Competition and Coexistence in Yeast Experimental Evolution

Competition and Coexistence in Yeast Experimental Evolution

A DISSERTATION PRESENTED
BY
EVGENI MIKHAILOVICH FRENKEL
TO
THE COMMITTEE ON HIGHER DEGREES IN BIOPHYSICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
BIOPHYSICS

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
MAY 2016

Thesis advisor: Professor Michael M. Desai Evgeni Mikhailovich Frenkel

Competition and Coexistence in Yeast Experimental Evolution

Abstract

Natural selection gives rise to biodiversity by purging the less-fit among variants that are too similar (a principle known as character displacement), but to predict how fit or different an organism needs to be to survive is hard. In the simplest theoretical case, the probability whether one lineage versus another survives depends only on their relative fitness and random fluctuations. In more complex scenarios, this probability may depend on the fitness of all the other lineages in the population, mutations that these and other lineages acquire before the outcome of competition is decided, and additional ecological interactions. These complexities evolve readily in laboratory microbial populations, suggesting that they are the norm in Nature, and have been extensively studied theoretically. This thesis provides one of the few empirical examples in which the evolution and mechanism of some of these complexities have been characterized and modeled sufficiently to make basic predictions, such as whether a mutation will fix or go extinct, which competing lineages may or may not coexist, and how do these processes relate? This work was carried out in an established system for experimental evolution, populations of asexual budding yeast (*S. cerevisiae*) in microtiter plates.

Chapter 2 demonstrates an experimental design and modeling approach to infer the distribution of fitness effects of beneficial mutations from the population-dynamics of genetic markers. The inferred distribution accurately predicts fixation probabilities of lineages and adaptation rates of populations. Chapter 3 describes a new example of spontaneously-evolved coexistence between types competing for the same resources, including the physical mechanism, genetic basis and a mathematical model of the coexistence. The conclusion provides additional analyses to connect the findings from these two chapters and discusses their implications for microbial evolution more generally and directions for future work.

Citations to Previous Work

Chapter 2 appears in its entirety as

> "The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations," EM Frenkel, BH Good, MM Desai. *Genetics*, 196(4), pp.1217-1226 (2014)

Chapter 3 appears almost in its entirety as

> "Crowded growth leads to the spontaneous evolution of semistable coexistence in laboratory yeast populations," EM Frenkel, MJ McDonald, JDV Dyken, K Kosheleva, GI Lang, MM Desai. *Proceedings of the National Academy of Sciences*, 112(36), pp.11306-11311 (2015).

Contents

1 Introduction **1**
- 1.1 What can be learned from experimental evolution? . . . 1
- 1.2 Dynamics of Adaptation . . . 6
- 1.3 Evolution of Coexistence . . . 11

2 The Fates of Mutant Lineages and the Distribution of Fitness Effects of Beneficial Mutations **16**
- 2.1 Introduction . . . 17
- 2.2 Materials & Methods . . . 20
  - 2.2.1 Strains . . . 20
  - 2.2.2 Experimental procedures . . . 20
  - 2.2.3 Simulations . . . 22
  - 2.2.4 DFE parameter estimation . . . 23
  - 2.2.5 Statistical tests . . . 25
- 2.3 Results . . . 26
  - 2.3.1 Tracking the fates of seeded lineages . . . 26
  - 2.3.2 Fates of seeded lineages reflect supply of competing beneficial mutations . . . 27
  - 2.3.3 DFE inferred from seeded lineage dynamics . . . 29
  - 2.3.4 Measurements of adaptation rate corroborate DFE inference . . . 32
- 2.4 Discussion . . . 34

3 Crowded Growth Leads to the Spontaneous Evolution of Semi-Stable Coexistence **39**
- 3.1 Introduction . . . 40
- 3.2 Results . . . 41
  - 3.2.1 Spontaneous Evolution of Stable Coexistence . . . 41
  - 3.2.2 Identification of Adherent and Bottom-Dweller Types . . . 43
  - 3.2.3 Role of Spatial Structure . . . 43
  - 3.2.4 Genetic Basis of Adherence . . . 44
  - 3.2.5 Dynamics of A-Type vs. B-Type Growth . . . 45
  - 3.2.6 Model of Coexistence . . . 46
  - 3.2.7 Interplay Between Ecological and Evolutionary Dynamics . . . 49
  - 3.2.8 Evolution of Filamentous Type that Disrupts Coexistence . . . 52

3.3 Discussion 53
3.4 Materials and Methods 57
3.4.1 Strains 57
3.4.2 Bulk Segregant Analysis 57
3.4.3 Cell Culture 58
3.4.4 Measurement of Lineage Frequencies and Relative Fitness 58
3.4.5 Direct Measurements of Maximal Growth Rates 59
3.4.6 Measurements of Equilibrium Frequency 59
3.4.7 Imaging 60
3.4.8 Modeling 60

4 Conclusion **63**

Appendix A Supplement to Chapter 2 **67**

Appendix B Supplement to Chapter 3 **72**

References **92**

Listing of figures

1.1 Early experimental evolution of protozoan populations in increasing temperature . . . 4
1.2 Fitness gains and mutations in a long-term *E. coli* evolution experiment . . 5
1.3 Schematic of fitness assay . . . 7
1.4 Marker-divergence data interpreted assuming periodic-selection . . . 9
1.5 Dynamics of adaptation when beneficial mutations are rare versus frequent 10
1.6 Schematic of the distribution of fitness effects . . . 11
1.7 General form of negative frequency-dependent selection . . . 14

2.1 Trajectories of seeded lineages . . . 28
2.2 Possible fates of seeded lineages . . . 29
2.3 Measured fates of seeded lineages as a function of their fitness . . . 30
2.4 Inferred DFE parameters . . . 32
2.5 Lineage dynamics data and simulations for $s_o = 2.8\%$ . . . 33
2.6 Rate of adaptation predicted and measured . . . 34
2.7 A screen for beneficial mutations . . . 37

3.1 Coexistence due to negative frequency-dependent fitness . . . 42
3.2 Adherence phenocopied using antifungal drugs . . . 44
3.3 Effect of crowding on growth of adherent and bottom-dwelling strains . . . 45
3.4 Population dynamics within growth cycles . . . 47
3.5 Predictably semistable coexistence . . . 50
3.6 Dependence of coexistence on fitness . . . 51
3.7 Competition of two adherent and one bottom-dweller strains . . . 52
3.8 Identification of filamentous F type . . . 53

4.1 Lineage dynamics with vs. without negative frequency-dependent selection 66

A.1 Lineage dynamics data and simulations for $s_o = 3.8\%$ . . . 68
A.2 Lineage dynamics data and simulations for $s_o = 5.0\%$ . . . 69
A.3 Lineage dynamics data and simulations for $s_o = 7.3\%$ . . . 70
A.4 Rate of adaptation data . . . 71

B.1 Coexistence due to negative frequency-dependent fitness (full data) . . . 73
B.2 Bottom-dwellers do not exhibit frequency dependent fitness . . . 74
B.3 Adherent vs. bottom-dweller dynamics within each growth cycle . . . 75

B.4 Growth curves of adherent and bottom-dwelling strains 76
B.5 Fitness in flat wells compared with the maximum growth rate in round wells 77
B.6 Competition between two adherent and bottom-dweller strain (full data) . 78
B.7 Transitivity of fitness among adherent strains 79

To my family.

Acknowledgments

My advisor, Michael Desai, once ceremoniously gave me an antique-ish razor because a year prior I vowed to a labmate not to shave until finishing "the paper" (Chapter 2). Throughout graduate school Michael has been an inspiring and attentive scientist-mentor. He made the highs higher and the lows shorter with his generous, free-flowing whiskey, coffee and some serious scientific guidance. I was incredibly fortunate to be his student and hope to make him proud in my future work. Andrew Murray also kept track of my progress, gave helpful suggestions, perspectives on science and was generally very supportive.

My labmates Alex Nguyen, Ben Good and Michael McDonald took time to teach me things and discuss my projects. They were like co-advisors, in addition to being great friends, as were Elizabeth Jerison, Ivana Cvijovic and Sergey Kryazhimskiy. David Van Dyken and Katya Kosheleva were also helpful colleagues and great volleyball teammates. Dan Rice and John Calarco were among the folks I looked forward to seeing any day. The Murray Lab generally took us little Desai Lab achievers under their wing, especially in the early years when we were a bunch of yeast amateurs. Beverly Neugeboren was especially kind and helpful, as were Quincey Justman, Melanie Müller, John Koschwanez, Gregg Wildenberg and a post-doc in our lab, Gabriel Perron. Patricia Rogers, who was responsible for the flow cytometers, was very kind and impressive at her job, also forgiving of my occasional absent-mindedness. Eric Solis, Patrick Stoddard, Vlad Denic and Erik Hom helped or worked with me on side projects and generally gave good advice, as did Chris Marx and Scott Wylie. Erel Levine was a great course instructor when I served as a teaching assistant. Several administrators were tremendously kind and helpful: Michele Jakoulov, Prof. Jim Hogle, Jenn Thomson, Karen Lachmayr and Bodo Stern.

Teachers in high school and college opened doors for me intellectually and practically. Without their support, I could not have made it this far. I especially thank Scott Rubin and Dwight Raulston at St. John's School in Houston; Josh Tenenbaum at MIT who participated in Canada/USA Mathcamp and Research Science Institute; Bob Austen, Bill Bialek, David Botstein, Konstantin Doubrovinski, Josh Rabinowitz, and Ned Wingreen at Princeton; and Rob Phillips at Caltech.

I especially thank my friends Phil, Adam, Antoni, Danny, Keisuke, Leon, and Michaele, as well as Alexandra, Alyssa, Anya, Arthur, Deborah, Reb Eitan, Lucas, Tal and Tom for moral support during graduate school.

Most importantly, I thank my parents, older brother and grandparents for raising me. My Dad drove me 20+ miles to and from highschool. My Brother taught me to stand, be confident and deal with setbacks, such as endlessly losing in ping-pong. My Mother taught me how to take care of myself and others. My Grandmother spoiled me, while my Grandfather imparted much, including that I wouldn't amount to anything if I didn't study. My nephew Leo has given me loads to look forward to as he grows up. After emigrating from the USSR, my family first settled in Denver, where the local Jewish community helped us assimilate. I thank them, especially Burt and Claire Seiden. Thank you also Yevgeny, Lisa, Julia and Maya Brudno for being the wonderful relatives I never knew I had. Finally, thank you Dunia for being my very special lady and bringing so much love and joy into my life.

1

Introduction

1.1 What can be learned from experimental evolution?

The general aim of evolutionary biology is to understand how organisms change over the course of generations, from the origin of life to the present and future. Laboratory evolution experiments provide relatively brief and simple instances of this process that can be dissected and rigorously understood. The motivations for this pursuit are three-fold: first, one may hope to derive principles of evolution that provide insight into the biodiversity of Nature. Secondly, different aspects of biology together determine any particular outcome of evolution; hence to understand and predict these outcomes can lead to a more-integrated understanding of biological systems. For example, populations of budding yeast

(*S. cerevisiae*) propagated at higher-than-normal temperatures are reproducibly taken over by mutants that have duplicated chromosome III.[137] Why chromosome III? It turns out that elevated expression of a cluster of genes on chromosome III improves growth at high temperature. Why not then only duplicate this cluster or elevate its expression specifically? That happens during subsequent evolution: the chromosome duplication reproducibly reverts, while elevated expression of certain genes remains. So why the aneuploid intermediate? This question points to a generalizeable principle: chromosomal rearrangements and copy number changes occur so frequently that when such mutations can produce a strongly favorable phenotype, they may typically take over a population before other kinds of beneficial mutations arise. Because such drastic genetic changes are likely to produce far-from-optimal solutions, they may also typically become partially reversed by subsequent evolution. In Nature, many fungal species are descendant from a whole genome duplication roughly 100 million years ago.[135,71] Why this ancient duplication was favored is uncertain, but it coincided with the evolution of angiosperms (plants with fruits), and the duplications that were retained suggest that it enabled yeasts to compete better for sugar.[21] Likewise in cancer, loss of tumor suppressors and activation of oncogenes often occurs through loss or duplication of chromosomes.[2] Together these examples illustrate how the study of cell biology and experimental and natural evolution inform one another.

Thirdly, we perform evolution in the laboratory to see it happen in detail.[15] One of Darwin's contemporaries, William Dallinger, was the first to realize that the rapid generation time of microorganisms could allow one to directly observe evolution in an experiment. He collected three species of plankton and grew them at increasing temperature for several years.[24] Initially, the organisms tolerated no more than 70°F, but eventually, he raised their growth temperature to 158°F and found that they now no longer survived at their natural temperature of 60°F. What can be learned from such an experiment? It was a foregone

conclusion, given the experimental design, that the organisms' thermotolerance would increase. Moreover, most organisms survive only in a relatively narrow range of conditions, so their loss of tolerance to low temperatures was not surprising. Darwin praised the work as a demonstration of how algae in hot springs are evolutionarily related to algae in cooler waters, but as Dallinger remarked, evolution did not need further proof. More interestingly, Dallinger periodically performed pilot experiments to determine the earliest moments and maximum extent to which he could raise the temperature without loss of viability and used a heated microscope stage to record morphological changes that accompanied these increases in thermotolerance (Fig. 1.1). His observation of increased vacuole number and size in all three species suggests that these morphological changes were adaptive. He also noted that the pace of adaptation, as indicated by these increases in temperature, varied between year-long periods when the temperature could not be raised at all and other periods when the populations tolerated increases of several degrees per month. Irregular, almost step-wise increases in stress tolerance have since been observed in other contexts, such as bacterial populations adapting to increasing antibiotic concentrations.[124] This pattern stands in contrast to observations from another evolution experiment started a century later by Richard Lenski, in which a dozen populations of *E. coli* have been propagated in batch culture since 1988 for over 64 thousand generations and counting.[134] This experiment established some general expectations for how adaptation proceeds in a constant environment: the pace of phenotypic change becomes progressively slower, but the pace of genetic changes may remain the same or even increase (Fig. 1.2), with the caveat that occasionally, these organisms evolve to change their environment or develop ecological interactions, initiating fresh phases of adaptation.[12,98]

The preceding examples illustrate some patterns of genetic and phenotypic changes in evolution, but what determines such patterns? The possible answers may be classed

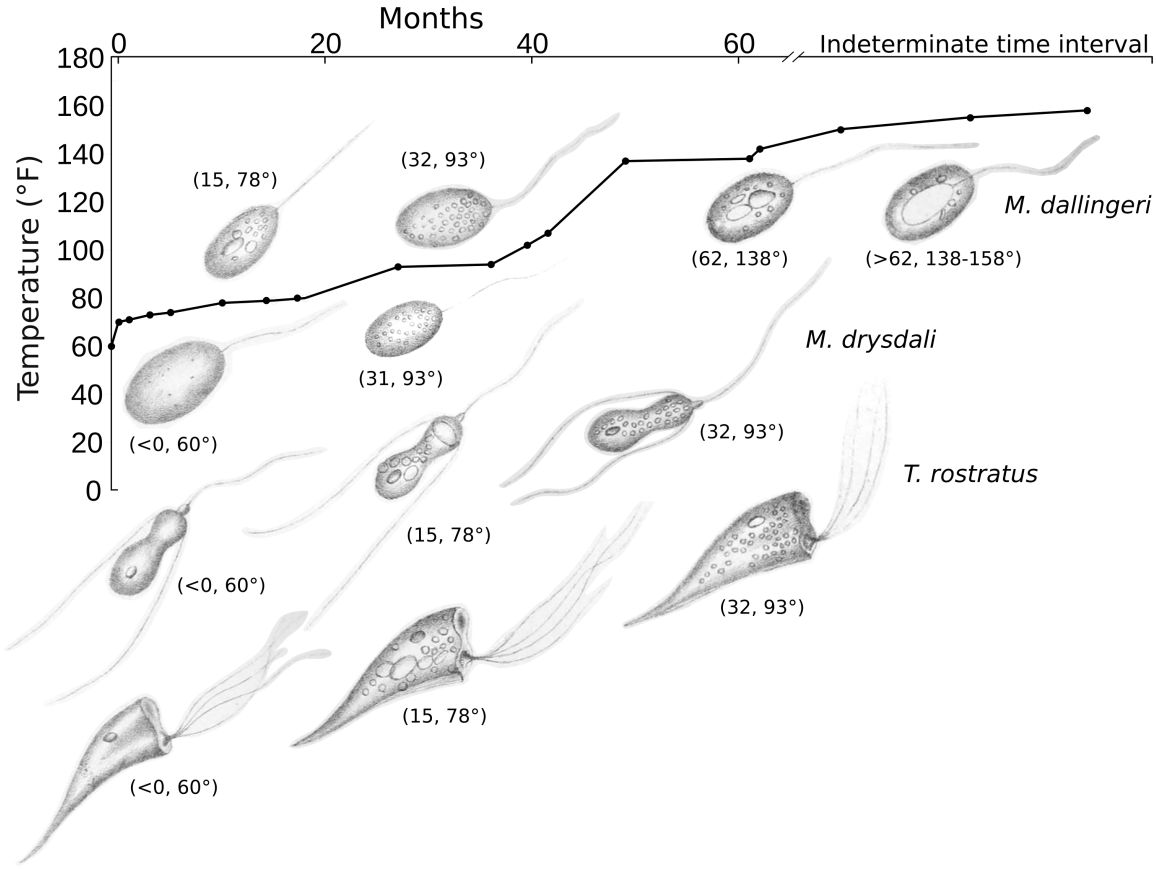

Figure 1.1: Increase in maximum-tolerable incubation temperature and morphological changes of three protozoan populations. The three species (*M. dallingeri, M. drysdali, T. rostratus*) were maintained in separate vessels at a common temperature that was raised at the maximum rate permitted by their viability. Each morphology was drawn at the (time point, temperature) indicated. The timing of the final three data points is unclear. Data from Dallinger[24].

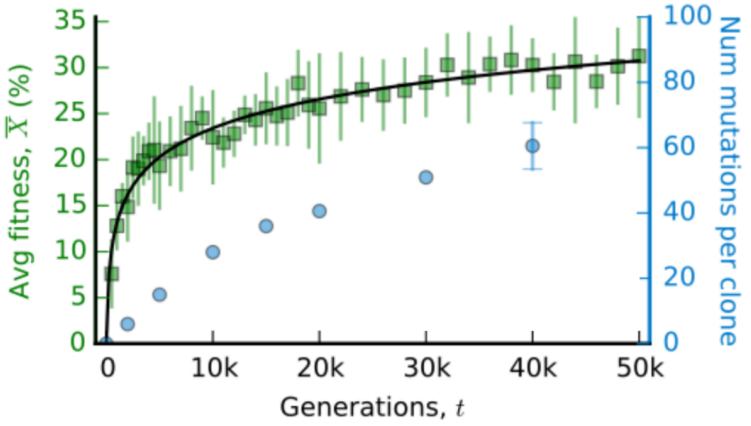

Figure 1.2: Fitness gains and number of fixed mutations in Richard Lenski's Long-term Evolution Experiment. Reprinted with permission from Good & Desai[48].

into two general categories: mathematical or population-genetic and biological, physical, chemical. For example in the case of yeast adapting to heat stress, one may predict that the observed pattern depends on population size. If the population were sufficiently larger, mutants with duplicated chromosome III might have been out-competed by rarer mutants that specifically activated the relevant genes on the chromosome. However, an increase in population size would also hasten the partial reversion of chromosome III duplication, so perhaps these double mutants would prevail. To predict the probabilities of these different outcomes, one needs to know how often different kinds of mutations arise and how quickly they spread through a population, a rate known as the selective effect or coefficient. This general process is referred to as the *dynamics of adaptation*, which is discussed in Sec. 1.2 below. Alternatively, one might wonder what biological, physical, chemical changes enabled cells to grow faster at higher temperature? In this case, altered sterol composition of the membrane appears to have conferred thermolerance.[17] There have been successes in bridging these two approaches, termed "functional synthesis."[30,35,34]

Experimental evolution can also provide insight into ecology. Populations adapt because offspring inherit traits from their parents, but not all offspring become parents themselves. Evolutionarily fitter traits are then those that increase the chances of producing more offspring, at least in the long run, and differences in traits arise by mutation. Hence over time, populations enrich for fitter lineages (an individual organism and its descendants). Given this process, how can so many different species, with vastly varied fecundities, survive in Nature? This is one of the basic questions of ecology. In the case of microorganisms, simple coexistence-ecologies have evolved *de novo* in laboratory environments, in which the hows and whys of their origins can be fully elucidated. These examples are discussed in Sec. 1.3 and Ch. 3.

This thesis takes an experimental-evolution approach to address two specific questions touched upon already: 1) How do the dynamics of adaptation determine which mutations in a population survive, and relatedly, how can we infer from these dynamics what mutations arose? 2) How can different types of microorganisms coexist while competing for the same resources? Some history and background to these questions is given in the next two sections, and the new results are described in chapters 2 and 3, respectively.

