamics framework ([54], [56]), but it remains unclear whether the predictions made under this framework are robust to the types of evolutionary constraints imposed by mutational landscapes (like those described in [60]). The utility of these models in predicting evolutionary trajectories of natural populations depends on these assumptions, so it is critical to understand how epistasis and genetic background interact to constrain trajectories predicted by these models.

Since diversification occurs when new phenotypes arise, the likelihood of diversification is constrained by the possible genetic changes that can give

rise to the necessary phenotypic variation. Typically, models of adaptive evolution assume that mutation effect sizes are symmetrical around a phenotype, so that random changes are just as likely to shift a particular trait in one direction as they are in the other ([34]). Due to complex epistatic networks, however, this is often not the case in nature ([60], [50]). For example, there may be many more ways to generate a null mutation than a variant with increased enzyme activity. Furthermore, while the probability of a mutation at any point in the genome is approximately uniform, the effect of that mutation on phenotype, and therefore adaptation, depends both on the genomic background in which it arises, and on the environment that the organism currently occupies. All of these aspects of the biology of real organisms could limit the utility of abstract models for predicting evolutionary outcomes.

In the current work, we use populations of *Escherichia coli* to investigate the interaction of genetic and ecological factors on the likelihood of diversification. This follows a number of experiments exploring the diversification of *E. coli* in a laboratory environment. When propagated in a well-mixed environment containing two carbon sources, originally isogenic populations diversify into two distinct, heritable types ([18]). Typical of bacterial growth in a medium with two resources, populations exhibit diauxic growth in which there are two distinct phases of growth separated by a period of little or slow growth. The first phase is comprised of the consumption of the preferred resource (glucose), followed by the consumption of the secondary resource (acetate). The two types observed to coexist after many generations differ primarily in the amount of time between phases. One type (referred to as SS for "slow-switcher") consumes glucose efficiently, but exhibits a long lag time between depleting glucose and consuming the secondary resource. The second type (FS for "fast-switcher") consumes glucose less efficiently than SS, but has a much shorter lag between phases.

Previous experiments found this diversification to be consistent, arising in 12 of 12 lines between approximately 200 and 400 generations of daily batch culture ([18]). Competition experiments determined that the divergent types are maintained by negative frequency-dependent selection due to trade-offs in resource consumption ([58]). Sequencing of FS and SS clones from three different divergent populations showed that the metabolic differences were due to some shared mutations and some unique mutations ([25]). Interestingly, SS clones were more similar to each other than FS clones were to each other, in terms of the mutations that were detected. A single point mutation in the *arcA* gene (which encodes a transcription factor for anaerobic respiration control) was found by [30] to be associated with many of

the genes that are differentially expressed between FS and SS. To quantify the fitness and phenotypic effect of this mutation, it was inserted into the genome of several SS and FS clones isolated from every 200 generations from the original evolution experiment ([31]), generating strains genetically identical except for this point mutation. When introduced into the genome of an FS individual, this mutation has little or no effect on phenotype ([31], this work). In the SS background, however, this mutation results in an intermediate lag time between growth phases (referred to here as MS, for "medium switcher"). Furthermore, between the clones isolated from different time points along the original evolution experiment, we found no measureable difference in the effect of the mutation on switch time (i.e. the $arcA^-$ allele reduced the switch time in an SS individual isolated from gen200 by about the same amount as in an individual from gen1000). Because of the clear ecotype-dependent phenotypic effect of this particular mutation, these engineered strains were used to investigate diversification in this system with a novel set of initial conditions; specifically, the effect of the $arcA^-$ mutation, initial ecotype, and number of generations previously evolved on the likelihood of subsequent diversification.

Here, we take isogenic colonies of these genotypes (SS$arcA^+$, SS$arcA^-$, FS$arcA^+$, and FS$arcA^-$) from five different time points from the experiment conducted by [31] and expose them to a second bout of evolution, lasting approximately 220 generations. To determine if populations underwent diversification, we measure the "switch-time" of several clones from the population, and from this calculate two parameters: 1) the variance in this phenotype present in the sample and 2) the range of phenotypes present in the sample (both measures of diversification). By comparing these statistics between lines, we can determine how the likelihood of diversification changes over time, between ecotypes and what effect the $arcA^-$ mutation has on this likelihood. With a full factorial design, we can determine the effects of these factors individually as well as the interactions between them in an attempt to understand the relative contributions of ecological and genetic factors to the process of diversification. Evidence from this experiment shows that both of these factors play a role in adaptation and that predicting diversification requires a careful integration of ecological and genetic complexity.

3.2   Methods

3.2.1   Bacterial strains

The bacterial strains used in this experiment were isolated from a previous long-term evolution experiment conducted by [58], in which the original founding ancestor was *Escherichia coli* B strain REL606. Originally, twenty replicate populations were propagated for approximately 1200 generations in a liquid media supplemented with glucose and acetate, in the same manner as described in [18], and also described in the Rediversification section below. One of these original populations, pop20, was chosen for further experiments by [31], and all genotypes selected for this work were from this population. This particular population was chosen partly because it diversified relatively early, with both SS and FS ecotypes being present in the population by 200 generations. [31] isolated both FS and SS clones from several different time points along the original evolution experiment. Then, using a pKO3 plasmid vector carrying a mutated version of the *arcA* allele, the *arcA* locus of the clones was modified (by [31]) to contain a thr81ala point mutation. The mutation was inserted into the genome by homologous recombination between the plasmid and genome in the regions flanking the point mutation, followed by excision of the plasmid DNA by a second recombination event. After sequencing to ensure that no other mutations had arisen, the allelic replacement resulted in genotypes that differed only at the *arcA* locus by the presence or absence of this point mutation. For the current work, to ensure that our founding populations were isogenic, we isolated individual FS and SS clones from five time points along the frozen fossil record, beginning at 200 generations, recurring every 200 generations (illustrated in Figure 3.1, panel A). As well as these 10 genotypes, we use the modified strains with the $arcA^-$ mutation, for a total of 20 unique genotypes. Before beginning the evolution experiment, the clones were screened for the presence or absence of the mutation using restriction enzymes.

3.2.2   Rediversification

The 20 founding genotypes, each replicated four times, were propagated in serial batch culture for approximately 220 generations, yielding 80 evolved lines in total (see Figure 3.1 for a schematic of the experimental design). Since these bacteria have been previously exposed to experimental evolution, we refer to this second bout of evolution as the rediversification period. The environmental conditions we exposed the bacteria to were identical to the one in which the populations were originally evolved; 18-mm diameter tubes