1.2 Dynamics of Adaptation

Mutations are random and therefore rarely beneficial, but microbial populations can be enormous: one milliliter of growth media may easily contain 10^8 yeast or 10^9 bacteria. The mutation rate is typically 10^{-9} - 10^{-10} per base pair per generation. Hence in a population of 10^8 yeast, every base pair has mutated on average once after roughly 20 generations or 1-2 days for laboratory strains given ample nutrients. So beneficial mutations may be extremely rare among all mutations and yet arise frequently in modest-sized laboratory

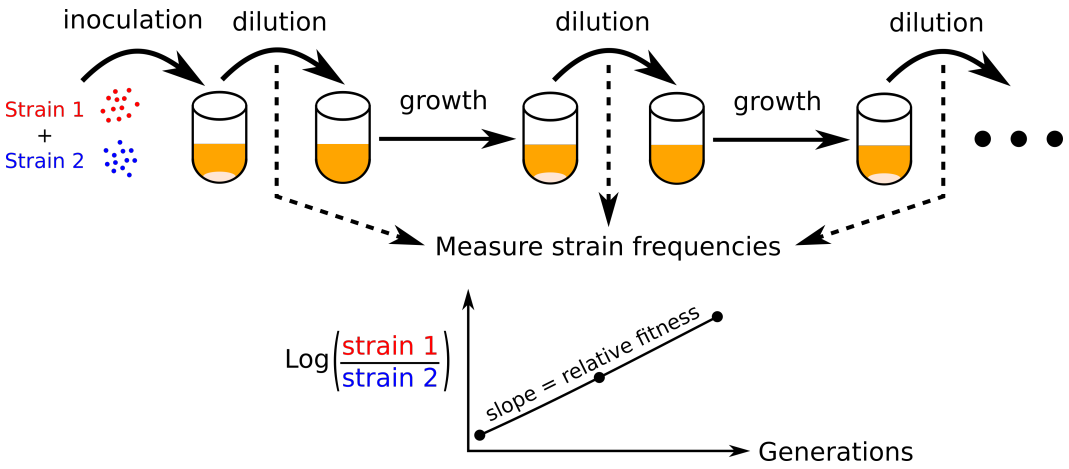

Figure 1.3: Schematic of competition assay to measure relative fitness. During the assay, populations are propagated using the same protocols and conditions as during longer-term experimental evolution.

microbial populations.

One might expect microorganisms to be near-perfectly optimized after billions of years of evolution, but typically, microbial species evolve to proliferate more quickly when propagated in laboratory conditions.[88,87] In other words, it is not necessary to apply a special stress or selection pressure to observe adaptation. As mentioned, an organism becomes more fit when it acquires beneficial mutations that enable it to produce more offspring; however, growth rate and fitness are not exactly the same thing. The difference is somewhat technical: fitness is the slope of a genotype's proportion in a population plotted through time, which empirically, is the time-average of the relative growth rate over some interval, measured using a competition assay (Fig. 1.3). Unless stated otherwise, the assumption here is that the relative fitness of genotypes is simply a constant, not dependent on their proportion or frequency, and that the population proliferates asexually, as most microorganisms usually do.

This definition of fitness implies two different ways one may observe the dynamics of adaptation: one is to compete an evolved population versus a reference strain, typically the ancestor of the population, which is how the fitness data in Fig. 1.2 were obtained, and the other is to track the frequency of genotypes in a population through time, an approach known as *marker divergence*. The first method reveals the average fitness increase of an evolved population but provides little information about variation in fitness within the population or the fitness effects of the mutations responsible for adaptation. To see beneficial mutations spread through a population, Novick & Szilard performed the first marker divergence experiment by tracking the proportion of bacteria resistant to phage in a chemostat.[103] Less than ten years prior, Luria & Delbruck had shown that phage resistance was caused by mutations,[94] and here, these mutations had little or no fitness effect because the chemostat was free of phage. Instead the resistance phenotype served as a convenient genetic marker to track the dynamics of adaptation in a population. Some of those data are shown in Fig. 1.4.

For nearly half a century, marker divergence data were interpreted in terms of periodic selective sweeps,[4] as illustrated in Fig. 1.4. The key assumption was that beneficial mutations must be so rare that populations essentially alternate between two distinct states: waiting for a beneficial mutation, during which time neutral genetic diversity accumulates, and experiencing a selective sweep, when this diversity is purged (Fig. 1.5A). Several implications follow from this assumption: except during relatively brief periods of selective sweeps, all genetic diversity within a population is neutral or deleterious; adaptation rate scales linearly with mutation rate and population size; and the probability a mutation fixes or goes extinct is determined solely by the population size and its selection coefficient. Despite a few reports,[55,1] it was not widely appreciated until the late 1990s that these expectations are largely inaccurate, at least for laboratory microbial populations.[40,29] Instead,

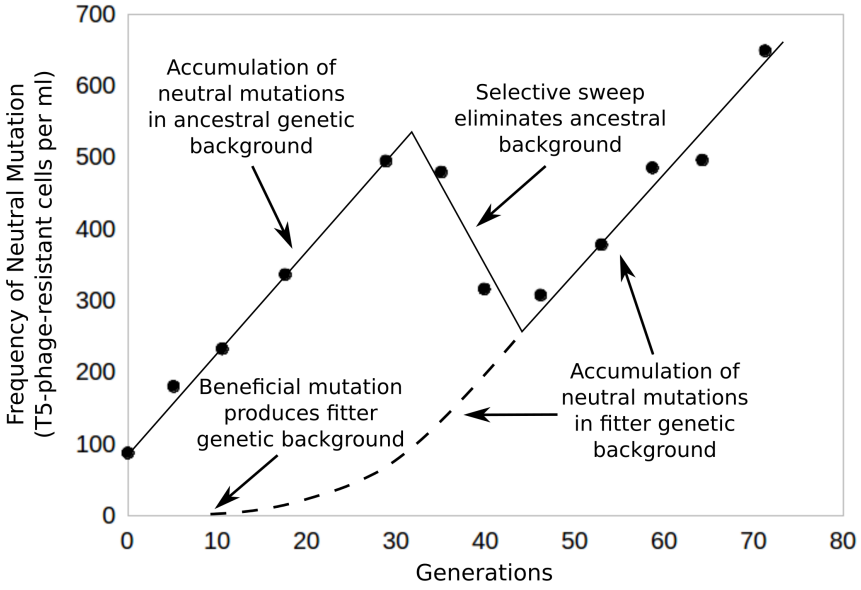

Figure 1.4: Marker-divergence data interpreted assuming periodic-selection (i.e. successive selective sweeps, as illustrated in Fig. 1.5A). Data from Novick & Szilard [103] and annotations adapted from Paquin & Adams [105].

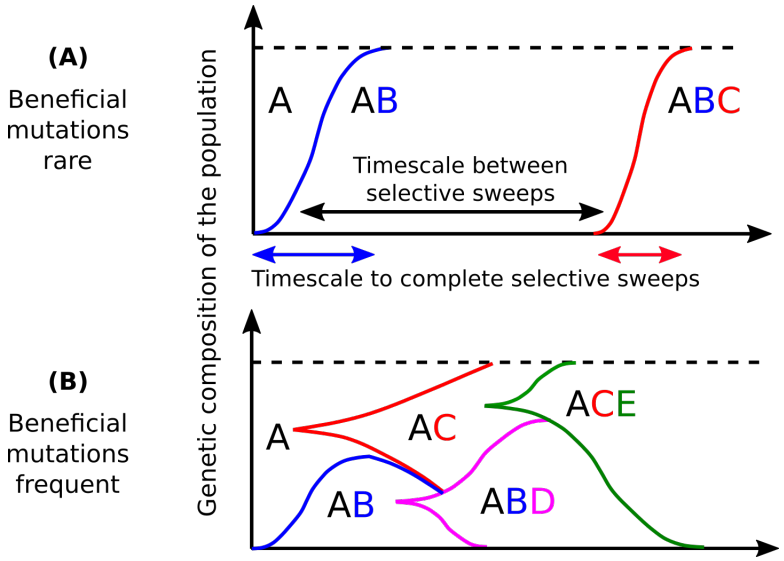

Figure 1.5: Muller diagrams illustrating the dynamics of adaptation when beneficial mutations are rare (A) versus frequent (B) in an asexual population with initial genotype A. Letters B-E correspond to newly-arising beneficial mutations that escape genetic drift.

beneficial mutations occur frequently enough that microbial populations routinely contain multiple beneficial mutations segregating simultaneously (Fig. 1.5B). Hence the adaptation rate has a complex dependency on mutation rate and population size, as well as other factors, and the probability individual mutations fix or go extinct depends more weakly on their individual fitness effects and more on the effects of all the other mutations segregating in the population.[32,33,81]

The main aim of marker divergence experiments has been to learn about the frequency and effect sizes of beneficial mutations. Are one in a million mutations beneficial or one in a billion? Stated more precisely, what is the *distribution of fitness effects* of beneficial mutations, drawn schematically in Fig. 1.6? Answering this question when beneficial mutations occur frequently enough to produce complex dynamics of adaptation like those illustrated in Fig. 1.5B is not feasible using the traditional approach of Fig. 1.4 because to

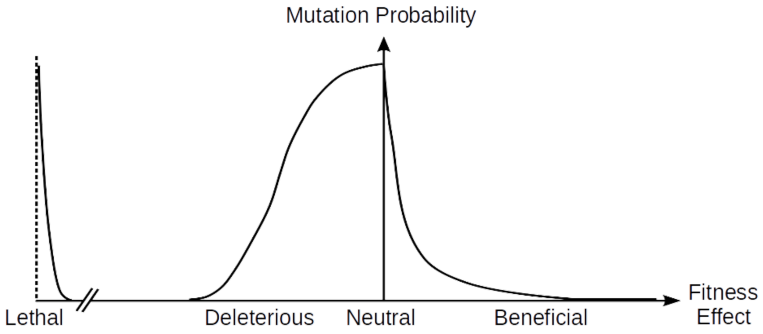

Figure 1.6: Schematic of the distribution of fitness effects of new mutations. In an adapting population, selection amplifies rare beneficial mutations from the right of the distribution, giving rise to dynamics of adaptation like those in Figs. 1.2, 1.4, 1.5.

do so essentially requires inferring the dynamics of multiple genotypes in the population from the time-series of only one.[54] The work described in Ch. 2 provides a solution to this problem, based on tracking frequencies of labeled lineages through time much like prior marker divergence experiments but with a new experimental design (carefully chosen initial conditions) to make these dynamics more informative. Since then, others have developed DNA-sequencing-based method to track thousands of lineages simultaneously, which also solves the problem because the number of simultaneously-segregating beneficial mutations is typically far fewer.[91,65]

1.3 Evolution of Coexistence

Consider a flask of two microbial species or strains growing together in nutritive broth: each day we let them grow until the nutrients are exhausted; then we take a small aliquot from the flask and inoculate it into another one of fresh broth and repeat. After a long time, does one expect two species or one in the flask? This experiment is the competition assay illustrated in Fig. 1.3, so we expect one survivor and can calculate when that outcome would

be reached by measuring the change in proportion over a short interval: with a measured relative fitness s, population size N, and roughly equal initial proportion, the expected time to fixation is $\frac{1}{s}\ln(N)$ generations.

But over the years, many researchers (including the author) sometimes found two not one and wondered, why? More generally, the question is how can different species or types coexist while competing for the same resources? Such coexistence needs explanation because of the *competitive exclusion principle*,[52] which may be stated as follows: if two species are locked in competition such that the increase of one necessarily implies the decrease of the other, then extinction of one species is inevitable, unless prevented by some *mechanism of coexistence*. Even if the two are perfectly matched, random fluctuations are expected to bring one species to extinction. The competitive exclusion principle is tautological, so its usefulness is only that whenever it appears to be violated, that indicates something to be discovered.

For example, Hutchinson's 1961 article "Paradox of the Plankton" pointed out that the diversity of prototrophic marine microorganisms is vastly greater than the diversity of their resources, yet no mechanisms of coexistence among them were known.[61] This spurred many efforts to identify such mechanisms[19] or revisit the assumption that they were necessary, since in a dynamic natural environment species may tend toward extinction without ever getting there.[60] As a result, theoretical conditions permitting stable coexistence of N species on less than N resources were systematically cataloged through analysis of mathematical-ecological models.[19,3,127,90] The question of this line of research then became, when are these theoretical mechanisms of coexistence applicable? Direct efforts to answer this attempted to recapitulate natural ecologies in controlled laboratory environments,[18,121] but fortuitously, coexistence sometimes also evolved spontaneously in experiments.[97,68] To clarify, in these cases, populations are initially clonal and then become genetically diverse

through acquisition of mutations during the experiment, and coexistence emerges between the recently-diverged lineages in these populations. Ecologically, such coexistence is like that of species because these populations are asexual (at least in experiments so far), but one would not classify these coexisting lineages as different species but rather as diverged strains.

Theoretically-derived mechanisms of coexistence generally assume trade-offs in fitness between species or types. The empirical hallmark of coexistence is that one may isolate at least two types; mix them together at different proportions (e.g. 99:1 and 1:99); propagate the mixtures and observe convergence to some common intermediate proportion (e.g. 50:50)—a phenomenon known as *negative frequency-dependent selection* because the fitness of each type varies inversely with its prevalence in the population (Fig. 1.7). A mechanism of coexistence is a model that explains the observed pattern of negative frequency-dependent selection, thereby delineating the necessary conditions for coexistence. For example, the first reported case of spontaneously-evolved coexistence was in populations of *E. coli* propagated in flasks as described above and was explained by a model postulating a so-called gleaner-opportunist trade-off.[89,122,50] The basic hypothesis is that different adaptations may be required to increase growth rate when nutrients are abundant versus scarce, so that the opportunist type grows faster when the population is first inoculated into a flask, while the gleaner type grows relatively faster as nutrients become depleted. This intuitive explanation turns out to be incorrect. Trade-offs between growth at rich versus scarce nutrient concentrations have so-far been found to be too weak to support coexistence,[126] and instead such coexistence is commonly maintained by cross-feeding interactions, typically in which one type partially metabolizes the carbon source and secretes a secondary metabolite scavenged by the other.[56,112,126] The secretion of secondary metabolites is a common feature of microbial metabolism and not as wasteful as it seems because the secreted metabolites may be

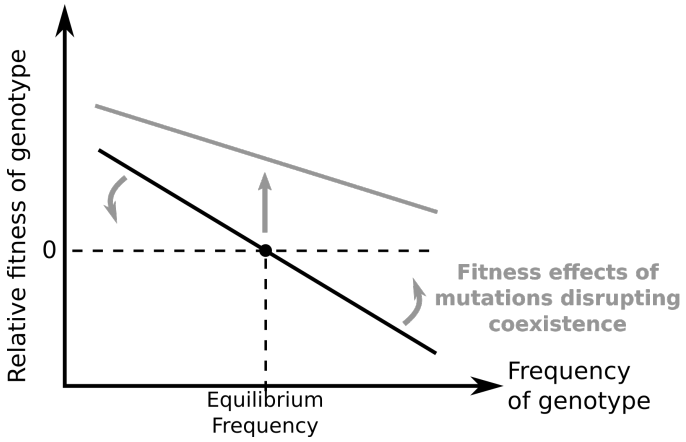

Figure 1.7: General form of negative frequency-dependent selection. A genotype whose fitness is positive when rare and negative when common, as illustrated by the solid black line, will tend to persist at a stable, intermediate equilibrium frequency. Mutations may push this equilibrium to either 0 or 1 by shifting the fitness-frequency relation as illustrated by the grey arrows and line.

taken up and consumed later.[37] Performing these different metabolic stages simultaneously within the same cell may be sub-optimal because different proteins are required.[8]

Trade-offs in fitness that enable coexistence can be eliminated by evolution to create one type uniformly better than the other. Hence the applicability of theoretically-derived mechanisms of coexistence depends on the strength of evolutionary constraints imposed by biophysical properties of organisms, their environments or other factors. Whenever coexistence evolves spontaneously, it suggests that perhaps the assumed trade-off has a biophysical basis or perhaps that one uniformly superior type is unlikely to arise for population-genetic reasons. For example in one of the populations of Richard Lenski's long-term evolution experiment, two cross-feeding types, L and S, have stably coexisted for more than 30,000 generations.[85] However the stability of the coexistence is revealed to be more tenuous when clones of each type isolated from different time points are competed against one another: S-type clones from earlier time points can be driven extinct by L-type clones from later

time points.[85] The specific adaptations that enable the L-type to sometimes eliminate the S-type are unknown, but conceptually, their effects may be either to weaken the trade-off or to increase the relative fitness of the L-type to such a degree that the trade-off cannot maintain coexistence, as illustrated in Fig. 1.7. The new example of spontaneously-evolved coexistence described in Ch. 3 also exhibits this property that two types stably coexist due to negative frequency-dependent selection, but adaptation of one may cause extinction of the other.

2

The Fates of Mutant Lineages and the Distribution of Fitness Effects of Beneficial Mutations

The outcomes of evolution are determined by which mutations occur and fix. In rapidly adapting microbial populations, this process is particularly hard to predict because lineages with different beneficial mutations often spread simultaneously and interfere with one another's fixation. Hence to predict the fate of any individual variant, we must know the rate at which new mutations create competing lineages of higher fitness. Here, we directly measured the effect of this interference on the fates of specific adaptive variants in laboratory *Saccharomyces cerevisiae* populations and used these measurements to infer the distribution of fitness effects of new beneficial mutations. To do so, we seeded marked lineages with different fitness advantages into replicate populations and tracked their subsequent frequencies for hundreds of generations. Our results illustrate the transition between strongly advantageous lineages which decisively sweep to fixation and more moderately advantageous lineages that are often outcompeted by new mutations arising during the course of

the experiment. We developed an approximate likelihood framework to compare our data to simulations and found that the effects of these competing beneficial mutations were best approximated by an exponential distribution, rather than one with a single effect size. We then used this inferred distribution of fitness effects to predict the rate of adaptation in a set of independent control populations. Finally, we discuss how our experimental design can serve as a screen for rare, large-effect beneficial mutations.

2.1 Introduction

Evolutionary adaptation is driven by the accumulation of beneficial mutations. There are two basic questions one can ask about this process. First, what are the set of mutations available to the population? That is, what is the overall mutation rate, U, and the distribution of fitness effects, $\rho(s)$, of new mutations? Second, what is the fate of those mutations that occur? In other words, how does the frequency of each mutation change over time until it eventually fixes or goes extinct?

When beneficial mutations are rare, these two questions are independent. Mutations of a given fitness effect, s, occur at rate $U\rho(s)$. The fate of each mutant is then decided entirely on its own merits: it increases in frequency (or is lost due to random drift) at a rate commensurate with its selective effect. Experiments, however, have shown that even for modestly sized laboratory populations of viruses and microbes, multiple beneficial mutations often spread simultaneously and interfere with one another, an effect known as *clonal interference*[29,101,66,107,33,67,86] (see Sniegowski & Gerrish[120] for a recent review). This means that the fate of each beneficial mutation depends not only on its own effect, but also on its interactions with the rest of the variation in the population.[81,83] In this regime, the mutation rate and the distribution of fitness effects of beneficial mutations (the DFE,

$\rho(s)$) controls the availability of competing mutations, which then play an important role in determining the fate of each new beneficial mutation.[40,49]

These factors highlight the importance of the DFE as a central parameter in adaptation, determining which new mutations occur and influencing their subsequent fate. Some theoretical work has argued that the DFE will typically be exponential.[44,104] However, this is fundamentally an empirical question, and in principle the details of the DFE could be highly system-specific. There has thus been extensive experimental effort devoted to measuring the DFE of beneficial mutations in a variety of laboratory populations[63,113,115,5,69,16,107,111,96,9,99] (a separate literature has used population genetic methods to infer the DFE in natural populations, reviewed by Keightley & Eyre-Walker[70]).

Experimental efforts to measure the DFE of beneficial mutations in laboratory populations have largely taken one of two complementary approaches. The first approach is to isolate mutants and directly assay their fitness. The difficulty with this method is that beneficial mutations are rare, so many clones must be screened to isolate comparatively few beneficial mutations.[115,69] To avoid this difficulty, some studies have imposed a harsh selection and studied the survivors, which by definition must have a beneficial mutation.[96,99] However, this approach is limited to harsh and typically narrow stresses (e.g. treatment with antibiotic), which may not be representative of adaptation to other conditions.

The second common experimental approach is to track the frequencies of genetic markers over time, and use the resulting dynamics to infer the underlying DFE. Such "marker divergence" experiments typically use two or more strains that differ by a single neutral genetic marker which can be easily tracked through time (e.g. antibiotic resistance or a fluorescent reporter). These strains are mixed, usually in equal proportions, and allowed to evolve in competition. The changes in frequencies of the neutral markers then reflect subsequent beneficial mutations that occur in one or the other genetic back-

ground.[103,4,55,105,1,63,5,28,107,54,67,6,81,45,7] Inferring the DFE from such data typically requires estimating the fitness effects of many mutations from the dynamics of relatively few markers, which is naturally quite difficult.[54,108,138,62,27] In principle, this difficulty could be removed by reducing the population size to such a degree that only one or zero beneficial mutations usually arise in each population.[107] However, this requires careful tuning of the population size, in order to make it small enough to minimize multiple mutations but also large enough to ensure that many replicates acquire a beneficial mutation.

Here, we introduce a twist on the traditional design of marker divergence experiments that produce dynamics more directly revealing of the underlying DFE. Rather than using neutral markers, we tracked the frequencies of marked lineages with a fitness *advantage* relative to a reference strain. We seeded these marked lineages at low frequency into populations of the reference, so that their subsequent dynamics are reflective of the fates of beneficial mutations with a particular selective advantage. Since the DFE controls the availability of competing mutations and hence the likelihood of clonal interference, we can exploit the observed fates of seeded lineages to infer the DFE. Using lineages with different fitness advantages enabled us to probe different corresponding portions of the DFE. This approach is particularly suited to infer those aspects of the DFE that are most important in determining the fates of new beneficial "driver" mutations, e.g. the high-fitness tail, which is otherwise hard to measure directly. In the process, we also directly measured how clonal interference alters a key quantity in adaptation: the fixation probability of a beneficial mutation as a function of its fitness effect.

2.2 Materials & Methods

2.2.1 Strains

All strains used in this study were derived from the base strain DBY15084, a haploid *S. cerevisiae* strain derived from the W303 background with genotype MAT**a**, ade2-1, CAN1, his3-11 leu2-3, 112, trp1-1, URA3, bar1Δ::ADE2, hmlαΔ::LEU2. Each experimental population included a resident and a seeded lineage. The resident lineage was DBY15108, a derivative of DBY15084 in which the fluorescent protein ymCherry was integrated at the URA3 locus.[81] The seeded lineages were descendants of strain DBY15104 isolated from timepoints of an earlier long-term evolution experiment.[81] To allow us to track their frequency using flow cytometry, we amplified a pACT1-ymCitrine pTEF-HISMX6 cassette from plasmid pJHK043 (provided by John H. Koschwanez) and integrated it at the HIS locus using oligos oGW137 (5'-TTGGTGAGCG CTAGGAGTC-3') and oGW138 (5'-TATGAAATGCTTTTCTTGTTGTTCTTACG-3') provided by Gregg Wildenberg. From this pool of transformants, we selected strains EFY11-17 based on fitness assays described below.

2.2.2 Experimental procedures

To obtain seeded lineage strains with a range of fitnesses, we isolated a large number of evolved clones and assayed their fitnesses as described in Lang et al.[81] Briefly, this protocol is to mix each strain in roughly equal proportion with a reference strain that bears a different fluorescent reporter, propagate these mixed populations for 30 generations, and measure the ratio of the strains at generations 10 and 30 using flow cytometry. Relative fitness was calculated as $s = (1/20) \cdot \log(\text{final ratio}/\text{initial ratio})$. From among these clones,

we chose EFY11-17 to use as seeded lineages and remeasured their fitnesses in 10 replicates. These additional assays showed that strains EFY12-14 and EFY15-16 had indistinguishable fitnesses, and so for the purposes of analysis, strains EFY11, EFY12-14, EFY15-16 and EFY17 were respectively grouped into the fitness classes indicated in Fig. 2.1.

To begin the evolution experiment, we grew up an individual resident clone to saturation in 3mL of standard growth media (YPD supplemented with 100 μg/mL ampicillin and 25 μg/mL tetracyclin). We transferred 128μL of this culture into each well of a 96 well-plate, diluted these cultures 2^{10}-fold into twelve 96-well plates containing fresh media, allowed these cultures to grow for 10 generations, and froze them at -80^{o}C in 15% glycerol. Later, these plates were thawed and propagated for 30 generations (as described below) to re-acclimate them to this environment. In parallel, we prepared the seeded clones in the same fashion. We then mixed seeded and resident populations to found a total of 1044 populations in twelve 96-well plates (see Table A.1). These populations were propagated at 30^{o}C in 128μl YPD per well and diluted every 24 hours by a factor of 2^{10} into new plates containing fresh media. This corresponds to an effective population size $N_e \approx 10^5$.[81,130] Each plate contained a set of 9 empty wells as cross-contamination controls. All control wells remained sterile throughout the experiment except for two accidents involving plate mixing. This contamination was resolved by restarting from glycerol stocks of an earlier time point. Transfers were carried out using a Biomek FX pipetting robot.

At approximately 50-generation intervals, seeded lineage frequencies were measured using flow cytometry. In particular, BD Biosciences Fortessa and LSR-II flow cytometers with high-throughput plate samplers counted ∼100,000 cells per population for the initial time point and ∼30,000 cells per population for time points thereafter. Repeated measurement of populations and blanks indicated that roughly ∼100 cell counts per sample were carry-over from previous samples. Therefore the uncertainty in frequency at the first time point was

$\sim$0.1% and $\sim$0.5% thereafter. These raw data were processed in FlowJo version 9.2. All processed data are provided in Table A.1.

We also assayed the fitness of 16 additional control populations founded with only the resident strain. To do so, these populations were thawed from frozen-archive plates, each was duplicated into 4 replicates, these were propagated for 30 generations to acclimate them, and then their fitness was assayed as described above.

Note that 386 of the populations were later excluded from analysis, leaving a total of 658 replicate populations, apportioned among the seven seeded lineages and controls as described in Table A.1. In 232 of these, frequency dependent selection emerged. We identified these by first investigating 15 populations in which lineages co-existed at constant proportion for hundreds of generations. We found that this co-existence was maintained by frequency dependent selection exclusively in populations having a characteristic pellet morphology (see Ch. 3), so we excluded from analysis all populations that also had this morphology. In the other cases, the initial frequency of the seeded lineage was so low that it could not be precisely determined or extinction due to drift was common. To exclude these without biasing the statistics of trajectories, we chose a cut-off for the initial frequency of each seeded strain such that in all replicates in which the initial frequency was above the cut-off, the seeded lineage rose to at least 5%. All replicates below the cut-off were excluded.

2.2.3 Simulations

For a given DFE, we simulated lineage trajectories using a forward-time algorithm designed to mimic the conditions of our experiment. Between each transfer, each cell expanded clonally for 10 generations at a deterministic exponential growth rate $r = r_0 + X$, where X is the fitness of the cell relative to the resident ancestor strain. At the transfer step,

the population was downsampled to $N_b = 10^4$ individuals with Poisson sampling noise. Mutations accumulate during the growth phase, but we assumed that they did not influence the fitness of the cells until the next transfer cycle. Thus, mutation was approximated by assuming that each individual has a probability $10U_b$ of gaining a beneficial mutation at the end of a transfer step, with additive fitness effects drawn from the underlying DFE. In order to speed computation, we binned the fitnesses of individual cells into discrete fitness classes of width $\Delta s = 0.01\%$ for all simulations except those in Fig. 2.7, which required information from individual mutations.

Each replicate simulation began at generation $t = 0$ with a homogeneous seed population with initial fitness s_0 and initial size $f_0 N_b$, and a resident population of size $(1 - f_0)N_b$, with s_0 and f_0 as measured experimentally. The initial genetic composition of each resident population was obtained by simulating deterministic growth from a single-cell to 3×10^8 cells, followed by a Poisson dilution down to N_b cells and four transfer cycles as described above. Simulated trajectories were then obtained by propagating the seeded lineage and the resident and recording the number of descendants of the seeded lineage at the same timepoints as the experiment, up to the time required for the fixation or first-peak used in the inference. Simulations for the rate of adaptation were carried out in a similar manner for populations consisting only of the resident (without the four transfer cycles prior to $t = 0$). A copy of our implementation is available upon request.

2.2.4 DFE parameter estimation

To determine the likelihood of the data for a particular set of DFE parameters, the 650 measured trajectories were partitioned into 13 classes such that the seeded lineages within each class shared the same initial fitness s_0 and differed in their initial frequency f_0 at most 2-4 fold (Table A.1). We classified each trajectory into one of 17 bins of $(f_{\text{peak}}, s_{\text{down}})$

values as described in the text. To estimate the relative probabilities of each of these bins, we simulated a large number of trajectories for each of the 13 seeded lineage classes and recorded the fraction of times that each trajectory bin was observed. The total likelihood of the data for a given set of DFE parameters was then estimated as the product of the trajectory bin probabilities for each of the 650 measured trajectories.

We determined the most-likely parameters for a particular DFE shape by scanning across a grid of U_b and $\bar{s}$ values, which was locally resampled at finer resolutions until the most-likely parameters could be identified with a reasonable level of confidence. We first simulated a coarse grid of parameter values with a mean rate of adaptation between 0 and 5% per 100 generations. We confirmed by visual inspection that the likelihood surface smoothly sloped toward the most-likely point identified in this coarse grid. We drew a rectangle around this peak and resampled points and adjusted the boundaries of this region until they satisfied the following criteria:

(1) Any infinitesimal area of the region contained at least one point whose likelihood uncertainty (due to the finite number of simulated trajectories) was less than 0.5 log-likelihood units (LLU). Here, infinitesimal areas correspond to 10% increments of U_b and 0.1% increments of $\bar{s}$. We estimated the uncertainty in the likelihood using the Wilson confidence interval[14] and employed a minimum of 10^4 simulated trajectories per parameter value.

(2) Each infinitesimal area on the border of the peak region contained a point with likelihood at least 10 LLU below the peak and whose uncertainty was less than 0.5 LLU. Once these criteria were met, the most-likely parameters for the candidate DFE shape were estimated to be the grid point with the highest likelihood value. We estimated the confidence regions in Fig. 2.4 by re-fitting the most-likely parameters for 10^4 bootstrapped datasets, which we obtained by resampling the observed trajectories with replacement in such a way that the total number of trajectories in each of the 13 trajectory classes was

preserved. Fig. 2.4 shows the scatter of parameters that were found to be most-likely for at least 1% of these bootstrapped data sets.

2.2.5 Statistical tests

We used a standard likelihood ratio test to evaluate whether the most-likely exponential DFE provided a significantly better fit than the most-likely δ-function DFE. To obtain the null distribution of the likelihood ratio, we simulated 10^4 data sets using the most-likely δ-function DFE and determined the most-likely parameters for each of these simulated datasets under the exponential and δ-function DFEs as described above. We then estimated the p value as the fraction of simulated data sets whose likelihood ratio was more extreme than the value obtained from the measured trajectories. A similar procedure was used to compare the exponential and truncated exponential DFEs, with the exponential DFE now taking the role of the null hypothesis.

To obtain an absolute measure of goodness-of-fit for the exponential and δ-function DFEs, we used the estimated maximum likelihood as a test statistic and generated 10^4 simulated datasets given the most-likely values of U_b and $\bar{s}$. We then estimated the p-value as the fraction of simulated datasets whose estimated maximum likelihood was lower than that of the actual data.

The significance of the slowdown in adaptation rate was assessed with a non-parametric bootstrap procedure. We generated 10^4 bootstrapped datasets obtained by resampling the 16 populations with replacement, and for each of these, further resampling from the four fitness measurements at each timepoint. The null distribution for the change in adaptation rate, Δv, was obtained by calculating the change in adaptation rate in each bootstrapped dataset and subtracting the observed value from the original data. We then estimated the p-value as fraction of bootstrapped datasets in which $|\Delta v|$ was greater than that of the

actual data.

2.3 Results

2.3.1 Tracking the fates of seeded lineages

Any beneficial mutation creates a new lineage that is more fit than the genetic background in which it arose. To systematically study the fates of such lineages, we prepared a set of fluorescently labeled haploid budding yeast strains (the *seeded lineages*) with measured fitness advantages, s_0, of approximately 3, 4, 5 and 7% relative to a closely related but separately labeled reference strain. We founded 658 replicate populations of the reference (the *resident*), and introduced one of the seeded lineages at low frequency into each replicate population. We propagated these populations asexually in batch culture for hundreds of generations at an effective population size of $N_e \approx 10^5$, measuring the frequency of the seeded lineage in each population approximately every 50 generations (see Methods). This allowed us to track the fate of the seeded lineages over time, as illustrated in Fig. 2.1.

Each seeded lineage was introduced at an initial frequency f_0 large enough that genetic drift is expected to be weak relative to natural selection (i.e. $f_0 \gg \frac{1}{Ns}$). In the absence of additional mutations, this implies that the frequency $f(t)$ of each seeded lineage should increase deterministically according to the logistic equation, $f(t) = \frac{f_0 e^{st}}{1+f_0(e^{st}-1)}$. This expectation is indicated by the dashed curves in Fig. 2.1. As is apparent from the figure, most seeded lineages initially conformed to this expectation (the exceptions are lineages whose initial frequencies were only several-fold greater than $\frac{1}{Ns}$, which is low enough that genetic drift could partially reduce their initial rate of increase). Subsequently, many lineages diverged into a variety of qualitatively distinct fates. Since both genetic drift and measurement errors are expected to be small relative to this divergence (see Methods), the

variation in the fates of seeded lineages indicates that their relative fitnesses were modified by new beneficial mutations arising during the experiment.

2.3.2 Fates of seeded lineages reflect supply of competing beneficial mutations

The trajectory of each seeded lineage provides information about the beneficial mutations that did (or did not) arise within the competing resident population. Consider for example the case where a seeded lineage of fitness s_0 peaks and then declines in frequency. This reflects a clonal interference event, where one or more new beneficial mutations in the resident population create a competing lineage with fitness greater than s_0 (see Fig. 2.2). By considering the range of outcomes in replicate populations, we can estimate the probability of these events (Fig. 2.3). A higher probability of clonal interference implies a larger supply of beneficial mutations that can generate successful competing lineages.

Comparing the fates of seeded lineages of different fitnesses provides additional insight into the mutations responsible for clonal interference. For example, the seeded lineage with fitness advantage $s_0 = 7\%$ always swept to fixation without any detectable deviation from the expectation in the absence of interference. In contrast, the lineage with $s_0 = 5\%$ swept in 84% of replicates. Together, these two results suggest that clonal interference in the $s_0 = 5\%$ case was primarily due to beneficial mutations in the resident that created competing lineages with fitness advantages between 5 and 7 percent. Extending this logic, comparing the fates of seeded lineages with $s_0 = 5$, 4, and 3 percent provides information about the probabilities that beneficial mutations create competing lineages of fitness between 4 and 5 percent and between 3 and 4 percent.

While this intuition is straightforward, quantitative inference of the DFE requires us to connect the rates of individual mutations with the fitnesses of competing lineages. This

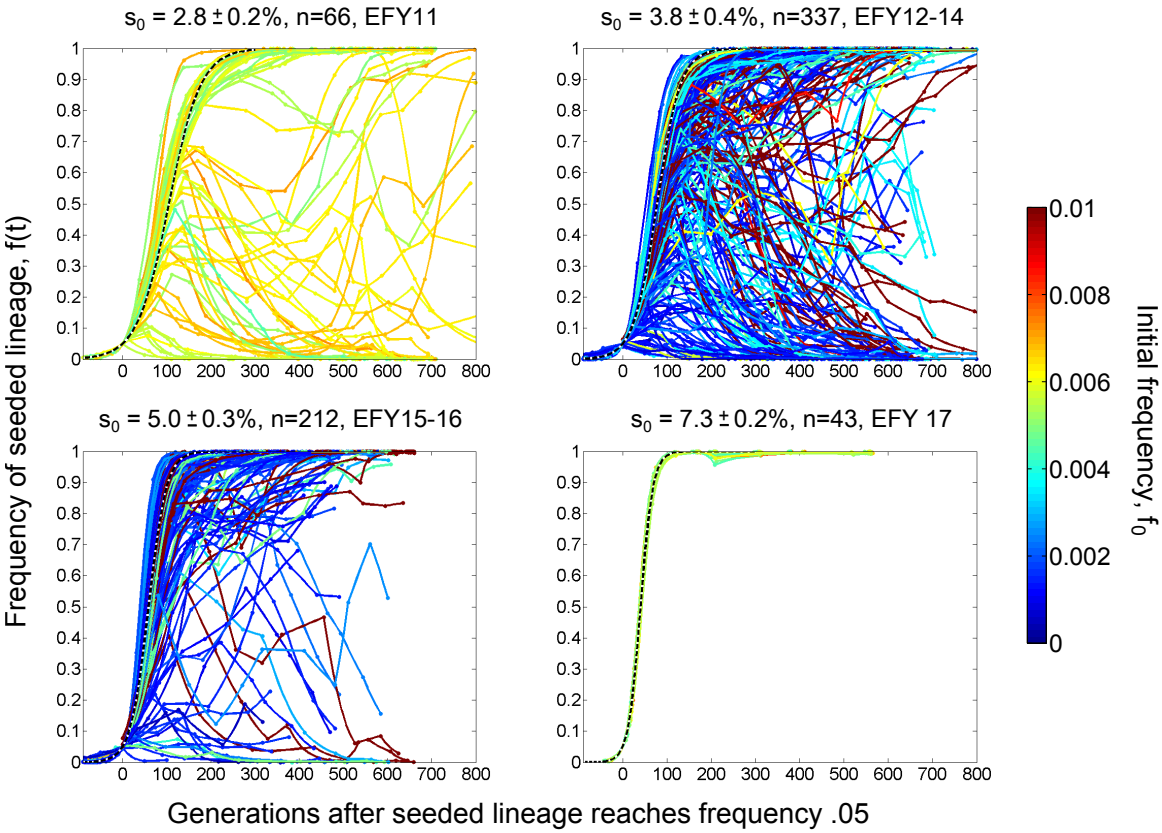

Figure 2.1: Trajectories of seeded lineages. Each line represents the frequency over time of a marked lineage with fitness advantage s_0 seeded into a replicate resident population. Colors correspond to the initial frequency f_0 of the seeded lineage according to the legend at right. Time is measured in generations, with $t = 0$ defined as the time at which each trajectory reached frequency 0.05. The dashed curves show the expected trajectories in the absence of new beneficial mutations (i.e. without clonal interference). (Note that the seeded lineages for $s_0 \approx 4\%$ and 5% consisted of multiple strains; see Methods).

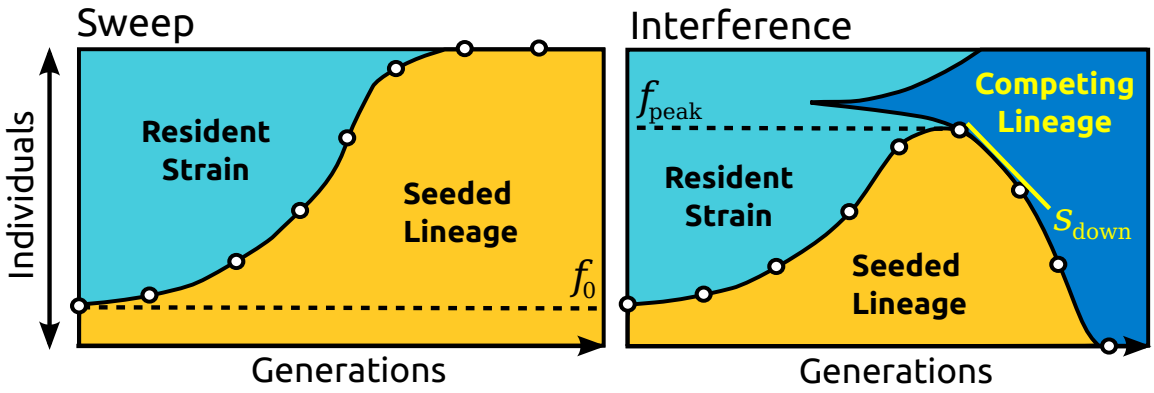

Figure 2.2: The fates of seeded lineages. We classified the trajectory of each seeded lineage according to whether it increased monotonically to fixation (a selective sweep, shown left) or peaked and subsequently declined in frequency (clonal interference, shown right). Each clonal interference event implies that the resident adapted fast enough to overtake the seeded lineage in fitness. These cases were further classified by the seeded lineage's peak frequency, f_{peak}, and relative fitness after this peak, s_{down}, as indicated in the schematic at right.

is complicated because competing lineages may often contain multiple beneficial "driver" mutations. In addition, beneficial mutations may also arise in seeded lineages, despite their initially much smaller population sizes. To fully account for these effects, we now introduce a computational method for inferring the DFE.

2.3.3 DFE inferred from seeded lineage dynamics

We implemented an approximate likelihood method which uses information from the shapes of the trajectories of seeded lineages to infer the DFE of beneficial mutations. Any particular trajectory only carries information about the beneficial mutations that rose to significant frequency in that population (i.e. the "contending" mutations; Rozen et al.[113]), but by modeling the trajectories of many populations together, we can learn about the overall distribution of possible beneficial mutations for the strains in our experiment. In order to make this inference tractable, we limited ourselves to single-parameter DFE shapes characterized

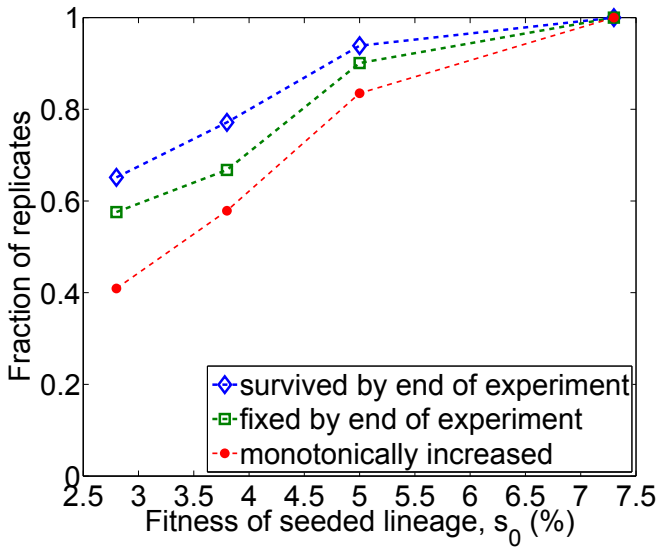

Figure 2.3: The fates of seeded lineages as a function of their fitness. We show the fraction of replicate populations in which the seeded lineage had the indicated fate.

by an average fitness effect $\bar{s}$ and beneficial mutation rate U_b. For concreteness, we considered three canonical distributions commonly used in the literature: an exponential DFE, $\rho_{\rm exp}(s) = \frac{1}{\bar{s}} e^{-s/\bar{s}}$, a uniform DFE $\rho_{\rm unif}(s) = \text{Heaviside}(2\bar{s} - s)/(2\bar{s})$, and a δ-function DFE where all beneficial mutations have the same fitness effect, $\rho_\delta(s) = \delta(s - \bar{s})$. We explain the significance of these choices in the Discussion.

To compute the likelihood of particular DFE parameters, we ran forward-time simulations of the experiment and estimated the likelihood as the fraction of replicate simulations that matched the data (see Appendix). In principle, we could use the complete trajectory of each seeded lineage for this comparison, identifying a match between simulations and data whenever the two were identical. However, in practice this was not computationally tractable. Instead, we focused on two features of the dynamics: the first peak frequency, $f_{\rm peak}$, of each seeded lineage (binned into quartiles, including fixed lineages) and the rate at which the seeded lineage declined in frequency following this peak, $s_{\rm down}$ (binned into

2% intervals). These are illustrated in Fig. 2.2. We chose to focus on these two quantities because we expect them to be particularly sensitive to the DFE: f_{peak} indicates how quickly a competing lineage arose in the resident population, while s_{down} measures how much the relative fitness of the resident population increased in this time. In addition, this focus on early-time dynamics ensures that most relevant mutations occur in the resident (due to its initially much larger population size), minimizing the effects of potential differences in the DFEs of the seeded genotypes.