with 10 mL of Davis Minimal (DM) media supplemented with 0.25 mg/ml glucose and 1.3225 mg/ml sodium acetate trihydrate (DMGA). Every 24 (+/- 1) hours, 100 μL of culture was transferred to fresh media, and these were stored in a shaking incubator maintained at 250 rpm and 37°C. Samples from all 80 lines were taken every four days and frozen in 20% glycerol.
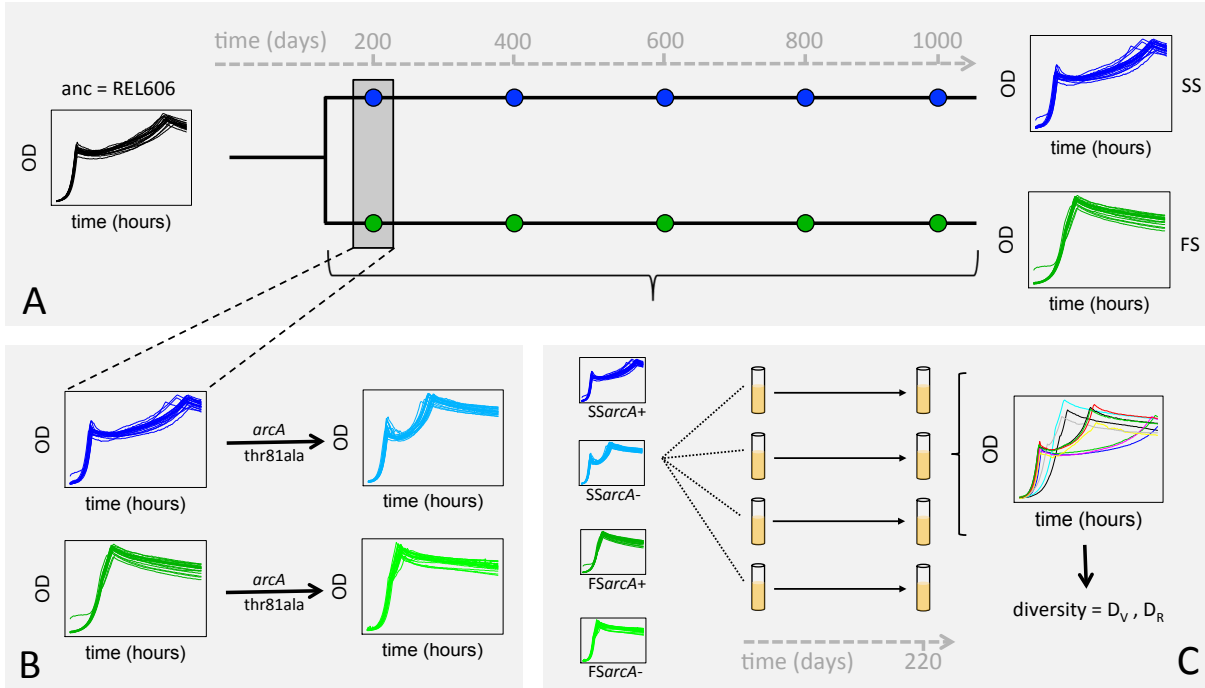
Figure 3.1: Schematic of experimental design. Individual clones were sampled every 200 generations from the frozen fossil record of the original 1000+ generation evolution experiment (panel A). When transformed with the thr81ala point mutation in the *arcA* gene ($arcA^-$), there is a large effect on the switch time phenotype of the SS type, producing the "medium-switcher" phenotype (panel B, top row), while having no observable effect on the phenotype of the FS type (panel B, bottom row). All four types (SS$arcA^+$, SS$arcA^-$, FS$arcA^+$, and FS$arcA^-$) from every 200 generations are subjected to a secondary bout of evolution (referred to as the rediversification period) lasting for approximately 220 generations, with each genotype evolved in four replicate lines (panel C). After the rediversification period, individual colonies are assayed for switch time, defined as the amount of time (in hours) between maximum density in the first phase of growth and maximum density in the second phase of growth, and this parameter is used to determine diversity present in a population.

3.2.3   Phenotypic assays

Following the rediversification period, all 80 lines were assayed for phenotypes present in the population. As a control, the 20 founding genotypes were also assayed (i.e. the lines before the rediversification period, or nonevolved). To determine phenotypes, we isolated and measured characteristics of several individuals within each population. First, 5 mL tubes of DMGA were inoculated with frozen samples of culture and incubated overnight. Then, each population was diluted, plated onto tryptone plates, and incubated overnight. Following that, individual colonies were picked randomly and transferred via sterile toothpick to 5 mL tubes of DMGA, to be incubated overnight. Lastly, after 24 hours of growth, 2 μL of culture was used to inoculate 200 μL of DMGA in a 96-well plate. Using a Biotek EL808 microplate reader, optical density was measured at 600 nm every 15 minutes over a period of 36 hours, to generate a plot of population size over time for each colony. We use this growth curve as the phenotype of a colony for further analysis. From all 80 lines following rediversification, we assayed 10 colonies per line, and from 20 lines prior to rediversification, we assayed 20 colonies, to yield a total of 1200 growth curves.

3.2.4   Statistical analysis

All analysis of the growth curve data was done using the statistical program R v 2.14.1 ([42]). We defined "switch time" as the distance (i.e. time) between the maximum optical density in phase 1 of growth and the maximum optical density in phase 2 of growth. For SS types, this quantity is relatively large ($\sim$25 hours), while for FS types it is small ($\sim$2 hours, illustrated in Figure 3.1, panel A). For the curves that did not appear to reach a second maximum within the 36 hour growth assay, as was the case with a small proportion of SS types, the end time point was used to ensure a conservative measure of distance between maxima. To quantify diversification, we calculated two measures from the switch time data. The first, "diversity by variance" (D_v), was defined as the variance in switch time between individuals within a population. The second, "diversity by range" (D_r), was defined as the difference between maximum and minimum switch time within each population. Using a threshold value, populations with $D_r \geq 15.0$ hours were considered to have diversified. To compare diversification between genotypes, a multi-factor ANOVA was performed, using generation (200, 400, etc.), ecotype (FS/SS), and mutation (presence/absence) as the explanatory variables.

3.3 Results

3.3.1 Naming convention

To avoid confusion between the evolved phenotypes and the original founding genotypes, we adopt the following conventions. Although SS and FS have been used to refer to a phenotype in previous papers, here we use these terms to refer to the founding genotypes. To refer to a phenotype, we call individuals *slow*, *medium*, or *fast*, according to the time it takes them to switch between phases of growth. Since the founding clones isolated from different time points are also distinct genetically, we use gen200, gen400, etc. to refer to the previous amount of generations they had evolved in the original experiment.

3.3.2 Phenotypic evolution

To assess phenotypic evolution, we sampled random individuals from each population and quantified their phenotype (switch time), using the 36 hour growth profile generated for each clone. To estimate phenotypes before the rediversification period, we sampled 20 colonies of each of the four types (SS$arcA^+$, SS$arcA^-$, FS$arcA^+$, and FS$arcA^-$) from all five time points. After the rediversification period, we sampled 10 colonies of each replicate population of each type from all five time points. These data are summarized in Figures 3.2 and 3.3, though it should be noted that since all timepoints and replicates are pooled together to show general changes in phenotypes, these two figures do not show information about the diversity present in individual populations. We found 12 out of the 20 populations initiated with the SS$arcA^+$ ancestor contained two distinct ecotypes after 220 generations, while only 1 out of 20 populations initiated with SS$arcA^-$ contained more than one ecotype (Table 3.1, Figure 3.2 bottom panel). Moreover, the populations initiated with SS$arcA^+$ that diversified, did so into *fast* and *slow* types, while populations initiated with SS$arcA^-$ did not diversify, but instead evolved into *fast* types. The one SS$arcA^-$ initiated replicate that did diversify was a gen1000 line and was found to be clearly dimorphic. Of the 20 populations initiated with FS$arcA^+$ and FS$arcA^-$, 8 and 11 populations diversified, respectively (Table 3.1, Figure 3.3).