For the three considered DFE shapes, we identified the most-likely parameters U_b and $\bar{s}$ by scanning a grid of candidate values. These parameters are shown in Fig. 2.4, along with confidence bounds estimated by bootstrapping (see Appendix). For each of these most-likely parameters, we show simulations of the $s_0 = 3\%$ seeded lineage trajectories in Fig. 2.5 and for the $s_0 = 4, 5$ and 7% lineages in Figs. A.1-3. Using a likelihood ratio test, we found that the exponential DFE provided a significantly better fit to the data than either the δ-function ($p < 10^{-4}$) or uniform distribution ($p < 10^{-4}$) and that the uniform provided a better fit than the δ ($p < 10^{-4}$).

Since the seeded lineage with $s_0 = 7.3\%$ always swept to fixation, indicating that larger-effect mutations must be rare, we checked whether truncating the high fitness end of the exponential DFE would improve its fit to the data. To do so, we considered an exponential DFE truncated at 7.3% and performed the same inference and statistical tests as above. We found that this truncated exponential provided a better fit to the data, but not significantly so ($p > 0.08$, likelihood ratio test). We also checked whether truncating the low-fitness end of the exponential would affect its fit to the data. We varied this truncation and found that, for the inferred exponential DFE parameters, discounting mutations with fitness effects below 2.1% improved these parameters' fit to the data, but only marginally so. This indicates that the seeded lineages were not strongly affected by mutations with fitness effects below

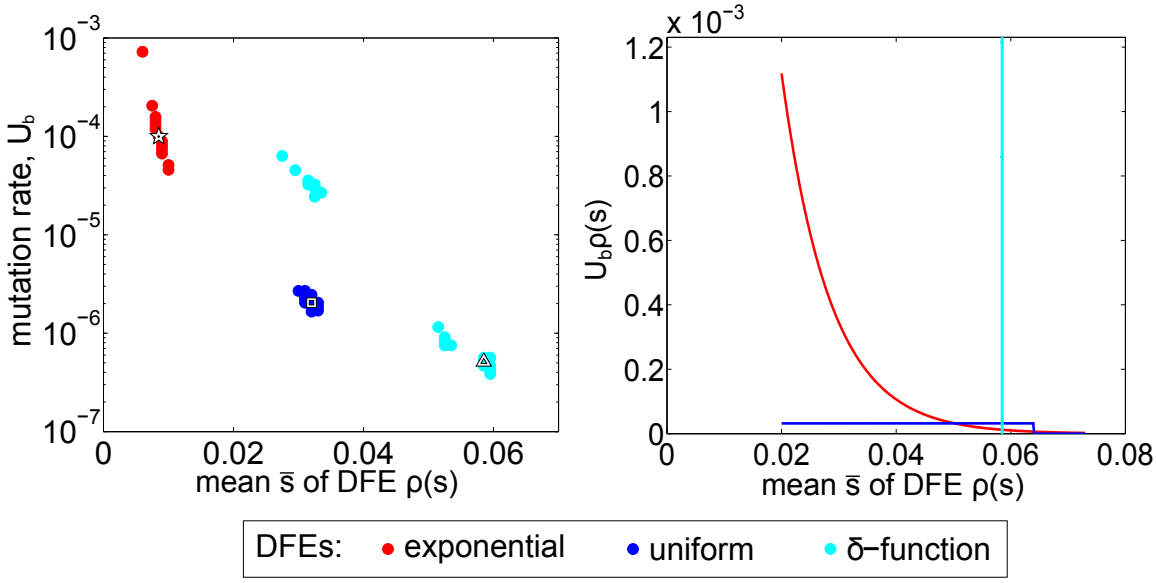

Figure 2.4: Inferred DFE parameters. Left: The most-likely parameters U_b and $\bar{s}$ for DFEs $\rho(s)$ given by exponential (star), uniform (square) and δ-function (triangle) distributions. Shaded circles indicate 1% confidence ranges of these parameters as estimated by bootstrapping (see Appendix). Right: The shapes of these distributions shown for $s \geq 2\%$ given their most-likely parameters.

$\sim 2\%$.

2.3.4 Measurements of adaptation rate corroborate DFE inference

In addition to determining the dynamics of seeded lineages, the DFE determines the rate of adaptation. Thus to test our inferences, we measured the changes in fitness over time of 16 control populations that consisted of the resident strain alone. We compared the average fitness of the control populations with the predictions of the most-likely exponential, uniform and δ-function DFEs. As seen in Fig. 2.6, the inferred exponential is fairly accurate in predicting these data, whereas the uniform and δ-function are less so.

Throughout our analysis, we have implicitly assumed that the DFE remained the same

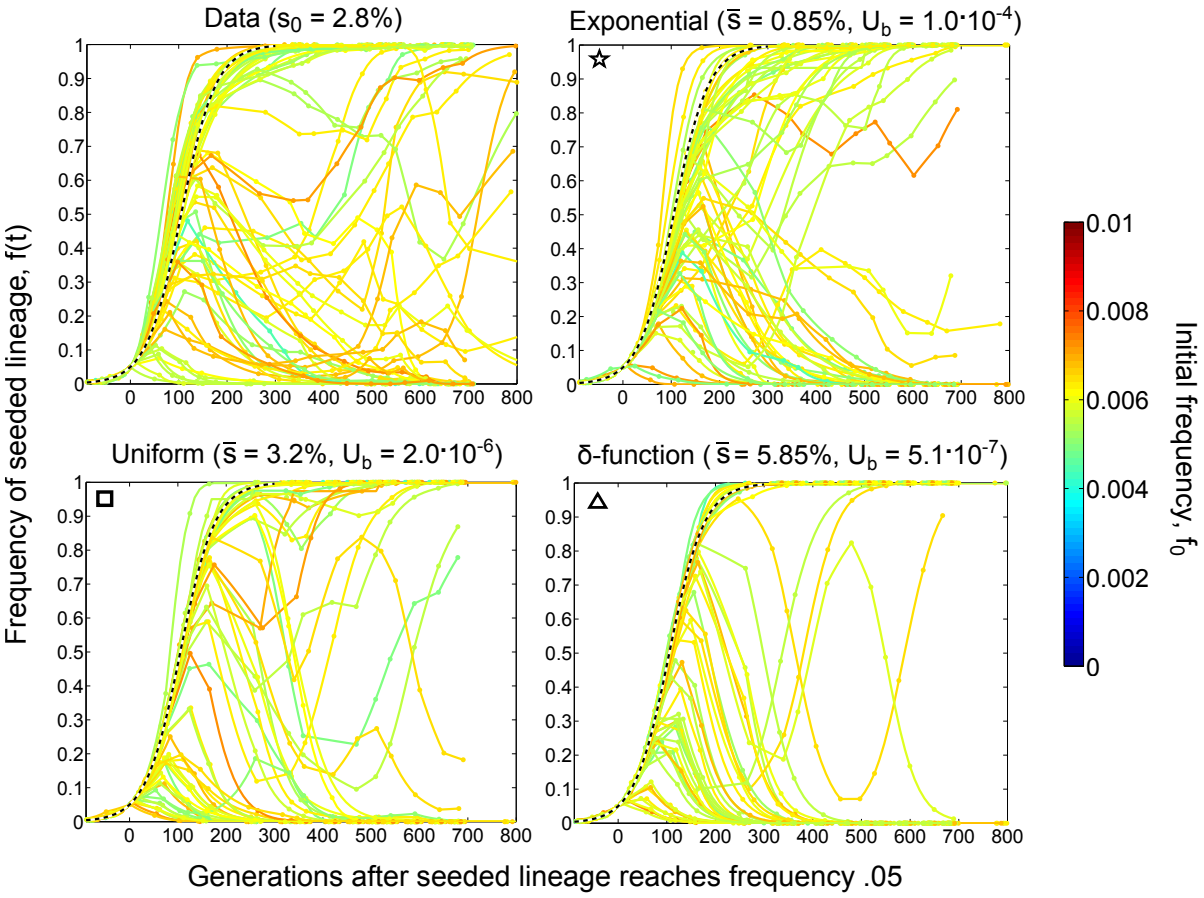

Figure 2.5: Lineage dynamics data and simulations for $s_0 = 2.8\%$. Each panel shows the trajectories of seeded lineages with initial fitness $s_0 = 2.8\%$ as observed in the experiment (top left) and as reproduced by simulations assuming the DFE parameters indicated above each panel.

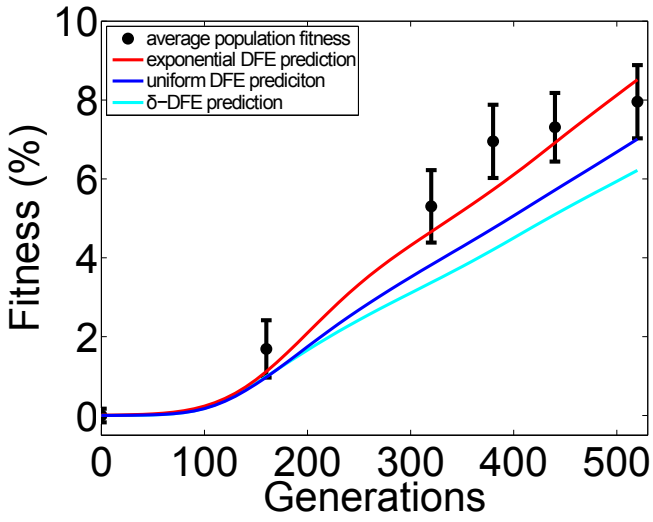

Figure 2.6: The rate of adaptation predicted and measured. The average fitness over time of 16 experimental control populations (± 1 s.e.m.) is shown in black. Solid curves are the predictions of these data given the most-likely exponential, uniform, and δ-function DFEs. The fitnesses of individual populations are shown in Fig. A.4.

across all genotypes in the experiment, which implies that the fitnesses of populations should increase linearly on average after some initial transient. In contrast, the rate of adaptation slowed after generation 380 ($p < 3 \cdot 10^{-3}$, see Appendix), which is reminiscent of declines in adaptation rate commonly observed in other evolution experiments[36]. Fortunately, we based our DFE inference on the early features of seeded lineage dynamics, most of which transpired prior to this time. Thus the change in adaptation rate is not inconsistent with our method.

2.4 Discussion

Interest in the DFE stems from a desire to know what beneficial mutations are available and which of these drive adaptation. In asexual populations, the DFE also determines the

distribution of competing mutations and the frequency of clonal interference. Here, we have described a simple experiment which exploits this connection in order to infer the DFE in experimental populations of *S. cerevisiae*. By introducing lineages with different fitnesses and tracking their subsequent dynamics, we inferred the DFE from the statistics of observed interference events. In the process, we directly observed how initial fitness advantages and clonal interference jointly influence the fixation or loss of adaptive lineages.

Previous experimental work has analyzed several other cases where an introduced lineage is outcompeted by a less-fit resident population.[43,131] Unlike our experiment, these earlier studies focus on the fates of a few key mutations (e.g., antibiotic resistance or microbial "cheaters") without attempting to infer the underlying DFE. Nevertheless, our results complement this earlier work by showing the transition between fitness effects that are susceptible to clonal interference and those that decisively sweep to fixation, which has previously been studied theoretically.[116,102,49] In our system, this transition occurs when the fitness of the seeded lineage is about 5 percent, which represents a critical effect size required for a mutation to drive adaptation. Of course, in natural populations some adaptive variants may arise in populations with substantial *standing* fitness variation, rather than the homogeneous resident populations employed here. In this case, the transition between mutations that sweep and those that experience interference is determined both by the DFE and by the distribution of fitnesses in the resident population. Further work is needed to address this situation.

Our computational inference method allowed us to distinguish between three representative DFE shapes: exponential, uniform, and δ-function (in which all mutations have the same effect). These represent idealized approximations to the actual DFE, and it is likely that a larger number of replicates or more sophisticated computational techniques could produce other DFE shapes with a significantly better fit. Yet one cannot continue this

process indefinitely without reaching a point where further determination of the fine-scale DFE becomes irrelevant for any particular application. In the end, certain features of the DFE matter for predicting certain aspects of the evolutionary process, and the required level of resolution is ultimately determined by the aspect of adaptation one wishes to study. Our present experiment, which focuses on the fates of advantageous mutants, provides a concrete illustration of this principle. Previous work has suggested that the dynamics of adaptation can be summarized by a single characteristic fitness effect, with a magnitude that depends on the actual DFE and the level of clonal interference within the population.[54,32,49] By rejecting the δ-function and uniform DFEs in favor of the exponential, we have shown that this assumption breaks down when one considers more detailed features of the lineage trajectories.

Given these caveats, the DFE that we inferred is worth pondering. We estimated an exponential distribution with mean $\bar{s} = 0.85\%$ and total beneficial mutation rate $U_b = 1.0 \cdot 10^{-4}$. Our modeling indicated that of these mutations, only those with effects greater than 2% affected the fates of seeded lineages, and that these mutations are predicted to arise at a rate of order 10^{-5} per individual per generation. If one assumes a per-genome point mutation rate of roughly $4{\cdot}10^{-3}$ [95], this would imply that of order 1 in 1000 mutations confer a fitness advantage of two percent or more. This is consistent with past work in a related system[33], and is also similar to DFEs reported for bacteria adapting to rich laboratory media.[107,69,134] In such permissive environments, other studies in yeast that have identified specific adaptive mutations report a mix of loss-of-function versus other kinds of beneficial mutations.[67,133,64,83,80] If a large fraction of beneficial mutations in our system are loss of function, and if roughly ten percent of spontaneous mutations in a gene cause loss of function,[82] our results would suggest that about 1 in 100 genes are beneficial to disrupt. This is at least qualitatively consistent with direct measurements using the yeast deletion

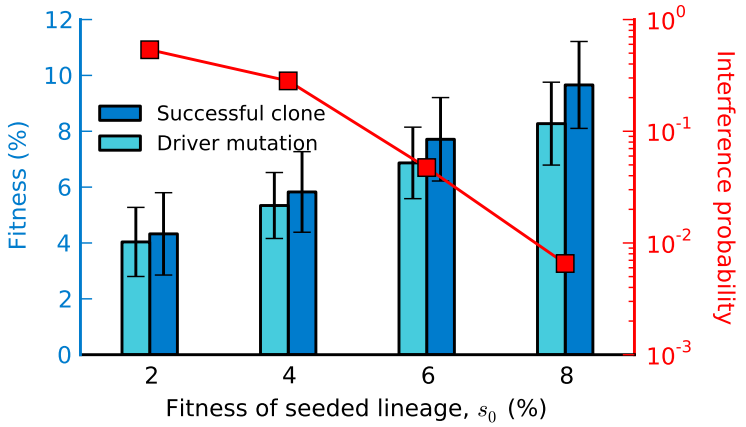

Figure 2.7: A screen for beneficial mutations. We simulated the evolutionary dynamics for a range of seeded lineages, and then simulated picking a single clone at random from the resident population immediately after a clonal interference event. The bars indicate the average fitness of this clone and its largest effect mutation (± 1 s.d., scale at left). We also show the fraction of replicate populations in our simulations in which clonal interference occurs (scale at right). The simulations assumed the most-likely exponential DFE inferred in the study.

collection.[119,10] Together, these results illustrate how inferences from lineage dynamics can combine with other lines of evidence to help build a more complete picture of adaptation.

Finally, we note that our experimental design has a potential practical application as a screen for beneficial mutations. Whenever a seeded lineage with fitness advantage s_0 experiences clonal interference, the resident must contain a mutant lineage at appreciable frequency with fitness greater than s_0. Thus, by picking clones from the resident immediately after a clonal interference event, we should in principle be able to isolate rare large-effect beneficial mutations. This is similar in spirit to earlier studies which used the dynamics of neutral markers to screen for adaptive clones (e.g. Rozen et al.[113]). However, because our seeded lineages are more fit than the resident, we can screen for beneficial mutations with particularly large effects. Further, since the resident must quickly generate a competing

lineage, our approach is more likely to find clones with fewer mutations of larger effect rather than many of smaller effect, as well as limit the number of non-beneficial hitchhiking mutations. To illustrate this idea, we simulated seeded lineage trajectories and then simulated picking a clone from the resident population after observed clonal interference events. In Fig. 2.7, we show the average fitness of each of these simulated clones and of the largest-effect mutation in each clone. As is apparent from the figure, it should be feasible to use this approach with a seeded lineage of the appropriate fitness to isolate large-effect beneficial mutations with specific fitness effects.

Acknowledgments

We thank Sergey Kryazhimskiy, Christopher S. Wylie, Andrew Murray, and Katya Kosheleva for useful discussions and comments on the manuscript; Melanie Muller, Gabriel Perron, John Koschwanez, and Gregg Wildenberg for help with strain construction; and Patricia Rogers for generous technical support of flow-cytometry. Simulations in this article were performed on the Odyssey cluster of the Research Computing Group at Harvard University. This work was supported by training grant GM831324 from the National Institutes of Health (NIH) and grant 1219334 from the NSF Physics of Living Systems graduate student network (E.M.F.), a National Science Foundation Graduate Research Fellowship (B.H.G.), and the James S. McDonnell Foundation, the Alfred P. Sloan Foundation, the Harvard Milton Fund, grant PHY 1313638 from the National Science Foundation, and grant GM104239 from the NIH (M.M.D.).

3

Crowded Growth Leads to the Spontaneous Evolution of Semi-Stable Coexistence

Identifying the mechanisms that create and maintain biodiversity is a central challenge in biology. Stable diversification of microbial populations often requires the evolution of differences in resource utilization. Alternatively, coexistence can be maintained by specialization to exploit spatial heterogeneity in the environment. Here, we report spontaneous diversification maintained by a related but distinct mechanism: crowding avoidance. During experimental evolution of laboratory *Saccharomyces cerevisiae* populations, we observed the repeated appearance of "adherent" lineages able to grow as a dispersed film, in contrast to their "crowded bottom-dweller" ancestors. These two types stably coexist because dispersal reduces interference competition for nutrients among kin, at the cost of a slower maximum growth rate. This tradeoff causes the frequencies of the two types to oscillate around an equilibrium over the course of repeated cycles of growth, crowding, and dispersal. However, further coevolution of the adherent and bottom-dweller types can perturb and eventually destroy their coexistence over longer timescales. We introduce a simple mathematical model of this "semi-stable" coexistence, which explains the interplay between

ecological and evolutionary dynamics. Since crowded growth generally limits nutrient access in biofilms, the mechanism we report here may be broadly important in maintaining diversity in these natural environments.

3.1 Introduction

The spontaneous evolution of stably coexisting lineages has been documented in several laboratory microbial evolution experiments. Such coexistence is typically maintained by some form of resource partitioning, either through specialization to different nutrients in the environment[39,12,22] or cross-feeding.[56,114,89,112,126,85] However, natural environments often harbor a greater diversity of microorganisms than nutrients, suggesting that many species stably coexist while competing for the same resources.[61] Understanding how this diversity is maintained despite the principle of competitive exclusion has long been an important challenge in biology.[52,50,123,117] An extensive body of theoretical work has sought to address this challenge by proposing a number of potential mechanisms that could maintain coexistence among lineages engaged only in exploitive competition (i.e., consumption of shared resources).[19,3]

Some of these mechanisms have been found to evolve spontaneously in experimental systems.[97,68] For example, coexistence can be maintained by cross-feeding of secondary metabolites in homogeneous environments containing a single limiting nutrient.[56] However, most experimental examples of the spontaneous evolution of coexistence on the same resources have involved spatially heterogeneous environments.[84,109,110] For example, laboratory Pseudomonas fluorescens populations diversify from a planktonic ancestor that occupies the liquid phase of statically incubated cultures into a second type that forms a mat at the broth surface.[110] This coexistence is stabilized by a tradeoff between the cost

of mat formation and its benefits in conferring privileged access to oxygen.[77] Coexistence based only on temporal heterogeneity is also possible[122] and has been investigated experimentally.[126] An evolved trade-off between nutritional competence and stress resistance has been found to maintain coexistence in the gut of gnotobiotic mice.[26]

Theory has suggested another mechanism of coexistence based on crowding avoidance: If access to resources is density-dependent and competing lineages positively assort, then competition within a lineage will be more intense than competition between lineages.[127,25] Here, we describe the first experimental observation of spontaneous diversification due to this crowding avoidance effect. As in earlier studies,[110] coexistence depends on spatial structure and a tradeoff between the ability to grow and maintain access to nutrients. However, nutrient access varies here according to cells' local density rather than the spatial heterogeneity of nutrients in the environment. This type of interference competition is widespread in microbial populations, particularly in biofilms.[73,136] We introduce a simple mathematical model of this mechanism, which explains our observations and shows that this coexistence is semistable: It can be changed or destroyed by evolution on longer time scales.

3.2 Results

3.2.1 Spontaneous Evolution of Stable Coexistence

In earlier laboratory evolution experiments,[38,81] we tracked the frequencies of fluorescently marked lineages over ∼1,000 generations in ∼1,000 haploid, asexual budding yeast populations propagated by serial passaging (with daily 1:2^{10} dilutions) in unshaken 96-well microplates. In most cases, natural selection eliminated diversity over time as these populations adapted, driving the marked lineage to either fixation or extinction. In 13 populations,

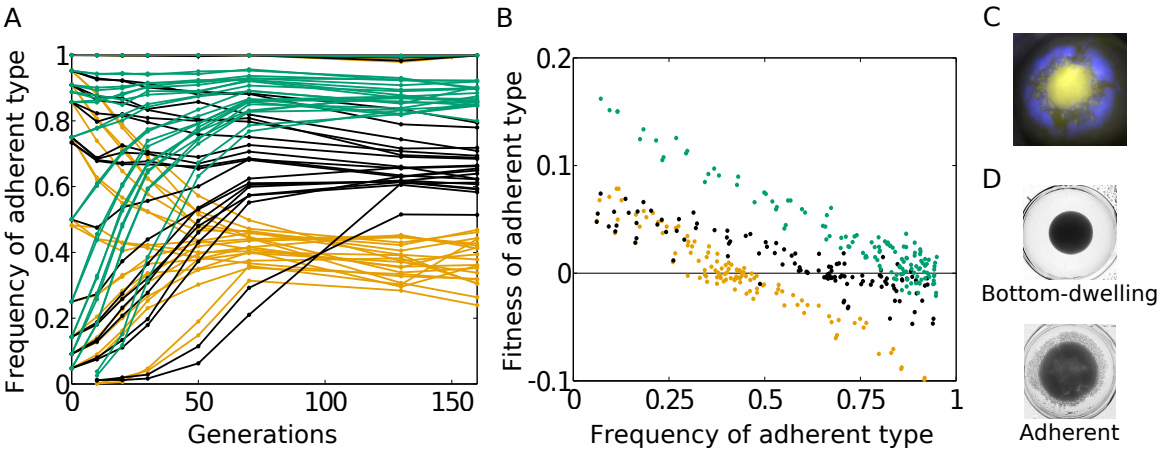

Figure 3.1: Coexistence due to negative frequency-dependent fitness. (A) Frequency of a fluorescently marked lineage over time. Colors represent independently evolved populations; each line is a replicate of the corresponding population in which the initial frequency of the marked lineage was perturbed to a given value. (B) Fitness of the marked strain as a function of its frequency, as calculated from the data in A. (C) Fluorescent image of one population shows that the marked lineage is located on the well walls, whereas the B type is located at the well bottom. (D) Microtiter wells containing isolated strains of the two types.

however, we observed marked lineages that remained at constant intermediate frequencies for hundreds of generations.

To test the stability of this coexistence, we used sorting cytometry to perturb the frequency of each marked lineage. Specifically, we varied the frequency of the marked lineage from 0 to 1 across 24 replicates generated from each of the original 13 populations. In many cases, the marked lineages returned to their original frequencies (Fig. 3.1A and Figs. B.1 and B.2), indicating that coexistence is stably maintained by negative frequency-dependent selection in these populations. However, lineages in different populations have different equilibrium frequencies and fitness-frequency relationships (Fig. 3.1B and Fig. B.1).

3.2.2 Identification of Adherent and Bottom-Dweller Types

Visual inspection of these lines revealed that all populations exhibiting stable coexistence also display a "dispersed" pellet morphology, with some cells able to grow across a greater range of the well surface (Fig. 3.1C). When we isolate clones from these populations, we find two types: adherent (A) types that can grow on the well walls and bottom-dweller (B) types that grow in a compact pellet at the well bottom (Fig. 3.1D). Yeasts are nonmotile, and these populations are grown in unshaken wells of media. Hence, adherence enables cells to colonize the walls only when they are diluted into a fresh well of media and adhere to the polystyrene well walls during initial sedimentation at the start of a new cycle of growth. In other words, unlike B cells, A cells tend not to tumble down to the bottom of the well once they land on a sloped surface, as seen in Movies B.1, B.2, B.3.

Competition assays revealed that the relative fitness of any pair of A and B clones is negative frequency-dependent, although different pairs have different fitness-frequency relationships (Fig. B.1). We found that the convergence of each marked lineage to its equilibrium frequency (Fig. 3.1A and Fig. B.1) is recapitulated by the pair of A and B clones isolated from the same population (Fig. B.1). We isolated clones from 59 other populations that had evolved dispersed morphology during the same experiment.[38] All of these populations contained A types, and we found that they could arise from any of the marked lineages (i.e., founding genotypes).

3.2.3 Role of Spatial Structure

The observation that cocultured A and B strains tend to occupy different regions of a microtiter well (Fig. 3.1C) suggests that their coexistence depends on spatial structure. In support of this view, we find that the frequency-dependent selection disappears and the A

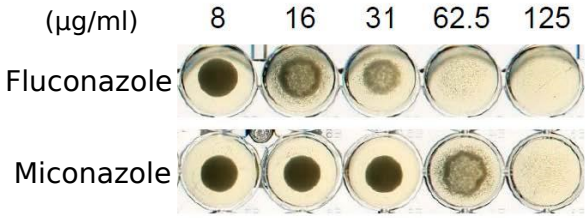

Figure 3.2: Adherence phenocopied using antifungal drugs. The B ancestral strain (DBY15108) was grown at a range of drug concentrations in round-bottom wells. After saturation, cells were resuspended by shaking, allowed to settle, and imaged (Materials and Methods).

and B types no longer coexist when microtiter plates are continuously shaken during growth or have a different well geometry (flat rather than round bottoms; Fig.B.1).

3.2.4 Genetic Basis of Adherence

Mating of A clones to their ancestors revealed that the A phenotype is a recessive Mendelian trait that segregates 2:2. To identify these mutations, we performed bulk segregant analysis of three independently evolved A clones,[118,13] and found a single mutation linked to the phenotype in each one (enzyme ERG3 and transcription factors UPC2 and HAP1, respectively). These genes are all related to the ergosterol biosynthesis pathway, and the nature of the mutations (Dataset B.1) led us to hypothesize that they reduce production of ergosterol. To test this hypothesis, we deleted ERG3 in the ancestral background, and the resulting strain is an A type that stably coexists with its B ancestor. We also found that adherence can be phenocopied by growing the ancestor at sublethal concentrations of azole antifungal drugs (Fig. 3.2), which inhibit another part of the ergosterol pathway (ERG11). Together, these data show that A lineages typically arise by single mutations that disrupt the ergosterol pathway and that such mutations alone are sufficient to give rise

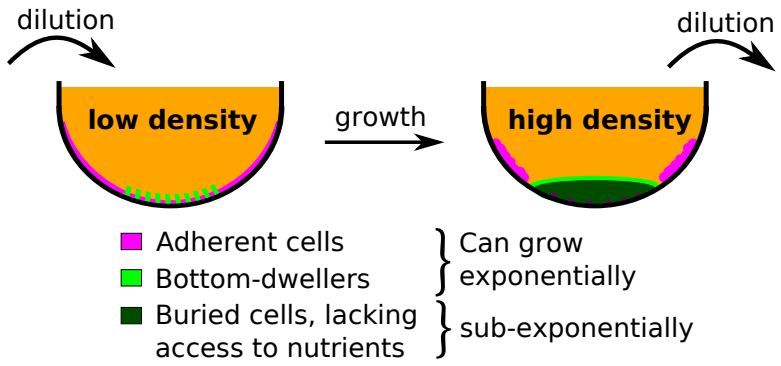

Figure 3.3: Effect of crowding on the growth of A and B strains. A schematic illustration of growth during one cycle is shown.

to coexistence.

3.2.5 Dynamics of A-Type vs. B-Type Growth

When cells are inoculated into a well at the beginning of each growth cycle, B cells accumulate at the bottom, whereas A cells also colonize the walls. We therefore hypothesized that adherence allows cells to escape crowding at the well bottom, dispersing them over a larger surface area, and therefore conferring a growth advantage during the high-density phase of each cycle of dilution and growth (illustrated schematically in Fig. 3.3). Coexistence then requires the A type to grow more slowly than the B type at low densities, as seen in Fig. 3.4. Accordingly, deleting ERG3 leads to a growth defect at low densities (Dataset B.1) and each A clone is less fit than the B clone isolated from the same population when grown in shaken or flat-bottom microplates (Fig. B.1).

The mechanism of coexistence proposed here implies a biphasic approach to the equilibrium frequency: The B type increases in frequency at low densities early in each cycle, but the A type later gains an advantage by better maintaining its growth rate as density increases. To test this hypothesis, we measured the relative frequencies of cocultured A and

B types over the course of two growth cycles (Fig. 3.4A and Fig. B.3). As expected, the frequencies of A and B types oscillate, with the B strain favored at the beginning of each cycle and the A type favored at the end. These dynamics are also qualitatively consistent with the growth curves of A and B strains cultured in isolation (Fig. 3.4B and Fig. B.4), which show that the B type grows faster at low density but slower at high densities.

3.2.6 Model of Coexistence

We can describe these dynamics using a simple mathematical model. At low density, both A and B types grow exponentially, at rates r_A and r_B, respectively. However, the growth rate of the B type declines at high densities due to crowding and burial. We model this decline as a transition to linear growth at rate $r_B n_B$, which reflects the continued growth of only a surface layer of n_B bottom-dwelling cells. In contrast, the A type continues to grow exponentially, because it is dispersed over a larger surface area. This mechanism does not require any differences in resource utilization, and accordingly, we find that the relative fitness of A and B types in shaken or flat-bottom plates is frequency-independent. For simplicity, we thus approximate the carrying capacity, K, of all strains to be the same. Together, this implies that during each daily growth cycle, the dynamics of the A and B types are given by

$$\begin{aligned} \frac{dA}{dt} &= Ar_A \cdot \Theta\left(1 - \frac{A+B}{K}\right) \\ \frac{dB}{dt} &= Br_B \cdot \left[1 + \frac{B}{n_B}\right]^{-1} \Theta\left(1 - \frac{A+B}{K}\right) \end{aligned} \tag{3.1}$$

Here, A and B represent the total number of A and B cells, and Θ is the Heaviside function (which states that all growth stops when the total population size reaches K).

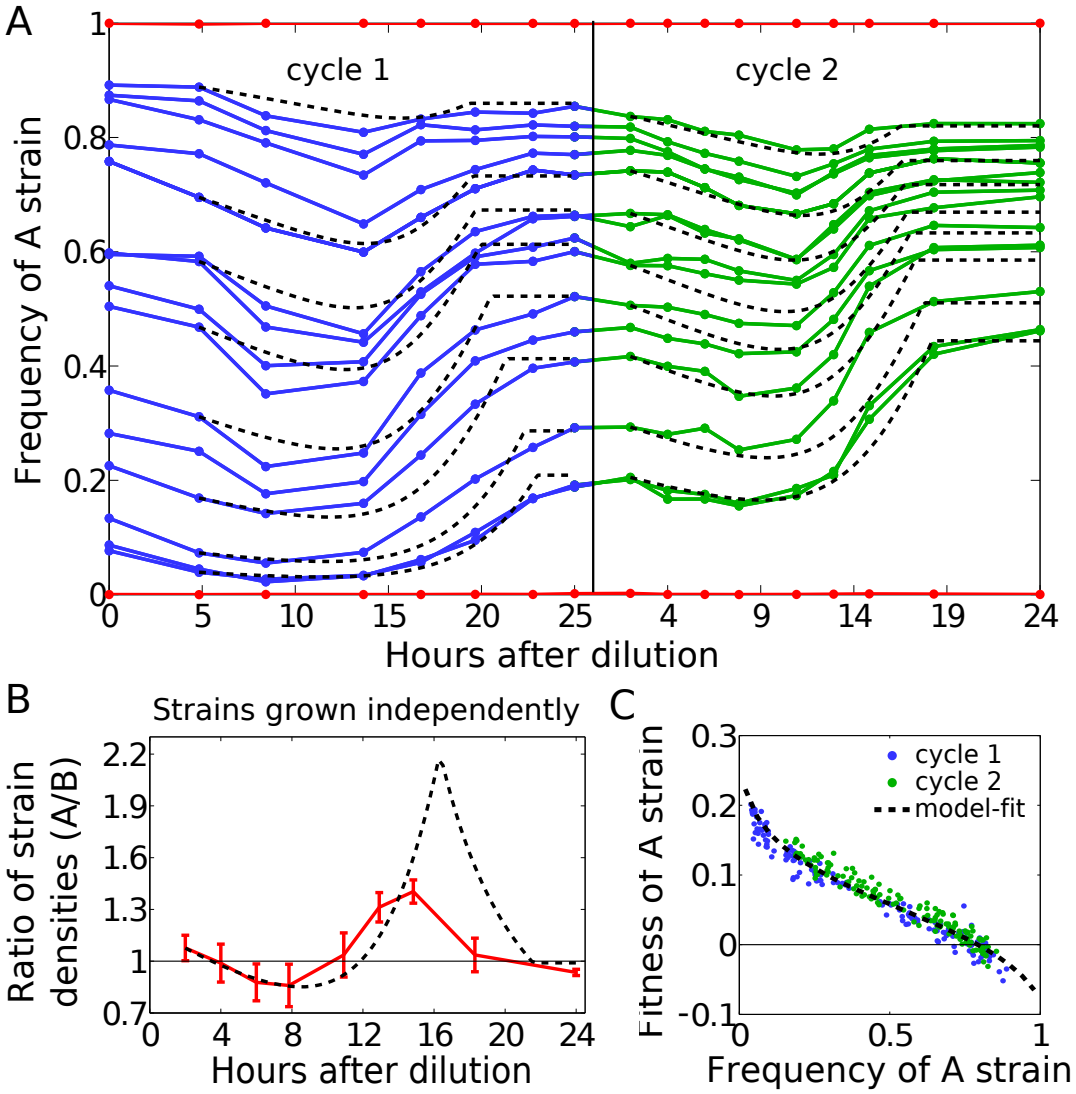

Figure 3.4: Population dynamics within growth cycles. (A) Frequency of the A type over the course of two growth cycles. Eighteen representative populations starting with different initial frequencies are shown (of 180 total populations; complete data are provided in Fig. B.3). (B) Ratio of the densities of the A type relative to the B type when each strain is grown in isolation (absolute density measurements are provided in Fig. B.4). (C) Net fitness of the A type across growth cycles in all populations as a function of starting frequency, calculated from the data in A and Fig. B.3. Dashed lines are the best-fit model prediction.

After each growth cycle, the cells are uniformly diluted into a fresh well (reducing both A and B by a factor of 2^{10}) and the cycle begins anew.

This model predicts that the A type declines in frequency during the initial phase of the growth cycle when the number of B cells is much less than n_B, because the B type has a faster maximum growth rate ($r_B > r_A$). If the A type is at higher frequency at the start of a growth cycle, this phase is prolonged (driving the A frequency down), because there are then fewer B cells initially, and hence they take longer to approach n_B. On the other hand, the A type increases in frequency later in the growth cycle, when the growth of the B type is limited by crowding. If the B type starts at high frequency, it more quickly approaches n_B and begins to feel these crowding effects, prolonging this second phase of the growth cycle (driving the A frequency up). Our mathematical model of these effects accurately reproduces the resulting biphasic convergence toward the equilibrium frequency (Fig. 3.4A) and the resulting fitness-frequency relationship (Fig. 3.4C). It also reproduces the qualitative differences in the growth curves of the two types when grown independently (Fig. 3.4B). To obtain these predictions, we set $K = 9 \cdot 10^6$ cells per well and $r_B = \ln(2)/80$ min^{-1} based on independent growth curve data (Fig. 3.4B and Fig. B.4), and fit the free parameters r_A and n_B, finding $r_A = \ln(2)/89$ min^{-1} and $n_B = 0.28K$. This inferred n_B corresponds to an $\sim 10\%$ slowdown of the B-type growth at $3 \cdot 10^5$ cells per well. Note that in Fig. 3.4B, the model overestimates the ratio of A- and B-type densities 12-20 h after inoculation; this transient discrepancy arises because we neglect the details of precisely how growth slows near saturation (Fig. B.4). These details can be incorporated into the model with additional parameters but are not necessary for accurately predicting the dynamics of the two types in co-culture.

3.2.7 Interplay Between Ecological and Evolutionary Dynamics

Although the two types occupy different spatial niches, they compete for shared nutrients. This competition implies that interplay between evolutionary and ecological factors will affect the long-term stability of coexistence. For example, both r_A and r_B can change over evolutionary time as the A and B types adapt. Our model predicts quantitatively how the equilibrium frequency of the two strains depends on the ratio of these growth rates, r_A/r_B (dashed curve in Fig. 3.5). Here, we use relative fitness in flat-bottom wells as a proxy for r_A/r_B (details are provided in Materials and Methods and Fig. B.5), because flat wells have vertical walls to which A cells hardly adhere, so that both types experience the same degree of crowding. Note this model prediction requires no additional fitting, because it depends only on the value of n_B/K estimated from Fig. 3.4. In principle, n_B and K could also evolve, and we could use our model to predict how this evolution would affect coexistence. However, we treat these parameters as constant because they are determined primarily by the fixed geometry of the microplate wells and the total nutrients in the media.

To test our predictions for how changes in r_A and r_B affect coexistence, we considered all possible pairwise combinations of 25 A strains and four B strains chosen to span a large range of fitnesses (Dataset B.1). Because we identified no A strains with fitness defects relative to B types greater than 10%, we also engineered a cycloheximide-resistant B strain and titrated the cycloheximide concentration in the media to decrease the relative growth rate of the A types artificially, and thus vary r_A/r_B over a broader range. For these competitions, we chose strains whose carrying capacities were unaffected by the cycloheximide for most of its concentration range (up to 100 nM). We found that our model accurately predicts how the equilibrium frequency depends on r_A/r_B (Fig. 3.5). Consistent with the model prediction, Fig. 3.5 shows that the B type will be driven extinct if the growth rate of the

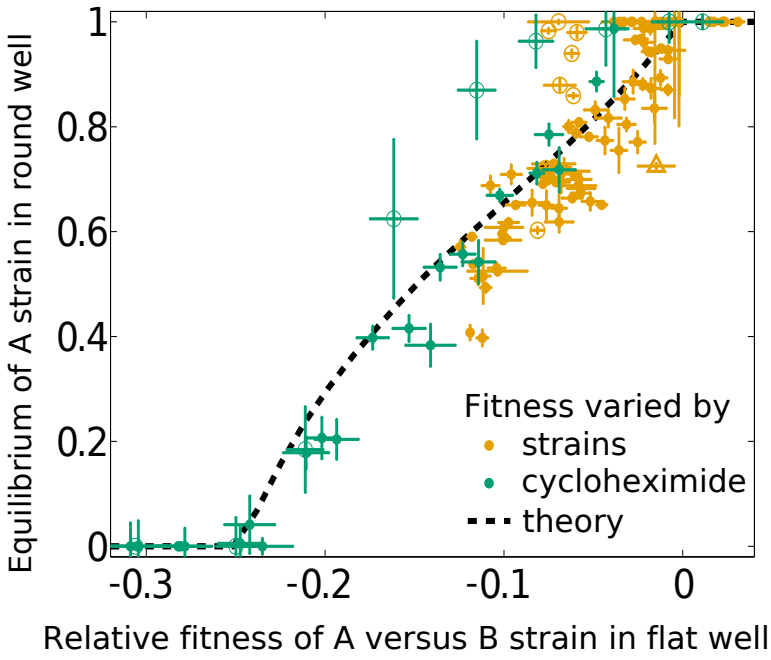

Figure 3.5: Semistability of the coexistence. We show the equilibrium frequency of A and B types as a function of their relative fitness in flat wells, which is a proxy for their ratio of growth rates, r_A/r_B. Each point represents the competition between one pair of A and B strains. Orange points are all pairwise combinations between 25 A strains and four B strains. The orange triangle is the constructed erg3Δ A strain vs. one of the B strains. Green points are competitions between four A strains and a cycloheximide-resistant B strain at a range of cycloheximide concentrations. Circled points correspond to two A strains that attained consistently higher than predicted equilibria, given their fitness. The dashed line is the model prediction; note this prediction involves no fitting.

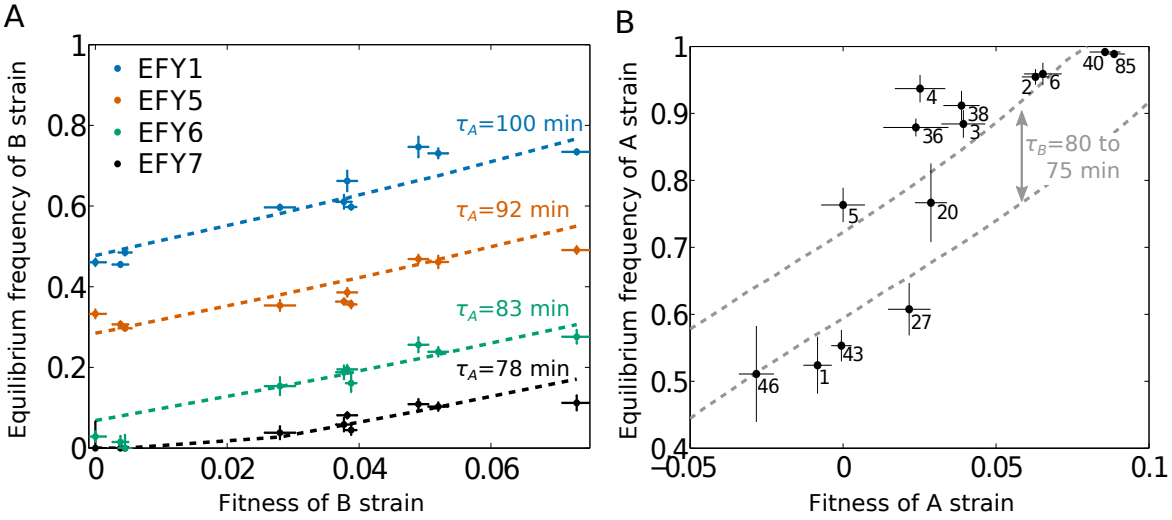

Figure 3.6: Dependence of coexistence on fitness. (A) All pairwise competitions between a set of A strains (color-coded) and B strains of varied fitness. These fitness measurements were included in an earlier publication.[38] (B) Similar data for A strains (strain numbers are indicated in the figure) vs. EFY64. Dashed curves show the model fit obtained by choosing the indicated values of $r_B = \ln(2)/\tau_B$. EFY5 is the A strain in Fig. 3.4; hence, its $r_A = \ln(2)/89$ min^{-1}. The r_A values of the other A strains are determined by their fitness relative to strain EFY5.