Initial genotype	Lines with $D_r \geq 15.0$
SS$arcA^+$	12
FS$arcA^+$	8
SS$arcA^-$	1
FS$arcA^-$	11
All genotypes	**32**

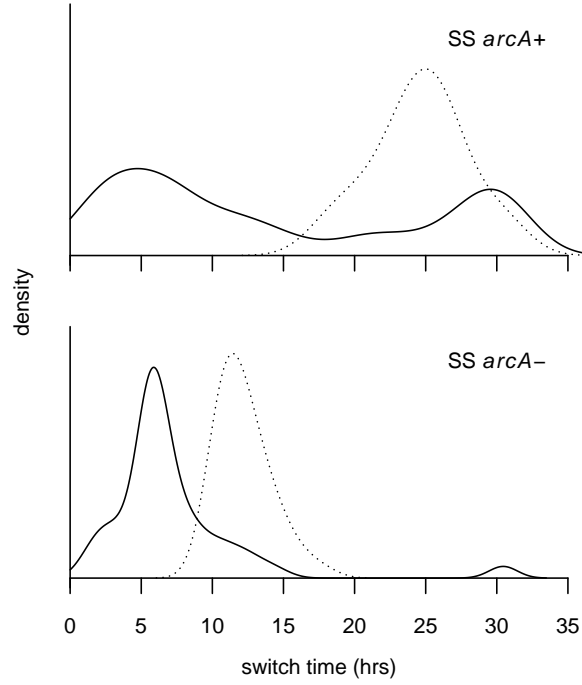Table 3.1: Summary of proportion of lines of each genotype that diversified.

Figure 3.2: Phenotypic evolution within SS types. Density distribution of the switch time trait in SS $arcA^+$ and SS $arcA^-$ both prior to (dotted lines) and following (solid lines) the rediversification period ($\sim$220 generations, showing the evolution of switch time within each type. All five timepoints and replicates are grouped together.
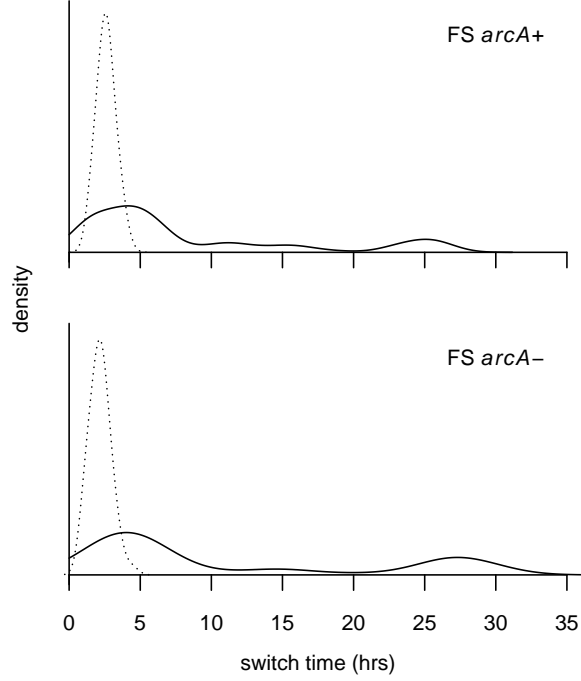
Figure 3.3: Phenotypic evolution within FS types. Density distribution of the switch time trait in FS $arcA^+$ and FS $arcA^-$ both prior to (dotted lines) and following (solid lines) the rediversification period ($\sim$220 generations, showing the evolution of switch time within each type. All five timepoints and replicates are grouped together.

3.3.3 Effect of $arcA^-$ mutation on diversification

To quantify diversity present in each population, we calculated both the variance in switch time of individuals within that population (D_v) as well as the range of switch times (D_r). Since the distribution of D_r values was clearly bimodal, we chose the middle point between the two peaks ($D_r \geq 15.0$ hours) as the threshold value to consider a population diversified. Given this criteria, after 220 generations of evolution, 32 out of the 80 evolution lines were found to be diversified (Table 3.1). Of these, many populations showed a clear dimorphism of *fast* and *slow* switcher types, while others showed a clear trimorphism of *fast*, *slow*, and *medium* types. See Appendix 1 (Figures A.1, A.2, and A.3) for examples of these populations. In a small

proportion of lines, individual curves varied widely, but did not form distinct clusters, though a lack of clusters could be due to a small sample size. The effect of the $arcA^-$ mutation on the propensity to diversify was found to be highly dependent on whether it was in an FS or an SS line. To quantify the effect of the mutation, we took the difference between D_v for $arcA^+$ populations and D_v for $arcA^-$ populations and called this ΔD_v (Figure 3.4). Within SS-initiated lines, arcA+ populations were much more likely to diversify than $arcA^-$ populations (Figure 3.4). For all five timepoints, the mutation dramatically hindered diversification, with a considerably stronger effect observed for SSgen200. The mutation had the opposite effect in FS-initiated lines; $arcA^-$ populations were much more likely to diversify than $arcA^+$ populations (Figure 3.4). This was the case for all timepoints with the exception of FSgen200.
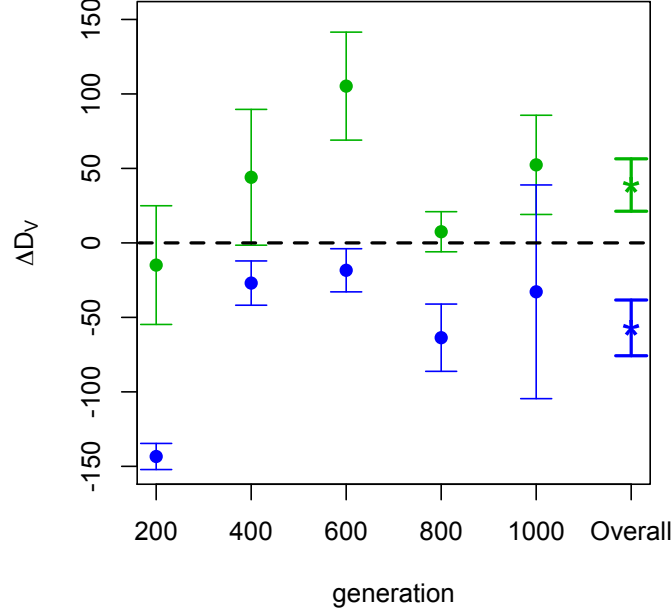
$$\Delta D_v = D_v(arcA^-) - D_v(arcA^+) \tag{3.1}$$

Figure 3.4: Effect of *arcA* mutation on diversity by variance (D_v) in both FS (green circles) and SS (blue circles). ΔD_v is the difference between types with the mutation and types without the mutation, so that a positive value denotes an increase in diversification associated with the *arcA* mutation and vice versa. Starred points denote average effect across all timepoints.

3.3.4 Effect of generation on diversification

We also investigated the effect of timepoint, or the amount of generations previously evolved, on diversification. Within populations without the $arcA^-$ mutation (i.e. SS$arcA^+$ and FS$arcA^+$ genotypes only), timepoint was found to be a significant predictor of the evolution of diversity. Specifically, in FS$arcA^+$, we observed a negative correlation between timepoint and D_v (Figure 3.5, second panel), showing that genotypes isolated from earlier in the original evolution experiment are more likely to diversify than those isolated later. Since the founding population diversified around 200 generations, this finding indicates that diversification is more likely for clones isolated closer to the original branching point. In SS$arcA^+$, however, the

relationship between generation and variance was found to be non-linear (Figure 3.5, first panel), showing that genotypes isolated from early and late in the original evolution experiment are more likely to diversify than those isolated from intermediate timepoints. Not surprisingly, we found no effect of generation on diversification within SS$arcA^-$ initiated lines. Since diversification in these lines was suppressed at all timepoints, no differences between timepoints could be observed (Figure 3.6, first panel). Despite moderate amounts of diversification in lines initiated with FS$arcA^-$, we found no effect of generation on diversification within those lines (Figure 3.6, second panel).
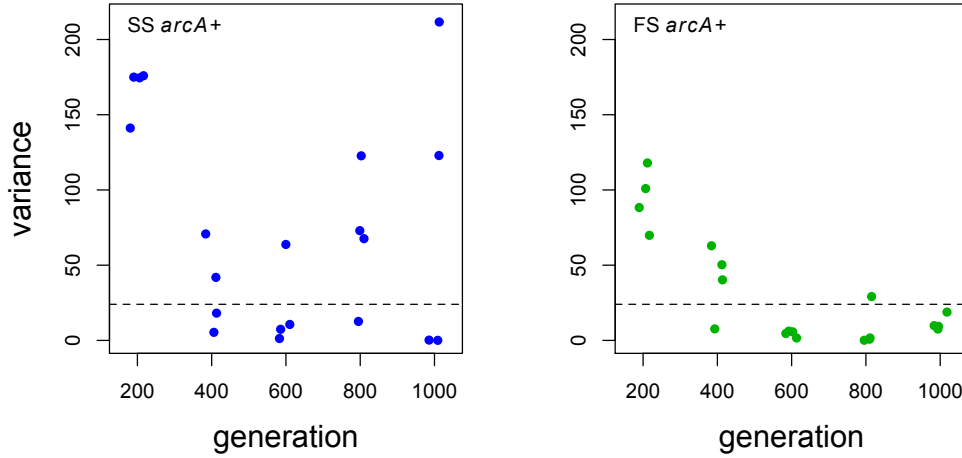

Figure 3.5: Effect of generation on diversification for both **wild type** lines. Each point represents a single replicate population and is the variance in the switch time phenotype present in that population (diversity by variance, D_v). The dotted lines separate populations which are comprised of a range of phenotypes (diversity by range, D_r) ≥ 15.0 hours from those in which $D_r < 15.0$ hours.
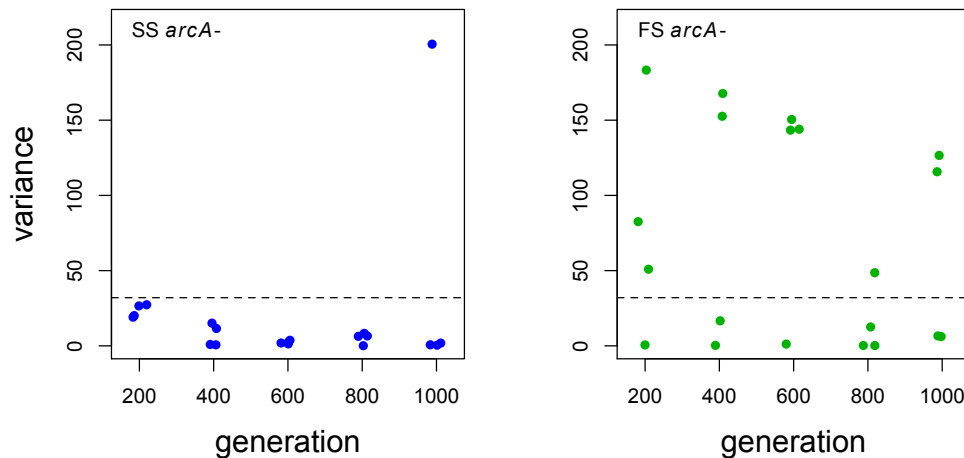
Figure 3.6: Effect of generation on diversification for both **mutant** lines. Each point represents a single replicate population and is the variance in the switch time phenotype present in that population (diversity by variance, D_v). The dotted lines separate populations which are comprised of a range of phenotypes (diversity by range, $D_r) \geq 15.0$ hours from those in which $D_r < 15.0$ hours.

3.4 Discussion

Understanding the factors that drive and/or constrain diversification can inform how populations will adapt to novel environments. Here, using bacteria isolated from experimental populations that had previously diverged, we examined the propensity to diversify in a number of ways. We asked how founding genotype, founding ecotype, and a specific point mutation affected diversification. We found that the effect of the $arcA^-$ mutation on diversification was highly dependent on the ecotype in which it was introduced; the mutation decreased diversification in SS types and increased diversification in FS types. We argue that this is evidence that adaptation in this system is constrained by both genetic and ecological factors and that it is important to consider the interaction between the two. Recently, investigation of the complexity of eco-evolutionary dynamics has been gaining momentum ([40], [49]) and the work presented here adds to a growing number of studies that

exhibit a clear feedback between ecological and evolutionary processes on the same time scale.

3.4.1   Genetic and ecological constraints

Our data shows that within lines initiated with the intermediate phenotype (SS$arcA^-$), very little diversification is observed and phenotypic evolution occurs in one direction; 19 out of the 20 populations that were initially *medium* types evolved to contain only *fast* types. Since the SS$arcA^-$ types have an intermediate phenotype, we might expect evolution to proceed in either direction along that phenotypic axis. We find consistent evolution towards a smaller switch time, indicating a constraint on phenotypic evolution, either genetically or ecologically.

Genetic factors can constrain evolution if there are differences in the availability of mutations that increase versus decrease the value of a given trait. Within the context of switch-time studied here, such differences in mutation availability could arise two ways. First, it could be the case that given the genes involved in the underlying metabolic network, there are more mutations that result in a shorter switch time than a longer switch time (for example, more loss-of-function than gain-of-function mutations). This hypothesis is consistent with results showing that populations initiated by SS types are more likely to diversify than FS types (Figure 5A), implying a greater availability of mutations that decrease the switch time. This lower diversification in FS is unlikely to be constrained by ecology alone, because FS types can readily invade SS populations ([18], [54]). The second way that an asymmetry in mutation effect could explain the results is through epistasis, whereby the $arcA^-$ mutation itself changes the distribution of mutation effects available to increase or decrease switch time. Since the protein made by $arcA$ regulates at least 30 operons ([55]), it seems plausible that a mutation changing the function of this gene could drastically alter the distribution of mutation effect. This explanation is consistent with results showing that the FS$arcA^-$ types had greatly increased diversification over FS$arcA^+$ types (Figures 3.5 & 3.6), despite having phenotypes that were highly similar ([31]). Since we know that mutations that result in *slow* types are beneficial in populations of *fast* types, this difference suggests that in an $arcA^-$ background, it is more likely that a new mutation will result in a *slow* phenotype. This finding demonstrates that a seemingly neutral mutation can turn out to be adaptive in future generations (like those described in [61]). Even though the interactions between metabolic genes in *E. coli* are relatively well understood, since these interactions are environment de-

pendent, we cannot make predictions about the effects of specific mutations, even given the underlying network of epistasis. For example, [26] found that LOF (loss-of-function) mutations in any one of three genes (*ppsA*, *sfcA*, and *maeB*) resulted in a longer transition lag in a mixed environment, but had no phenotypic effects when grown in either glucose alone or acetate alone. It is this environmental dependence that prevents the mutational landscape from remaining constant after a single mutation and provides yet another obstacle in predicting evolutionary trajectories.