A type, r_A, evolves to match or exceed the growth rate of the B type, r_B. Conversely, the B type will drive the A type extinct if it achieves a growth rate $[1 - n_B \ln(d)/K]^{-1} \approx 25\%$ higher (here, d is the dilution factor).[122]

Furthermore, these data imply that adaptation within the subpopulations of coexisting A and B types tends to increase their equilibrium frequency vs. one another. To test this hypothesis, we confirmed that the fitness of A and B types, when competed against reference strains of the same type, positively correlates with their equilibrium frequency vs. strains of the opposite type (Fig. 3.6). We also conducted competition assays between combinations of two A strains and one B strain, and found that the A type with higher equilibrium frequency eliminates the one with lower equilibrium (Fig. 3.7 and Fig. B.6).

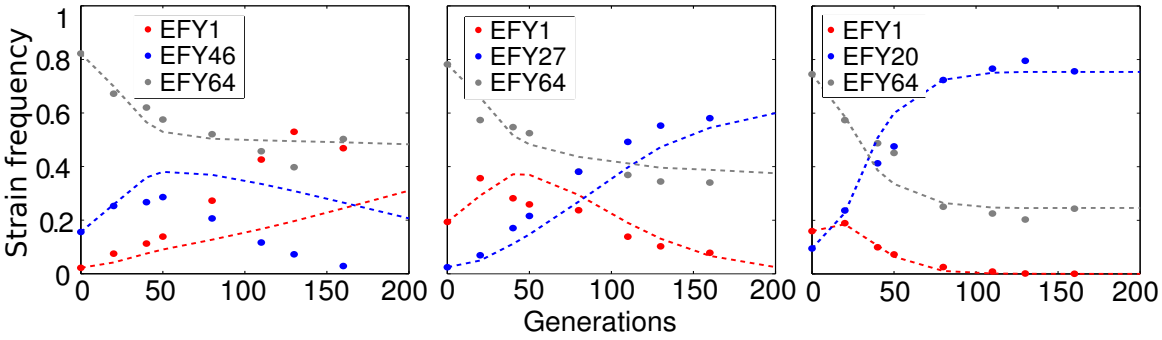

Figure 3.7: The outcomes of competition for three combinations of strains. Data points are color-coded by strain. EFY64 is type B, and the others are type A. Dashed curves are model predictions based on the measured frequencies at generation 0, the model parameters n_B and K inferred from Fig. 3.4, and growth rates inferred from the equilibrium frequencies in Fig. 3.6B (Materials and Methods). The outcomes of other three-strain competitions are shown in Fig. B.6.

3.2.8 Evolution of Filamentous Type that Disrupts Coexistence

During prior evolution experiments,[38,81] another distinct pellet morphology evolved in addition to the adherent A type, which we term filamentous (F) type (Fig. 3.8A). F-type pellets weakly stick together like a delicate fabric when aspirated from the bottom of a well or resuspended by shaking, and the strains appear to floculate in suspension. Several clones were isolated from two F-type populations (BYB1-B1 and BYB1-H6 in Lang et al.[81]) and mixed in roughly-equal proportion with populations containing coexisting A, B types (two populations from each round-well panel of Figs. B.1, B.2 at ∼130 generations). One of the F clones drove both A and B types extinct in all 26 populations within 40 generations. Another clone did so more slowly, as shown in (Fig. 3.8B). Preliminary data (not shown) indicate that F-type fitness versus the other types is positive frequency dependent. Fewer than half of spores inherited the phenotype in crosses with the ancestor, indicating multiple causal mutations, which is consistent with the F-type's relatively rare occurrence across populations: roughly 1 in 1000, compared to roughly 20% for the A phe-

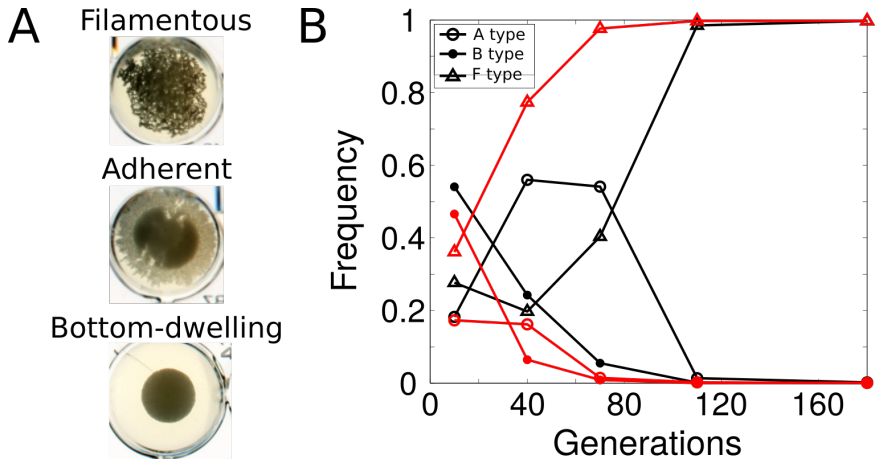

Figure 3.8: Identification of filamentous F type. (A) Pellet morphology of the F type compared to adherent and bottom-dwelling types. (B) Dynamics of all three types co-cultured in two independent populations (color-coded red, black).

notype. The F-type's ecological properties and their basis remain largely unknown, but its lattice-like pellet morphology and fitness suggest perhaps this phenotype is another means of crowding-avoidance.

3.3 Discussion

The spontaneous evolution of stable coexistence between competing lineages generally involves some type of tradeoff between traits affecting fitness.[117] A number of specific cases have been characterized in laboratory evolution experiments. For example, the growth of experimental *P. fluorescens* populations near the surface of a static broth creates an oxygen gradient, and coexistence arises between an evolved type able to colonize the oxygen-rich broth surface and an ancestral type able to grow faster but confined to oxygen-poor regions below.[110,77]

Coexistence between A and B types in our system arises due to a conceptually similar

tradeoff between the ability to grow vs. maintain access to nutrients. However, in our system, the A type achieves better access through crowding avoidance (i.e., reducing its density), rather than by occupying a privileged position near a nutrient source. Because growing nonmotile cells tend to become surrounded by their progeny, density-dependent interference competition for nutrients is typically more intense among kin. Hence, the A and B types meet the classic condition for competitive coexistence: Intraspecific competition within types is greater than the interspecific competition between them.[59] The two types exhibit different strategies for coping with this common form of intraspecific competition: The B type produces progeny as fast as possible but faces crowding sooner and more severely, whereas the A type mitigates crowding to grow at a slower but more sustainable rate.

Here, crowding avoidance improves access to a fixed pool of nutrients. This effect is closely related to other ecological mechanisms, such as dispersal, which, instead, typically involve gaining access to new resources. For example, the growth of microbial colonies in nutrient agar favors faster spreading away from regions of existing growth where nutrients are already depleted.[132,75,51] If cells are nonmotile, then spreading is achieved by growth itself, and there is essentially a single selection pressure for faster growth.[75] However, motile cells may face a trade-off between growth and motility, which can sustain coexistence.[93] Thus, crowding avoidance and dispersal can lead to similar selection pressures and ecological consequences.

One might expect the link between adherence and slower growth to represent a fundamental tradeoff if adherence required production of proteins (e.g., adhesins) or secretion of an extracellular matrix. However, we have no evidence that such is the case in our system, and overexpression of adhesins is typically associated with flocculation,[129] which we do not observe. Instead, the association between adherence and slower growth may be incidental to mutations disrupting the ergosterol pathway. Previous studies have shown that such

mutations reduce growth rate[41] and that deletions of ERG3 and ERG4 can increase and decrease adherence, respectively.[128] Most likely, A mutants are more hydrophobic, because adherence to polystyrene is itself a measure of cell-surface hydrophobicity.[74,53] Consistent with this speculation, pellet morphologies of A cells become indistinguishable from B cells in the presence of detergent (<0.05% Tween-20, a nontoxic concentration). However, the precise link between the ergosterol pathway and adherence is unknown (although this pathway is targeted by most antifungal drugs and adherence is a virulence factor).

Together, our findings show how an apparent tradeoff between faster growth and crowding avoidance evolves repeatedly and fosters ecologically stable coexistence that can be shifted and even eliminated by evolution on longer time scales. Here, this "semistability" is explained by a simple mathematical model that accurately predicts the full range of proportions at which the A and B types coexist as a function of their growth rate and related parameters. Semistable coexistence has also been studied in the context of cross-feeding.[85] Our findings suggest the following recurring evolutionary scenario. First, A types evolve from B types by loss-of-function mutations disrupting ergosterol biosynthesis. The resulting A types stably coexist with their B-type progenitors because these mutations incur a pleiotropic growth defect or occur on unfit genetic backgrounds. Subsequently, if both types adapt at the same rate, r_A/r_B will be constant and the equilibrium frequency will not change. However, due to the inherent randomness in evolution, and possibly due to differences in the available mutational spectrum (e.g., the potential for compensatory adaptation in the A lineages), r_A and r_B will typically evolve at different rates. Thus, evolution can cause the equilibrium frequency to change over time, and potentially even destroy coexistence. These effects could explain our observations in previous work of populations sometimes gaining and then losing a dispersed pellet morphology.[81] Furthermore, as these changes occur, additional A mutants can arise in the adapting B population. If the B type

has been adapting more rapidly than the A type, these new mutants are likely to be able to displace the original A type (a similar phenomenon of repeated recolonization of an established niche is observed in experimental *Burkholderia* biofilms[125]). These expectations are supported by competition assays between two A strains and one B strain, which show that A strains with higher equilibrium frequencies (relative to the B strain) displace those A strains with lower equilibria (Fig. 3.7 and Fig. B.6).

Our results and mathematical model highlight a simple mechanism for semistable diversification via crowding avoidance, which arises repeatedly even in a very simple laboratory system. Although the precise form of frequency-dependent selection we observe here is specific to the details of our system, this general mechanism of coexistence is likely to be broadly relevant. For example, microbial biofilms are usually characterized by periodic growth, competition, and dispersal. Crowding and adherence are key to these processes, and play a particularly important role in pathogenicity. Opportunities to reduce crowding, and hence relieve density-dependent competition within a lineage, may therefore play an important role in maintaining the extensive diversity of microbes in nature.

However, it is important to note that our system harbors more complexity than our model describes. For example, fitness differences among some B strains are affected by well geometry (Fig. B.2 and Dataset B.1), which cannot be explained in terms of growth rate and saturation density. A cells also often collapse in large numbers to the well bottom during later stages of growth, suggesting that they too experience some effects of crowding and may even actively bury the B cells below (Movies B.1, B.2, B.3). Possibly for this reason, some A strains attain consistently higher than predicted equilibria (circled points in Fig. 3.5). Given these caveats and complexities, these populations are a rich and tractable system for studying microbial ecology and eco-evolutionary dynamics.

3.4 Materials and Methods

3.4.1 Strains

All strains in this study are derived from a haploid MATa W303 ancestor. Specifically, strains EFY1-6, EFY50, and EFY67-85 are descendants of DBY15108 isolated from populations that evolved dispersed pellet morphology during an earlier long-term evolution experiment.[38] EFY20, EFY27, EFY37, EFY38, EFY40, EFY43, and EFY46 are descendants of EFY10-17 similarly isolated from the same experiment. EFY10-17 and EFY61-64 are descendants of DBY15104 and DBY15105, respectively, which were isolated from a different evolution experiment.[81] EFY10-17 were engineered to express a fluorescent reporter as previously described.[38] To construct the erg3Δ strain (MJM179), we amplified regions flanking the G418 resistance cassette of the yeast deletion collection erg3Δ mutant[42] and integrated the PCR product into the ERG3 locus of DBY15108. The deletion was confirmed by PCR and Sanger sequencing.

3.4.2 Bulk Segregant Analysis

To identify mutations causing the A phenotype, we sequenced three A clones and backcrossed each with an α-ancestor, resulting in three diploids heterozygous for all mutant sites present in the corresponding clone. We sporulated these diploids and dissected tetrads to isolate 40 recombinant haploids. Each haploid was inoculated into a round-bottom well and scored for the presence or absence of the A phenotype (Materials and Methods, Assay for Adherence). These haploids were split into two pools of clones containing either only A types or B types. Genomic DNA was obtained from each pool and genotyped by Sanger sequencing at each candidate locus. Causal loci were identified as those loci for which the

A pool exhibited only the mutant allele and the B pool exhibited only the ancestral allele.

3.4.3 Cell Culture

Yeast populations were propagated as previously described,[81] with slight modifications. Briefly, this protocol is to grow cells in 96-well polystyrene microplates containing 128μL of rich media (YPD) at 30°C and dilute them daily 2^{10}-fold into fresh plates using a pipetting robot (Beckman BioMek FX). To ensure consistent dispersal of the A type, plates were shaken 10-30 s (1,100 rpm; Heidolph Titramax 100) immediately after inoculation. Microplates had round or flat bottoms and were incubated either with or without continuous shaking (1,350 rpm; Titramax 100) as indicated above.

3.4.4 Measurement of Lineage Frequencies and Relative Fitness

Cell densities and frequencies of fluorescently labeled lineages were simultaneously determined by flow cytometry (BD Biosciences LSRII or Fortessa with attached high-throughput samplers). Because cytometry sample preparation requires washing cells in PBS, which disrupts population-level spatial structure, populations were assayed after their daily dilution or, in the case of Fig. 3.4 and Figs. B.3, B.4, B.5, by destructive sampling of replicates. To measure relative fitness, two strains of the same type are mixed together in equal proportion and propagated in co-culture for 30 generations.[81] The fitness is determined from their frequencies at generations 10 and 30 according to the formula: $s = \ln((f_x/f_y)t = T/(f_x/f_y)t = 0)/T$, where f_x and f_y are frequencies of lineages x and y and T is the number of population doublings (in this case, equal to 20). This formula is also used to calculate fitness from other lineage frequency data (e.g., Fig. 3.1A).

3.4.5 Direct Measurements of Maximal Growth Rates

To measure maximum growth rates (r_A and r_B in our model) directly, we maintained strains in the same way as described above, and assayed lineage frequencies and cell densities by destructively sampling replicates before saturation. For Fig. 3.4B and Fig. B.4, we measured total cell densities of one pair of A- and B-type strains at several time points. Because of the need to count many cells at very low densities in populations maintained in small volumes, it was not practical to measure direct growth rates accurately in this way for many strains. In addition, systematic errors can easily arise because the relevant growth rate differences are small, and these differences must be measured within the correct narrow time window during the first hours of growth before crowding effects become significant. Thus, to avoid these difficulties, we used fitness in flat-bottom wells as a proxy for r_A/r_B to test the model predictions shown in Fig. 3.5. To verify that these two quantities are correlated, we directly measured differences in maximum growth rates for 50 strain combinations (half of those strain combinations in Fig. 3.5). We did so by preparing replicates of strain pairs in coculture and assayed their relative frequencies 2 and 5 h after inoculation by destructive sampling of replicates. We then calculate $(1/\tau_x - 1/\tau_y) = \log_2((f_x/f_y)t = T/(f_x/f_y)t = 0)/T$, where $T = 180$ min and τ_x and τ_y are the doubling times of strains x and y. Although these measurements are noisy and may contain systematic biases for the reasons described above, we find that they are significantly correlated with relative fitness in flat-bottom wells (Pearson correlation = 0.55, $P < 10^{-5}$; Fig. B.5).

3.4.6 Measurements of Equilibrium Frequency

Equilibrium frequency is measured by mixing A- and B-type strains at a range of ratios and propagating the populations in parallel, as shown in Fig. 3.1 and Fig. B.1. Equilibrium

and its uncertainty then equal the mean and SE of the final frequencies. For Fig. 3.5, the propagation was truncated before populations converged to a common ratio. In these cases, the equilibrium is estimated from the fitness-vs.-frequency data (illustrated in Fig. 3.1B) using least-squares linear regression, and the uncertainty is computed by bootstrapping (SD of equilibria inferred from data sampled with replacement). In Fig. 3.5, the fitness and equilibrium between a cycloheximide-resistant B strain (DVD101) and two sensitive A strains (EFY3, EFY5, EFY68, and EFY85) were varied by competing these strains in media supplemented with cycloheximide concentrations ranging from 12.5 to 150 nM, in 12.5 nM increments.

3.4.7 Imaging

To obtain the fluorescent and time-lapse images of pellet morphologies (Fig. 3.1C and Movies B.1, B.2, B.3 respectively), we used a Zeiss LumarV12 stereoscope. For bright-field images (Fig. 3.1D), we used an Epson Perfection V700 transparency scanner.

3.4.8 Modeling

Our model predicts that equilibrium frequencies are determined by the ratios of parameters r_A/r_B and K/n_B. To obtain the predictions in Fig. 3.3, we set $r_B = \ln(2)/80$ min^{-1} and $K = 9 \cdot 10^6$ cells per well based on the growth curve data for single-strain populations (Fig. 3.4B and Fig. B.3). We then determined r_A and n_B by minimizing χ^2 error with respect to lineage frequencies and densities at the end of each cycle, assuming measurement uncertainties of 2.5% and 10% for lineage frequencies and cell densities, respectively. These computations were performed by a commercial differential equation solver and minimization algorithm [ode45 and fminsearch in MATLAB (MathWorks), version 2012a].

The value $K/n_B \approx 3.5$ inferred from Fig. 3.4 is assumed to be the same for all B strains. The model, given this assumption, implies a one-to-one correspondence between the fitnesses of strains of the same type and their ratio of growth rates, as well as between the equilibrium frequency of strains of the opposite type and their ratio of growth rates. Consequently, data in different figures can imply different values for the same parameters. For example, in Fig. 3.6A, the theoretical ratio of growth rates for strains EFY1 and EFY5 is 1.09 based on their different equilibria, but in Fig. 3.6B, it is 1.01 based on their difference in fitness. Likewise, in Fig. 3.7 and Fig. B.6, the predictions may be based on either the equilibria or fitnesses shown in Fig. 3.6B. We based predictions on the equilibria because, overall, this method produced better agreement with the data, but there are discrepancies when A strains have a larger difference in fitness than implied by their equilibria or vice versa. For example, EFY46 has a slightly lower equilibrium than EFY1, but a relatively large fitness defect (Fig. 3.6B). Hence, EFY46 declines faster than predicted by the model in Fig. 3.7 (Left).

The fitness values of the 14 A-type strains in Fig. 3.6B were inferred from 49 pairwise fitness measurements (Fig. B.7). We chose EFY5 as the zero-fitness reference and inferred the remaining 13 fitness values by minimizing χ^2 error under the assumption that fitness is transitive. The horizontal error bars show the parameter variation that increases χ^2 error by one unit around its minimum.

Acknowledgments

We thank Sergey Kryazhimskiy, Benjamin H. Good, Elizabeth Jerison, Chris Marx, and members of the M.M.D. laboratory for useful discussions; Andrew Murray for suggesting the cycloheximide experiment; and Patricia Rogers and Christian Daly for technical sup-

port of flow cytometry and liquid handling. This work was supported, in part, by Training Grant GM831324 from the NIH and Grant 1219334 from the National Science Foundation (NSF) Physics of Living Systems graduate student network (to E.M.F.); an NSF postdoctoral research fellowship (to J.D.V.D.); an NSF graduate research fellowship (to K.K.); the Charles E. Kaufman Foundation of The Pittsburgh Foundation (G.I.L.); and the James S. McDonnell Foundation, the Alfred P. Sloan Foundation, NSF Grant PHY 1313638, and NIH Grant GM104239 (to M.M.D.).

4

Conclusion

This thesis addresses some basic evolutionary questions—essentially, which lineages are likely to fix, go extinct or stably persist?—in a very specific system: populations of asexual budding yeast passaged in microtiter plates. Chapter 2 focuses on the question of fixation or extinction by directly measuring this process and using the data to infer the distribution of fitness effects of beneficial mutations, which generally governs lineage dynamics in an adapting population. Chapter 3 dissects a new example of coexistence that evolved spontaneously during the experiments of chapter 2, finding that it arose from the common fact that during growth, microorganisms compete for nutrients with nearby cells, which tend to be their own kin. The coexisting types, named bottom-dweller (B) and adherent (A), exhibit different strategies of dealing with this dilemma: cells may grow fast initially be-

fore becoming crowded by their progeny or employ some means (in this case adherence) to reduce crowding and grow at a slower but more sustainable rate.

Mathematical modeling and experiments reveal how the coexistence between these subpopulations and competition within them both largely depend on growth rate. As a result, the size of each subpopulation is coupled to the dynamics of adaption within them, and using the distribution of fitness effects inferred in chapter 2, one can predict how the equilibrium proportion of these coexisting types changes over time. Simulations of these dynamics are shown in Fig. 4.1A, with the corresponding non-frequency-dependent dynamics in Fig. 4.1B. Note that the time to fixation or extinction in the non-frequency dependent case is dominated by waiting for the first selective sweep, whereas with negative frequency-dependent selection, fitness gains within lineages incrementally shift their frequency, such that for every 1% increase in the average fitness of a lineage, its frequency may increase by ∼4% (see Fig. 3.5). As expected, negative frequency-dependent selection prolongs the coexistence of lineages (Fig. 4.1C). The coupling between adaptation and population size advantages the majority lineage because it is more likely to gain further beneficial mutations (a positive feedback loop). Hence negative frequency-dependent selection on average shortens coexistence at very low or high frequencies (Fig. 4.1C) and increases the fixation probability of the majority lineage (Fig. 4.1D). Given the dynamics in Figs. 4.1AB, the stability of experimentally-evolved lineages seen in Figs. B.1, B.2 appears highly improbable. However, these dozen populations were selected among ∼1000 specifically for the unusual stability of their lineages, so the accuracy of predictions such as those in Fig. 4.1 would need to be evaluated in future work. In any case, the dynamics predicted by Figs. 4.1AC are likely too fast because they assume a fixed distribution of fitness effects, whereas the rate of adaptation typically slows as fitness increases.[79,88,20,72]

These results complement a larger body of empirical[83,79,100,81,78,33] and theoretical

work[32,76,47,46,49,23] related to this experimental system, but how may they relate to evolution more generally? Some answers are provided in the discussion sections of the body chapters, but here this question is addressed more broadly. The typical way of thinking about evolution and ecology in Nature is to imagine that different types of organisms correspond to distinct niches or populations. Does this hold for the A and B types? In this case, relatively weak beneficial mutations may sweep only within whichever type they occur, but sufficiently strong mutations can drive the other type extinct, as illustrated by the data in Sec. 3.2.8 and Fig. B.6. So to some extent they are two populations, and to some extent they are one. The balance between these two extremes depends on the distribution of fitness effects of beneficial mutations and quantitative details of the mechanism of coexistence. The evolution of this type of scenario has also been identified in *E. coli* populations competing for glucose and acetate,[85,39,58,57] but its wider prevalence and evolutionary implications are open questions. The dynamics of adaptation in natural microbial populations are increasingly accessible through DNA sequencing, which also reveals an incredible diversity of microorganisms that largely remains to be explained.[11,92,106,31] If one were to generalize based on the results of this thesis, it would suggest that in these populations, the difference between stably-coexisting species or ecotypes and directly-competing lineages varies continuously rather than discretely.