Though the availability of mutations is an important factor in determining adaptation, the selective environment determines which mutations and phenotypes will increase versus decrease in frequency, which modifies predictions that might be made based on the above discussion of genetic constraints. For example, while our results suggest that the *arcA*⁻ mutation causes an increase in the availability of mutations that result in a slower switch time, all of the SS*arcA*⁻ populations evolved towards a *fast* phenotype (Figure 3.2). In this experiment, the selective environment is partly determined by the composition of the current population. The different ecotypes have alternative strategies for consuming resources, which in turn determines which resources (glucose and acetate) will be more or less abundant. Previous experiments have shown that *fast* types can invade a population of *slow* types, as well as vice versa, and that these types are maintained at an intermediate frequency ([18]). From the current experiment, it is clear that mutants with a *fast* phenotype can invade a population of *medium* types, but instead of reaching a dimorphic state, *medium* types are outcompeted, and the *fast* types completely replace the ancestral population (Figure 3.2). This is indicative of directional selection, as opposed to the disruptive selection experienced by populations of either *fast* or *slow* types. We hypothesize that since the *medium* types (initially SS*arcA*⁻) have an intermediate phenotype, there is only one niche available to occupy. Initially, the two typical niches (occupied by *fast* and *slow* types) are not available for colonization until the resident population is far enough away in phenotypic space from the other potential resident morph. This pattern of directional evolution followed by evolutionary branching was observed in a single population (initiated by gen1000-SS*arcA*⁻ and we hypothesize that if allowed to evolve for longer, more of the SS*arcA*⁻ initiated populations would show this same pattern of diversification.

A lack of diversification in SS*arcA*⁻ initiated lines suggests that *medium* types cannot coexist with other types. We did, however, observe that some of the lines initiated by other strains did have *medium* individuals present after the rediversification period. It is unlikely that the occurrence of these

intermediate types is the result of recurring mutations that are selected against, because they reach high frequency (>20%) in certain populations. There are two other possible explanations for the persistence of *medium* types: either they have diverged into an open niche along a phenotypic axis other than switch time that we did not measure, or the *slow* and *fast* types could be phenotypically divergent enough to create an open niche at an intermediate switch time. We did not find a difference between the switch time of *fast* and *slow* types in populations with *medium* types present and the switch time of *fast* and *slow* types in populations with no *medium* types present, suggesting that the second of these possibilities is less likely, though it may be the case that the difference occurs on a finer scale than we measured. It would be interesting to isolate and compete different *medium* clones to evaluate which of these explanations is most likely.

The findings of this experiment, especially the evolution of the *medium* types, combined with previous experiments suggest that there is a complex underlying ecology guiding evolution. The interactions between different phenotypes could be visualized by a pairwise invasibility plot, or PIP ([21]), shown in Figure 3.7. Previously, [54] showed that the ancestral ecotype first evolved an increase in switch time before diversifying, indicating a phase of directional selection towards a branching point. This can be interpreted on a PIP as an evolutionary attractor (point B, Figure 3.7), meaning that the branching point would always be reached, irrespective of the starting phenotype. The finding that *medium* types evolve in one direction (towards a *fast* phenotype only) allows us to infer that there is another important feature in this PIP. We can include this information by adding an evolutionary repellor somewhere between the ancestral and *medium* phenotype (point A, Figure 3.7), meaning that phenotypes will always evolve away from that point in either direction, depending on the starting phenotype. Because we know that *slow* types cannot invade populations of *medium* types, the initial *medium* phenotype must lie on the *fast* side of the evolutionary repellor. Though the picture could contain more complexity, this PIP summarizes the invasion fitnesses for which we have evidence. Understanding how selection regimes change with evolving phenotypes, together with a clear genetic picture of which phenotypes are attainable is crucial to predicting evolutionary trajectories.
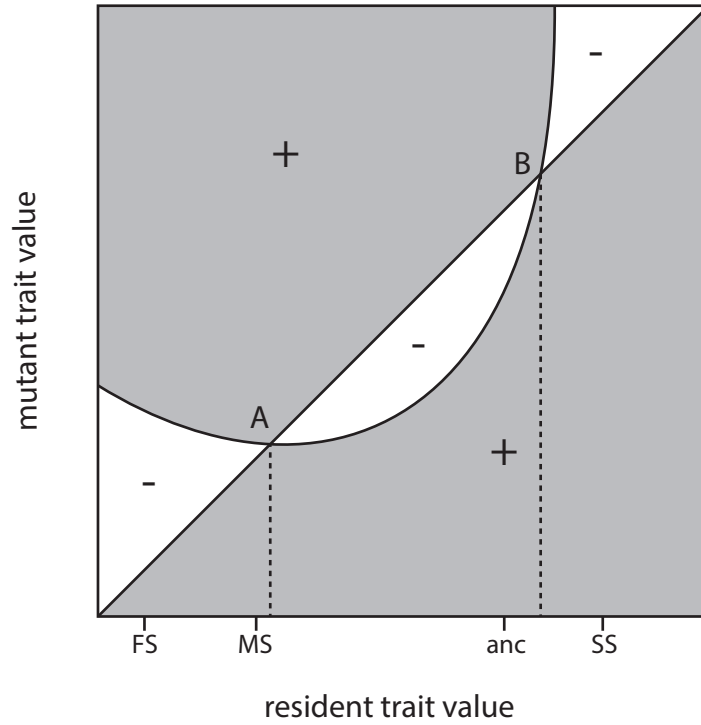
Figure 3.7: Pairwise invasibility plot, showing regions of positive invasion fitness (grey) and negative invasion fitness (white) for all combinations of resident (x-axis) and mutant (y-axis) phenotypes (in this case, "switch time"). Invasion fitness is zero on the diagonal (since mutant and resident have the same fitness), and evolutionary equilibria are the intersection points of the diagonal and the invasion fitness 0-isocline. The slope of this line at the intersection determines the evolutionary dynamics, so that point A is an evolutionary repellor and point B is an evolutionary attractor. Since there are multiple equilibria, the direction of evolution depends on the initial conditions.

3.4.2 Effect of timepoint

The decrease in the propensity to diversify with the number of generations previously evolved in the FS$arcA^+$ initiated types shows that, within this type, as timepoint moves later from the original branching point (which was before 200 gens), diversification becomes less likely. This finding is in accordance with theoretical predictions that attribute this decrease in diversification to the increasing phenotypic distance between divergent types. Though we do not observe an increase in phenotypic distance (i.e. the *fast* types from gen1000 do not have statistically shorter switch times than those from gen200), it is likely that there are differences in other measures. Furthermore, it is possible that the accumulation of mutations has affected this picture of diversification probability. For example, as evolution proceeds, it could be the case that the mutations that fix in the population are those that are redundant with the mutations that decrease switch time. Under this hypothesis, what changes over evolutionary time is the distribution of mutation effect in this particular trait dimension, therefore decreasing the likelihood that a mutation arising will result in a *slow* phenotype. It is interesting to observe the opposite, an increase in diversification in later timepoints within the SS$arcA^+$ lines, and perhaps this is indicative of a second branching point further along in evolutionary time, though this is complete speculation.

Though we observed an effect of timepoint on diversification in both SS$arcA^+$ and FS$arcA^+$ derived populations, diversification did not depend on timepoint in SS$arcA^-$ and FS$arcA^-$ derived populations (Figure 3.5 and 3.6). While we did not have an *a priori* prediction for how the effect of the $arcA^-$ mutation on diversification would depend on the amount of time previously evolved, this result could make sense in light of expected epistatic interactions between $arcA^-$ and mutations in the different founding clones. In SS$arcA^-$, an effect of timepoint is not observed, possibly because the mutation has such a dramatic effect on phenotype, resulting in reduced diversification, thus trumping the effect that any genetic differences might have on diversification. In FS$arcA^-$, however, while the mutation increases diversification on average, there is no clear effect of timepoint, suggesting that there may be some interaction between the artificially introduced $arcA^-$ mutation and the unique mutations that are present in the individuals isolated at different timepoints. By inserting $arcA^-$ artificially, the lines with the mutation are exposed to a genetic change that has not been previously tested in the populations, varying the effect of the timepoint on diversification.

3.4.3 Conclusions

The patterns observed in the experiments described here demonstrate the importance of the constraints of genetics and the influence of ecology on the process of adaptation. We have shown evidence that the diversification of *E. coli* in a glucose-acetate batch culture environment depends both on the availability of mutations as well as the selective pressure determined by the resident population. Since both FS and SS individuals can invade populations of the other type, the asymmetry in likelihood of diversification shows that certain mutation effects are more likely than others. Furthermore, a single point mutation that does not affect phenotype (like $arcA^-$ in the FS background) can strongly influence diversification rates, suggesting that epistasis also plays an important role. Diversification in this system is also constrained ecologically because the direction and strength of selection is determined by the composition of the population: the fitness of novel types depends on the resident types. Both ecological and genetic constraints play an important role, and models of adaptation that consider only one type are insufficient for predicting evolutionary trajectories. While the effect of ecology can be represented by adaptive dynamics models, predicting the evolutionary outcomes in real-world systems requires a careful integration of ecology and genetics.

The effect of the interaction between ecology and genetics on evolutionary outcomes shown here highlights the importance of considering both of these factors simultaneously. Even though all biologists are familiar with how common and complicating gene interactions can be, we often conceptualize evolution in models as being a linear accumulation of mutations with independent effects. To simplify the process, we imagine that mutations arise and are "tested" in the environment in which they arise, and then are either removed from the population or increase to fixation. As shown here, a mutation could change the selective environment as it changes in frequency in a population, creating a feedback loop between changes in allele frequencies (evolution) and the environment (ecology). Because the particular mutation studied here ($arcA$) has such a clear affect on resource metabolism, and therefore on the selective environment, this connection is less obscure than in many well-studied systems of adaptation. Due to this mechanistic link between ecology and evolution, this study provides a clear example of the eco-evolutionary feedback loops that may become increasingly important in understanding evolutionary biology.

Bibliography

[1] Marta Barluenga, Kai N Stölting, Walter Salzburger, Moritz Muschick, and Axel Meyer. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, 439(7077):719–723, February 2006.

[2] Jeffrey E Barrick and Richard E Lenski. Genome dynamics during experimental evolution. *Nature Reviews Genetics*, October 2013.

[3] Jeffrey E Barrick, Dong Su Yu, Sung Ho Yoon, Haeyoung Jeong, Tae Kwang Oh, Dominique Schneider, Richard E Lenski, and Jihyun F Kim. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*, 461(7268):1243–1247, October 2009.

[4] Daniel I. Bolnick and Benjamin M. Fitzpatrick. Sympatric speciation: Models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):459–487, 2007.

[5] Angus Buckling, Matthew A Wills, and Nick Colegrave. Adaptation limits diversification of experimental bacterial populations. *Science*, 302(5653):2107–2109, December 2003.

[6] Lin Chao and Edward C Cox. Competition between high and low mutating strains of Escherichia coli. *Evolution*, 37(1):125–134, 1983.

[7] H H Chou, H C Chiu, N F Delaney, D Segre, and C J Marx. Diminishing Returns Epistasis Among Beneficial Mutations Decelerates Adaptation. *Science*, 332(6034):1190–1192, June 2011.

[8] Tom M Conrad, Nathan E Lewis, and Bernhard Ø Palsson. Microbial laboratory evolution in the era of genome-scale science. *Molecular Systems Biology*, 7:509, 2011.

[9] Tim F Cooper, Susanna K Remold, Richard E Lenski, and Dominique Schneider. Expression profiles reveal parallel evolution of epistatic interactions involving the CRP regulon in Escherichia coli. *PLoS Genetics*, 4(2):e35, February 2008.

[10] Jerry A Coyne, H Allen Orr, et al. *Speciation*. Sinauer Associates Sunderland, MA, 2004.

[11] Jeremy R Dettman, Nicolas Rodrigue, Anita H Melnyk, Alex Wong, Susan F Bailey, and Rees Kassen. Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Molecular Ecology*, 21(9):2058–2077, May 2012.

[12] U Dieckmann and M Doebeli. On the origin of species by sympatric speciation. *Nature*, 400(6742):354–357, July 1999.

[13] Theodosius Dobzhansky and Theodosius Grigorievich Dobzhansky. *Genetics of the evolutionary process*, volume 139. Columbia University Press, 1970.

[14] M Doebeli and U Dieckmann. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *The American Naturalist*, 156(S4):S77–S101, 2000.

[15] Santiago F Elena and Richard E Lenski. Microbial genetics: Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, 4(6):457–469, June 2003.

[16] J L Feder, C A Chilcote, and G L Bush. Genetic differentiation between sympatric host races of the apple maggot fly Rhagoletis pomonella. *Nature*, 336(6194):61–64, 1988.

[17] Anders Folkesson, Lars Jelsbak, Lei Yang, Helle Krogh Johansen, Oana Ciofu, Niels Høiby, and Søren Molin. Adaptation of pseudomonas aeruginosa to the cystic fibrosis airway: an evolutionary perspective. *Nature Reviews Microbiology*, 10:841–851, 2012.

[18] Maren L Friesen, Gerda Saxer, Michael Travisano, and Michael Doebeli. Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in Escherichia coli. *Evolution*, 58(2):245–260, February 2004.

[19] T Garland and Rose M R, editors. *Experimental Evolution*. University of California Press, 2009.

[20] Sergey Gavrilets. *Fitness landscapes and the origin of species (MPB-41)*. Princeton University Press, 2004.

[21] S.A.H. Geritz, E. Kisdi, G. Mesze NA, and J.A.J. Metz. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary ecology*, 12(1):35–57, 1998.

[22] Aleeza C Gerstein and Sarah P Otto. Cryptic Fitness Advantage: Diploids Invade Haploid Populations Despite Lacking Any Apparent Advantage as Measured by Standard Fitness Assays. *PLoS ONE*, 6(12):e26599, December 2011.

[23] Mel Greaves and Carlo C Maley. Clonal evolution in cancer. *Nature*, 481(7381):306–313, 2012.

[24] Daniel L Halligan and Peter D Keightley. Spontaneous Mutation Accumulation Studies in Evolutionary Genetics. *Annual Review of Ecology, Evolution, and Systematics*, 40(1):151–172, December 2009.

[25] Matthew D Herron and Michael Doebeli. Parallel evolutionary dynamics of adaptive diversification in Escherichia coli. *PLoS biology*, 11(2):e1001490, 2013.

[26] Katy C Kao, Linh M Tran, and James C Liao. A global regulatory role of gluconeogenic genes in Escherichia coli revealed by transcriptome network analysis. *The Journal of biological chemistry*, 280(43):36079–36087, October 2005.

[27] Tadeusz J Kawecki, Richard E Lenski, Dieter Ebert, Brian Hollis, Isabelle Olivieri, and Michael C Whitlock. Experimental evolution. *Trends in ecology & evolution*, 27(10):547–560, October 2012.

[28] Aisha I Khan, Duy M Dinh, Dominique Schneider, Richard E Lenski, and Tim F Cooper. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, 332(6034):1193–1196, June 2011.