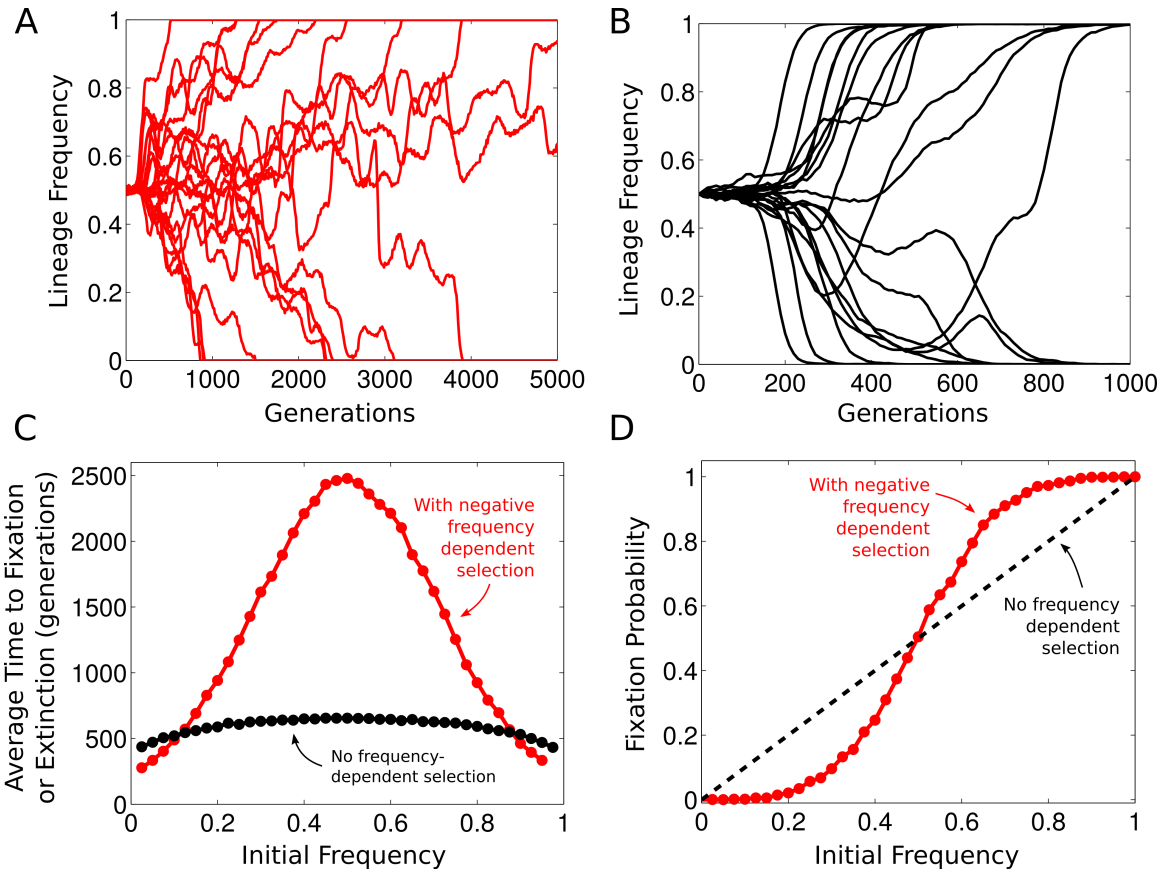

Figure 4.1: Lineage dynamics with vs. without negative frequency-dependent selection. (A) and (B) show simulations of 20 populations initially consisting of 2 clonal lineages at equal proportion. In (A), the lineages initially stably coexist at equal proportion according to the model in Ch. 3, in which case one type has a fitness advantage of ∼12.5% (see Fig. 3.5). Beneficial mutations arise within both lineages according to the distribution of fitness effects inferred in Ch. 2, shifting their equilibrium proportion over time. (B) shows the corresponding dynamics in absence of frequency-dependent selection, in which case the lineages have equal initial fitness. (C) and (D) show the effect of negative frequency-dependent selection on average duration of lineage coexistence and fixation probability, determined by 1000 simulations like those in (A) and (B) for each starting frequency.

A

Supplement to Chapter 2

Table A.1: Supplementary file containing all the data plotted in Fig. 2.1.

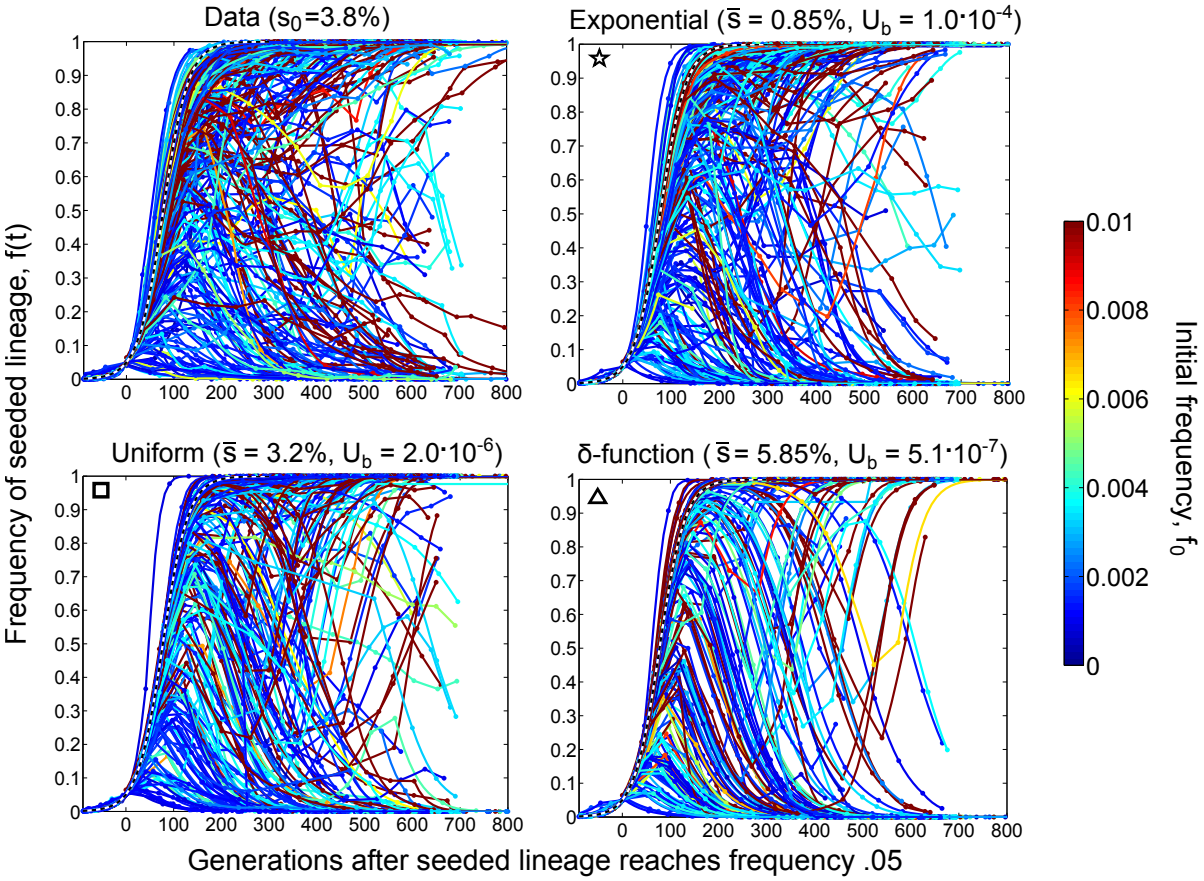

Figure A.1: Lineage dynamics data and simulations for $s_0 = 3.8\%$. Each panel shows the trajectories of seeded lineages with initial fitness $s_0 = 3.8\%$ as observed in the experiment (top left) and as reproduced by simulations assuming the DFE parameters indicated above each panel.

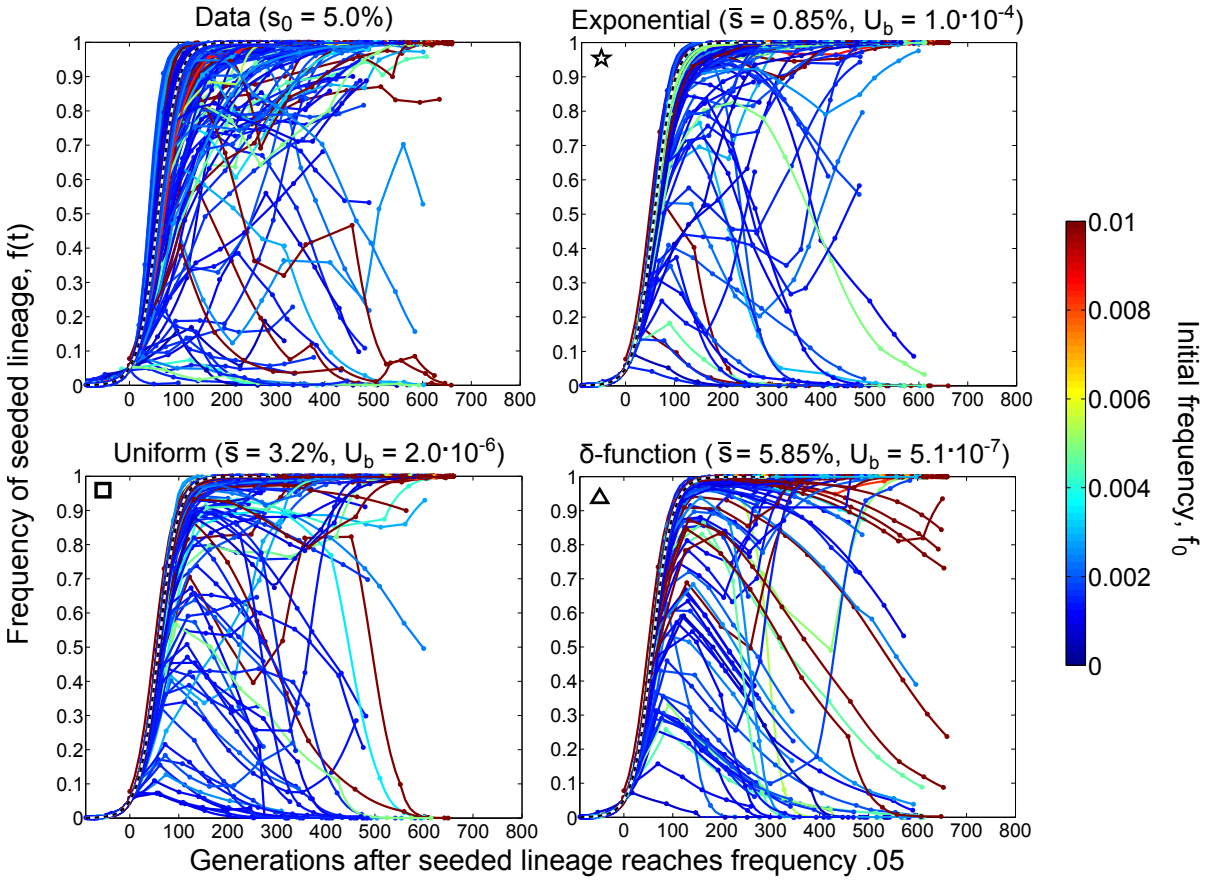

Figure A.2: Lineage dynamics data and simulations for $s_0 = 5.0\%$. Each panel shows the trajectories of seeded lineages with initial fitness $s_0 = 5.0\%$ as observed in the experiment (top left) and as reproduced by simulations assuming the DFE parameters indicated above each panel.

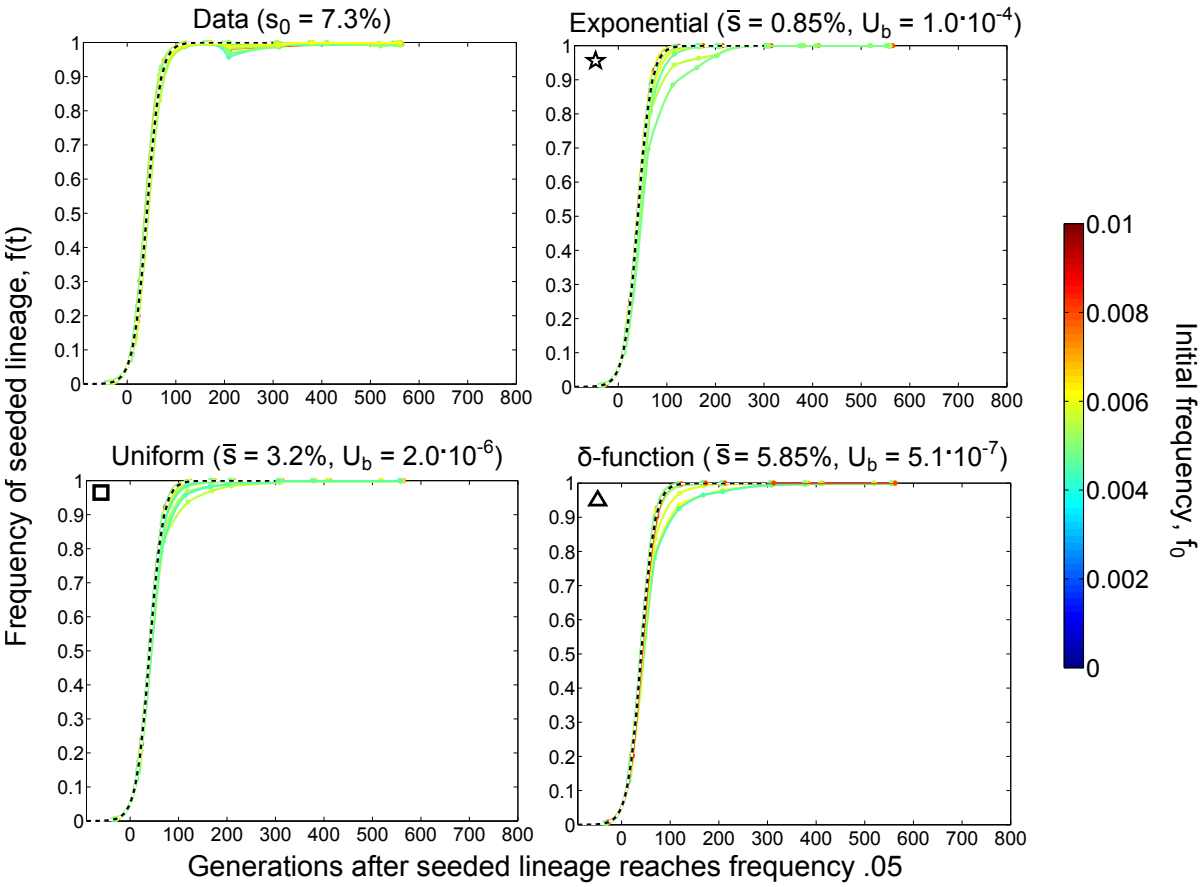

Figure A.3: Lineage dynamics data and simulations for $s_0 = 7.3\%$. Each panel shows the trajectories of seeded lineages with initial fitness $s_0 = 7.3\%$ as observed in the experiment (top left) and as reproduced by simulations assuming the DFE parameters indicated above each panel (which are also the ones indicated by the star, triangle and square in Fig. 2.4).

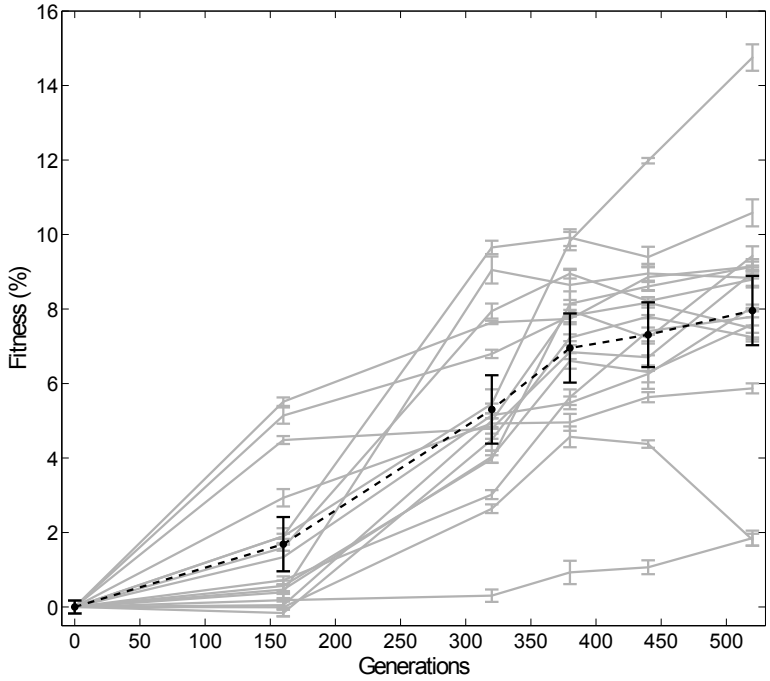

Figure A.4: The fitness over time of 16 experimental control populations (grey curves) and their mean (black curve). The error bars are ±1 s.e.m.

B

Supplement to Chapter 3

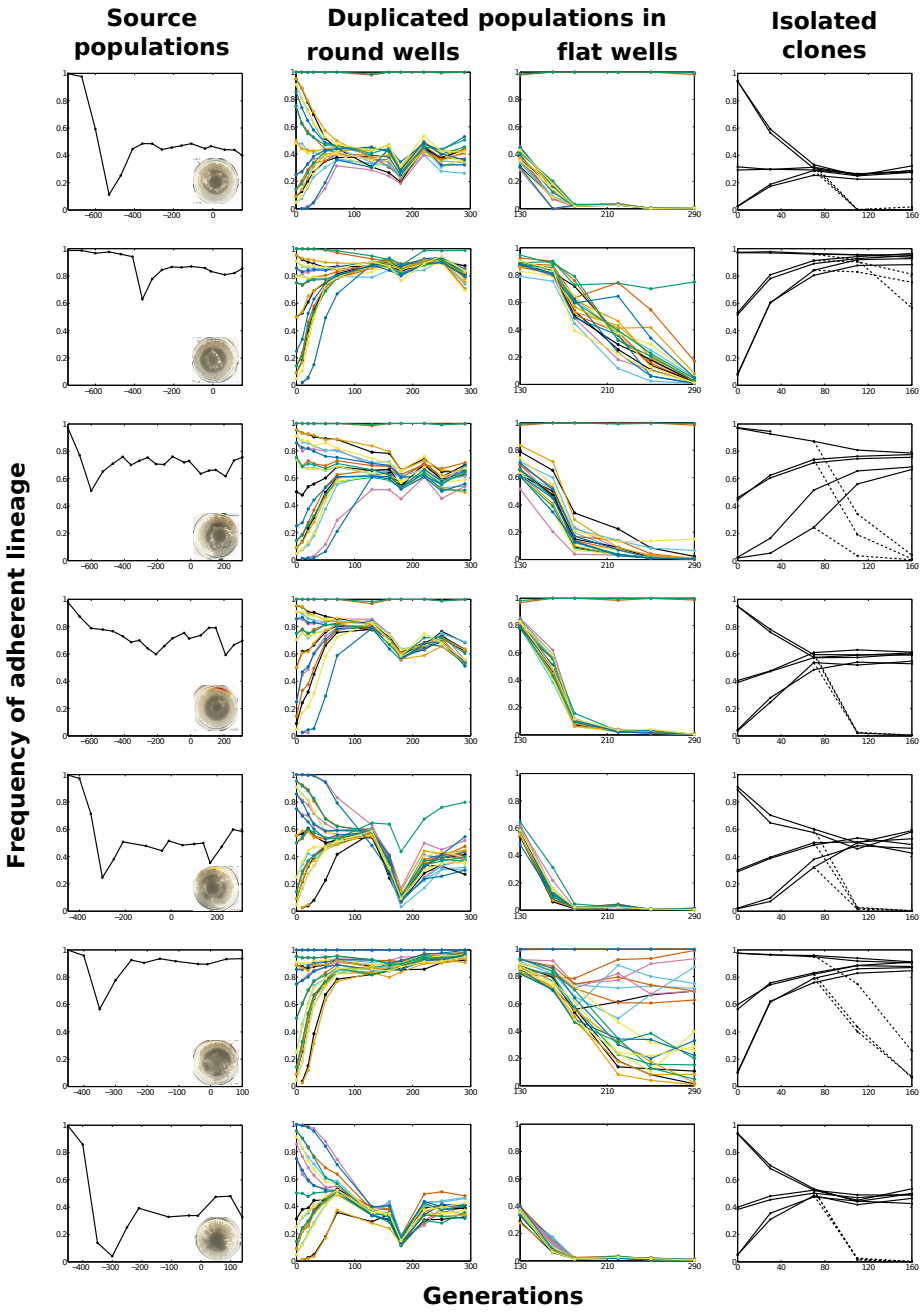

Figure B.1: Coexistence due to negative frequency-dependent fitness (full data). Column 1 shows the original observation of marked lineages remaining at intermediate frequencies, suggesting frequency-dependent selection. (Insets) All these populations had dispersed pellet morphology. Columns 2 shows these lineages returned to their original frequencies after perturbation by sorting cytometry (Materials and Methods; note that the sharp decline in frequencies around generation 180 is an artifact of an experimental error, in which populations were temporarily propagated at dilution factors $\gg 2^{10}$. Column 3 shows the resulting lineage dynamics when these populations were duplicated from round-bottom into flat-bottom wells (A cells do not adhere to the vertically sloped walls of flat wells, which eliminates the crowding avoidance effect by causing both types to be similarly confined to the well bottom, and hence prevents coexistence). Column 4 shows the competition of clones, one A and one B, drawn from the source population of each row and propagated in unshaken (solid lines) and shaken (dashed lines) round-bottom wells.