[29] T Kingston and S J Rossiter. Harmonic-hopping in Wallacea's bats. *Nature*, 429(6992):654–657, 2004.

[30] M Le Gac, M D Brazas, M Bertrand, J G Tyerman, C C Spencer, R E W Hancock, and M Doebeli. Metabolic Changes Associated With Adaptive Diversification in Escherichia coli. *Genetics*, 178(2):1049–1060, February 2008.

[31] Mickael Le Gac and Michael Doebeli. Epistasis and frequency dependence influence the fitness of an adaptive mutation in a diversifying lineage. *Molecular Ecology*, 19(12):2430–2438, June 2010.

[32] Richard E Lenski and Michael Travisano. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences*, 91(15):6808–6814, 1994.

[33] J B Losos. Contingency and Determinism in Replicated Adaptive Radiations of Island Lizards. *Science*, 279(5359):2115–2118, March 1998.

[34] Michael Lynch, Bruce Walsh, et al. *Genetics and analysis of quantitative traits.* Sinauer Sunderland, 1998.

[35] Guillaume Martin, Santiago F Elena, and Thomas Lenormand. Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nature Genetics*, 39(4):555–560, April 2007.

[36] Ernst Mayr et al. Animal species and evolution. *Animal species and their evolution.*, 1963.

[37] Anita H Melnyk and Rees Kassen. Adaptive landscapes in evolving populations of Pseudomonas fluorescens. *Evolution*, 65(11):3048–3059, November 2011.

[38] Johan AJ Metz, Stefan AH Geritz, Géza Meszéna, Frans JA Jacobs, and JS Van Heerwaarden. Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. *Stochastic and spatial structures of dynamical systems*, 45:183–231, 1996.

[39] J R Meyer, D T Dobias, J S Weitz, J E Barrick, R T Quick, and R E Lenski. Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda. *Science*, 335(6067):428–432, January 2012.

[40] F Pelletier, D Garant, and A P Hendry. Eco-evolutionary dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1523):1483–1489, May 2009.

[41] Patrick C Phillips. Epistasis — the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews Genetics*, 9(11):855–867, November 2008.

[42] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.

[43] Fisher RA. The genetical theory of natural selection, 1930.

[44] P B Rainey and M Travisano. Adaptive radiation in a heterogeneous environment. *Nature*, 394(6688):69–72, July 1998.

[45] Susanna K Remold and Richard E Lenski. Pervasive joint influence of epistasis and plasticity on mutational effects in Escherichia coli. *Nature Genetics*, 36(4):423–426, April 2004.

[46] Daniel E Rozen, Dominique Schneider, and Richard E Lenski. Long-Term Experimental Evolution in Escherichia coli. XIII. Phylogenetic History of a Balanced Polymorphism. *Journal of Molecular Evolution*, 61(2):171–180, June 2005.

[47] Gerda Saxer, Michael Doebeli, and Michael Travisano. The Repeatability of Adaptive Radiation During Long-Term Experimental Evolution of Escherichia coli in a Multiple Nutrient Environment. *PLoS ONE*, 5(12):e14184, December 2010.

[48] Dolph Schluter, Elizabeth A Clifford, Maria Nemethy, and Jeffrey S McKinnon. Parallel evolution and inheritance of quantitative traits. *The American Naturalist*, 163(6):809–822, June 2004.

[49] T W Schoener. The Newest Synthesis: Understanding the Interplay of Evolutionary and Ecological Dynamics. *Science*, 331(6016):426–429, January 2011.

[50] Daniel Segre, Alexander DeLuna, George M Church, and Roy Kishony. Modular epistasis in yeast metabolism. *Nature genetics*, 37(1):77–83, 2004.

[51] Pedro Simões, Josiane Santos, Inês Fragata, Laurence D Mueller, Michael R Rose, and Margarida Matos. How repeatable is adaptive evolution? The role of geographical origin and founder effects in laboratory adaptation. *Evolution*, 62(8):1817–1829, August 2008.

[52] P D Sniegowski, P J Gerrish, and R E Lenski. Evolution of high mutation rates in experimental populations of E. coli. *Nature*, 387(6634):703–705, June 1997.

[53] Christine C Spencer, Melanie Bertrand, Michael Travisano, and Michael Doebeli. Adaptive diversification in genes that regulate resource use in Escherichia coli. *PLoS Genetics*, 3(1):e15, January 2007.

[54] Christine C Spencer, Jabus Tyerman, Melanie Bertrand, and Michael Doebeli. Adaptation increases the likelihood of diversification in an experimental bacterial lineage. *Proceedings of the National Academy of Sciences*, 105(5):1585–1589, February 2008.

[55] Alejandro Toro-Roman, Timothy R Mack, and Ann M Stock. Structural analysis and solution studies of the activated regulatory domain of the response regulator ArcA: a symmetric dimer mediated by the alpha4-beta5-alpha5 face. *Journal of molecular biology*, 349(1):11–26, May 2005.

[56] Jabus Tyerman and Michael Doebeli. Variation in the propensity to diversify in experimental populations of Escherichia coli: consequences for adaptive radiation. *Evolutionary Ecology Research*, 12:803–820, 2010.

[57] Jabus Tyerman, Naomi Havard, Gerda Saxer, Michael Travisano, and Michael Doebeli. Unparallel diversification in bacterial microcosms. *Proceedings Biological sciences / The Royal Society*, 272(1570):1393–1398, July 2005.

[58] Jabus G Tyerman, Melanie Bertrand, Christine C Spencer, and Michael Doebeli. Experimental demonstration of ecological character displacement. *BMC Evolutionary Biology*, 8(1):34, 2008.

[59] Sara Via. Sympatric speciation in animals: the ugly duckling grow up. *Trends in Ecology & Evolution*, 16(7):381–390, 2001.

[60] Daniel M Weinreich, Nigel F Delaney, Mark A Depristo, and Daniel L Hartl. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, 312(5770):111–114, April 2006.

[61] R J Woods, J E Barrick, T F Cooper, U Shrestha, M R Kauth, and R E Lenski. Second-Order Selection for Evolvability in a Large Escherichia coli Population. *Science*, 331(6023):1433–1436, March 2011.

[62] S Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics*, 1(6):356–366, 1932.
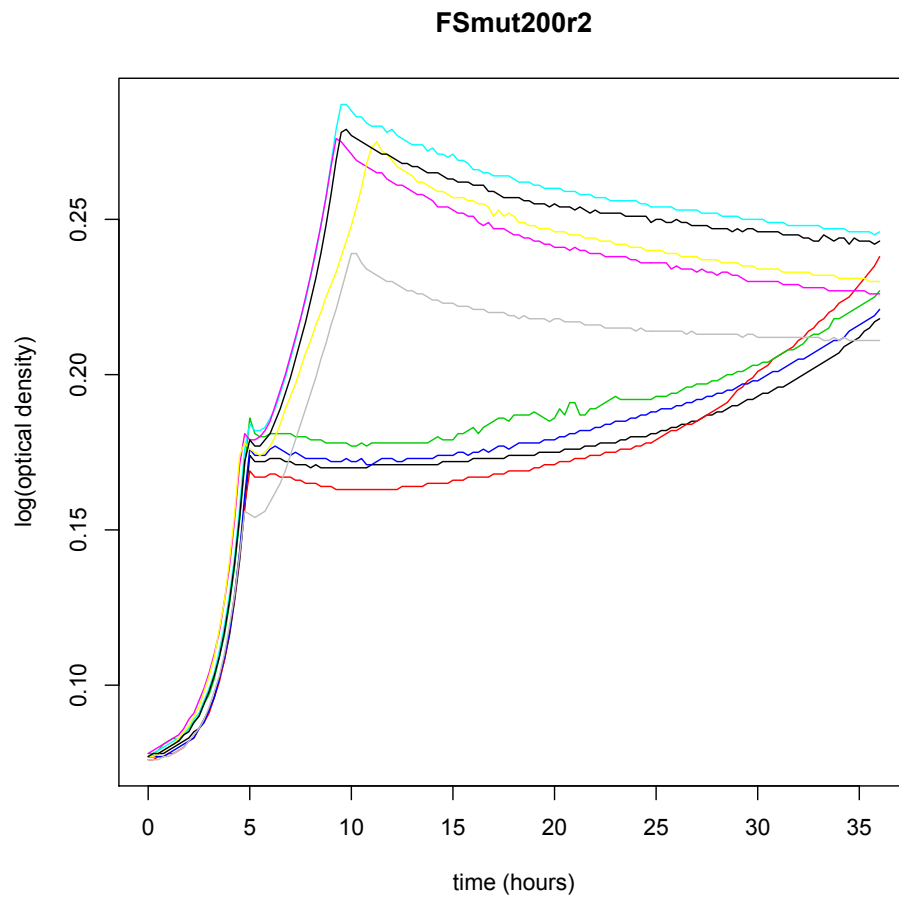
Appendix A

Supplementary Figures

FSmut200r2

Figure A.1: Growth curves of clones from gen200-FS$arcA^-$-rep2. Optical density of individual clones measured every 15 minutes over a 36 hour period.
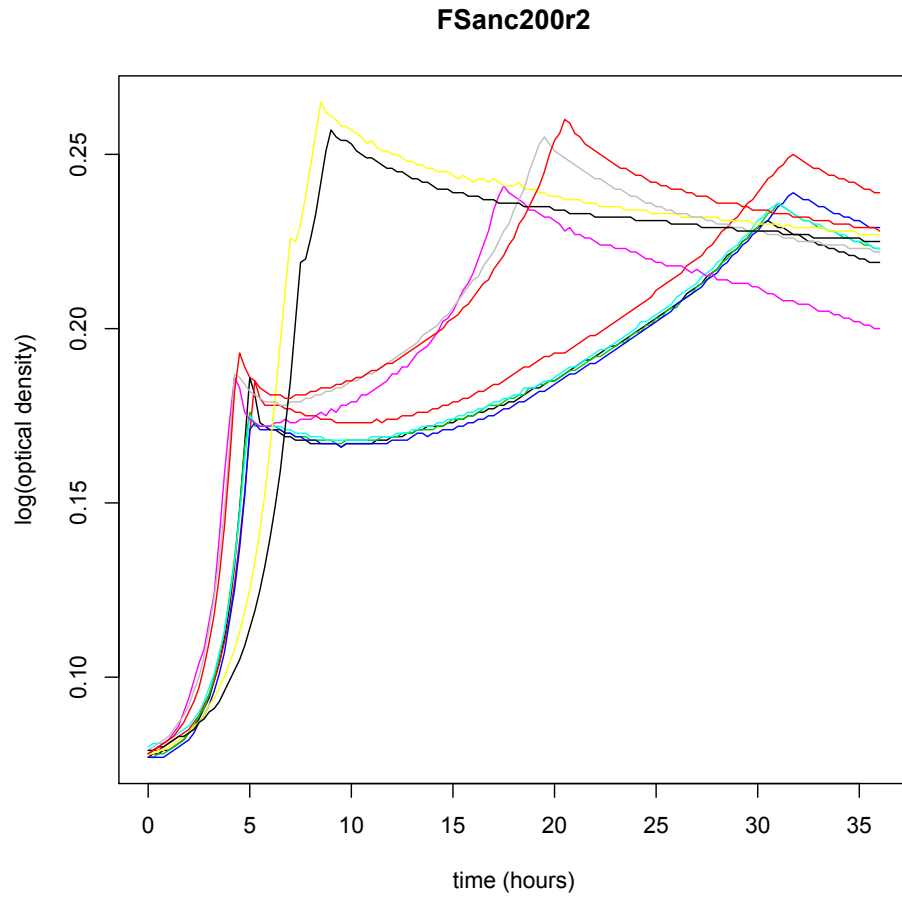
Figure A.2: Growth curves of clones from gen200-FS$arcA^+$-rep2. Optical density of individual clones measured every 15 minutes over a 36 hour period.
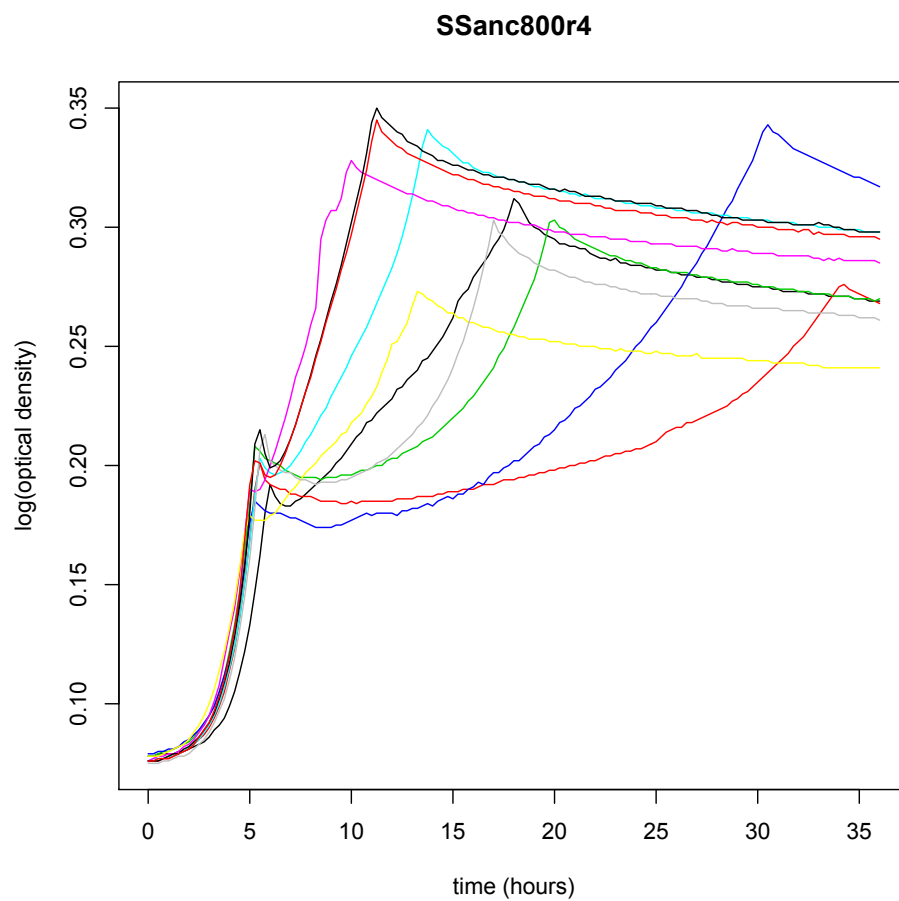
SSanc800r4

Figure A.3: Growth curves of clones from gen800-SS$arcA^+$-rep4. Optical density of individual clones measured every 15 minutes over a 36 hour period.