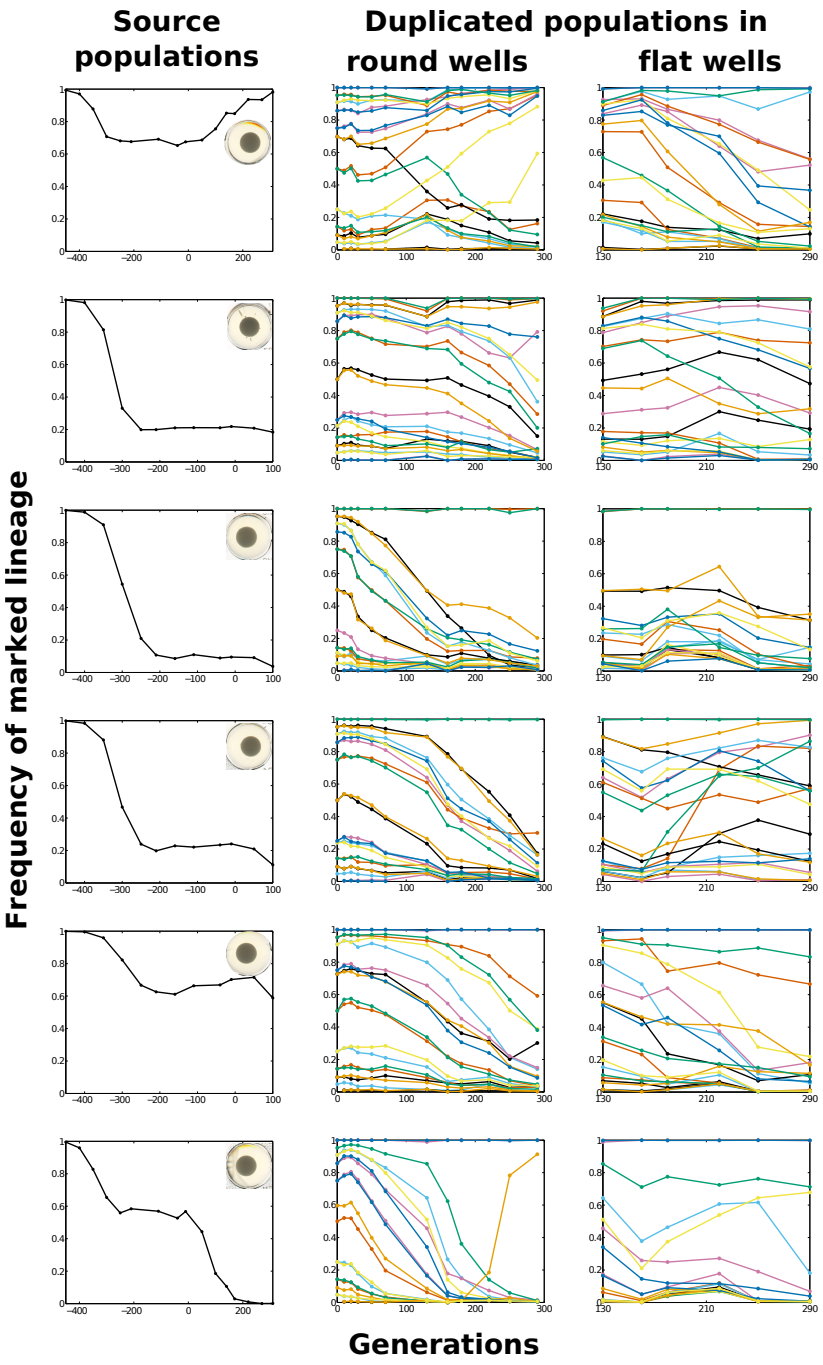

Figure B.2: B types do not exhibit frequency dependent fitness. Column 1 shows the observation of marked lineages remaining at intermediate frequencies in populations subsequently shown to lack frequency-dependent selection. (Insets) None of these populations had dispersed pellet morphologies. Column 2 shows the dynamics of these lineages after perturbation by sorting cytometry, and column 3 shows their dynamics after duplication from round-bottom to flat-bottom wells.

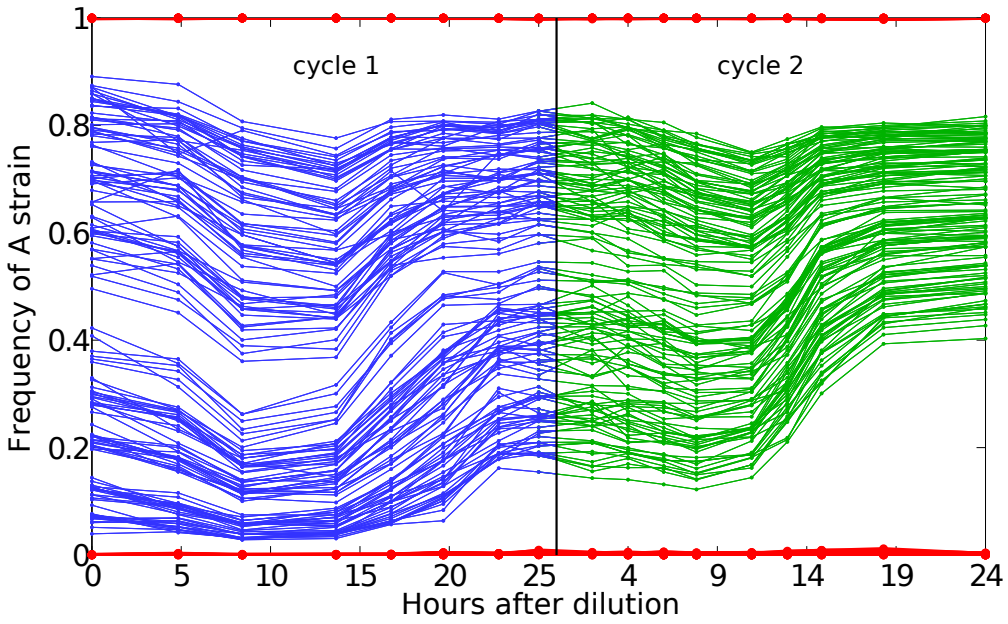

Figure B.3: A vs. B dynamics within each growth cycle. A strain frequencies in 162 populations are shown, in addition to the 18 shown in Fig. 3.4A.

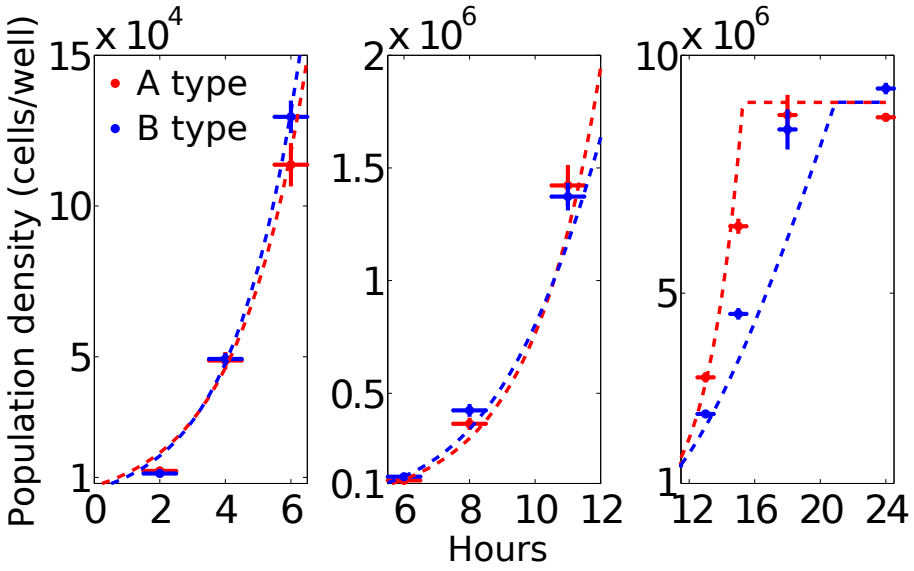

Figure B.4: Growth curves of A and B strains. The densities of the A and B strains in Fig. 3.4 (EFY5 and EFY10, respectively) as measured in 15 replicates (mean ±SEM, horizontal error bars of ±30 min). The dashed lines are simulated growth curves given eqns. 3.1 and parameters $K = 9 \cdot 10^6$ cells per well, $r_B = \ln(2)/80$ min^{-1}, $r_A = \ln(2)/89$ min^{-1} and $n_B = 0.28K$ (see Materials & Methods). The ratio of these measured and simulated growth curves is shown in Fig. 3.4B.

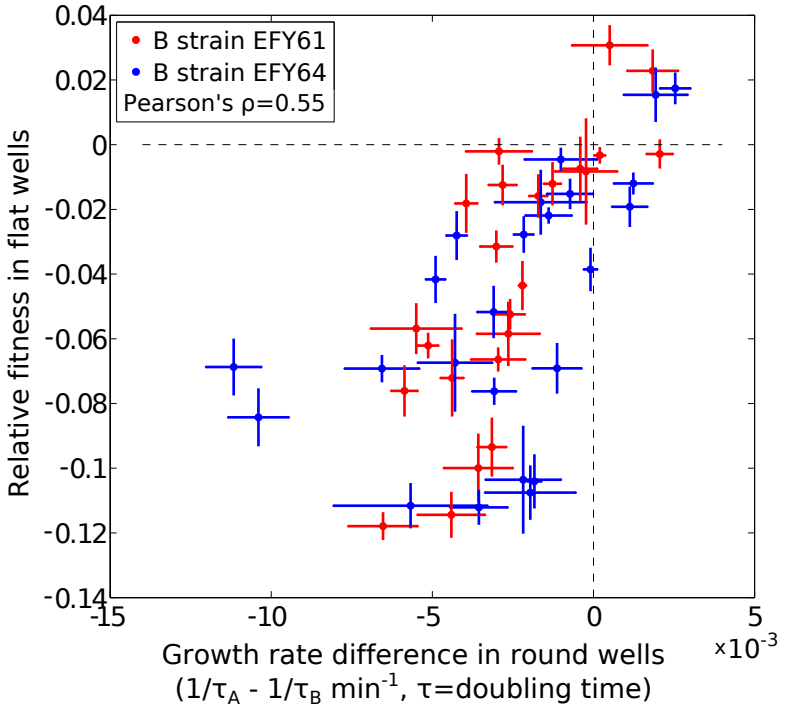

Figure B.5: Fitness in flat wells compared with the maximum growth rate in round wells. Absolute differences in growth rate are plotted on the x axis for 48 A and B strain pairs. These strains were measured at low density ($< 10^6$ cells per milliliter) in round wells by mixing pairs of strains and assaying their proportions at the beginning and end of a 3-h interval (error bars are SEM for four biological replicates). The measurements of relative fitness in flat-bottom wells, plotted on the y axis, are the same as in Fig. 3.5.

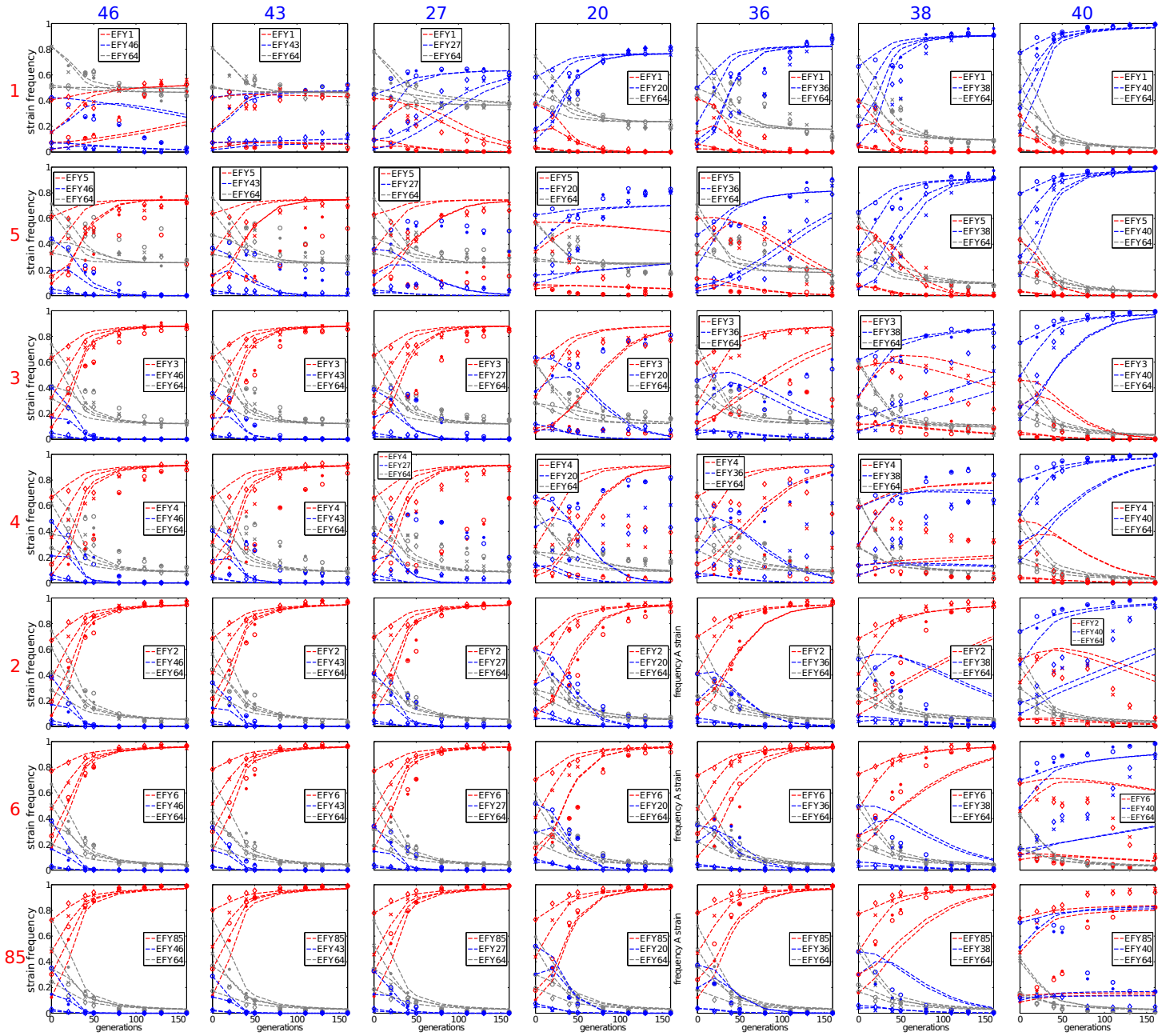

Figure B.6: Competition between two A strains and one B strain. Dynamics in populations containing all pairwise combinations between two sets of A strains, indicated by numbering of rows and columns, and one B strain (EFY64). Data points are color-coded according to strain, and symbols (×, ◇, ○, •) correspond to independent populations. Dashed curves are model predictions based on measured initial frequencies and parameters inferred from Figs. 3.4 and 3.6B (Materials and Methods). Measurements at generations 40 and 50 were performed on samples stored for ∼24 hours at 4°C, which appears to have biased those frequencies in favor of the nonfluorescent strain (EFY64).

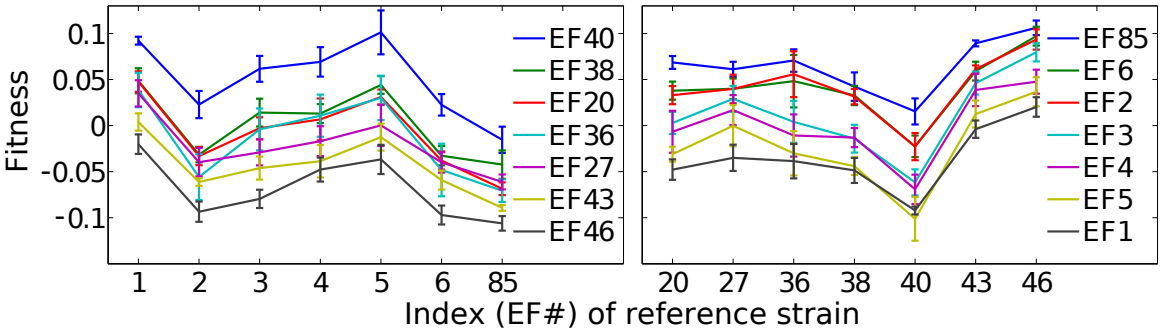

Figure B.7: Transitivity of fitness among A strains. All pairwise fitnesses between two sets of seven A strains measured in triplicate are shown. These data were used to determine the absolute fitnesses shown in Fig. 3.6B and were found to be consistent with transitive fitness among A strains ($\chi^2 = 0.6$; Materials and Methods).

Movie B.1: Time-lapsed growth in round-bottom wells. An A-type strain (EFY5) grown in isolation is shown. The time-lapse movie consists of images taken every 20 min for 22 h in standard growth conditions.

Movie B.2: Time-lapse movie of the growth of a B-type strain (EFY10) in a round-bottom well. Images taken every 20 min for 22 h in standard growth conditions.

Movie B.3: Time-lapse movie of the growth of A-type and B-type strains (EFY5 and EFY10) in coculture, mixed at equal initial proportions, in a round-bottom well. Images taken every 20 min for 22 h in standard growth conditions.

Data Set B.1: Summary of fitness data for all strains in Ch. 3; equilibrium frequencies for all pairs of strains tested; and the specific mutations in genes ERG3, UPC2, and HAP1 identified by bulk segregant analysis.

References

[1] Adams, J. & Oeller, P. (1986). Structure of evolving populations of Saccharomyces cerevisiae: adaptive changes are frequently associated with sequence alterations involving mobile elements belonging to the Ty family. *Proceedings of the National Academy of Sciences*, 83(18), 7124–7127.

[2] Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2014). *Molecular Biology of the Cell*, chapter Cancer. Garland Science.

[3] Armstrong, R. A. & McGehee, R. (1976). Coexistence of species competing for shared resources. *Theoretical Population Biology*, 9(3), 317–328.

[4] Atwood, K., Schneider, L., & Ryan, F. (1951). Periodic selection in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 37(3), 146.

[5] Barrett, R. D., MacLean, R. C., & Bell, G. (2006). Mutations of intermediate effect are responsible for adaptation in evolving Pseudomonas fluorescens populations. *Biology Letters*, 2(2), 236–238.

[6] Barrick, J., Kauth, M., Strelioff, C., & Lenski, R. (2010). Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects. *Molecular Biology and Evolution*, 27(6), 1338–1347.

[7] Barroso-Batista, J., Sousa, A., Lourenço, M., Bergman, M.-L., Sobral, D., Demengeot, J., Xavier, K. B., & Gordo, I. (2014). The first steps of adaptation of escherichia coli to the gut are dominated by soft sweeps. *PLoS Genet*, 10(3), e1004182.

[8] Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J. R., & Hwa, T. (2015). Overflow metabolism in escherichia coli results from efficient proteome allocation. *Nature*, 528(7580), 99–104.

[9] Bataillon, T., Zhang, T., & Kassen, R. (2011). Cost of adaptation and fitness effects of beneficial mutations in Pseudomonas fluorescens. *Genetics*, 189(3), 939–949.

[10] Bell, G. (2010). Experimental genomics of fitness in yeast. *Proceedings of the Royal Society B: Biological Sciences*, 277(1687), 1459–1467.

[11] Bendall, M. L., Stevens, S. L., Chan, L.-K., Malfatti, S., Schwientek, P., Tremblay, J., Schackwitz, W., Martin, J., Pati, A., Bushnell, B., et al. (2016). Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *The ISME journal*.

[12] Blount, Z. D., Borland, C. Z., & Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of escherichia coli. *Proceedings of the National Academy of Sciences*, 105(23), 7899–7906.

[13] Brauer, M. J., Christianson, C. M., Pai, D. A., & Dunham, M. J. (2006). Mapping novel traits by array-assisted bulk segregant analysis in saccharomyces cerevisiae. *Genetics*, 173(3), 1813–1816.

[14] Brown, L. D., Cai, T. T., & DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, (pp. 101–117).

[15] Buckling, A., Maclean, R. C., Brockhurst, M. A., & Colegrave, N. (2009). The beagle in a bottle. *Nature*, 457(7231), 824–829.

[16] Burch, C. L., Guyader, S., Samarov, D., & Shen, H. (2007). Experimental estimate of the abundance and effects of nearly neutral mutations in the RNA virus ϕ6. *Genetics*, 176(1), 467–476.

[17] Caspeta, L., Chen, Y., Ghiaci, P., Feizi, A., Buskov, S., Hallström, B. M., Petranovic, D., & Nielsen, J. (2014). Altered sterol composition renders yeast thermotolerant. *Science*, 346(6205), 75–78.

[18] Cermeño, P., Lee, J.-B., Wyman, K., Schofield, O., & Falkowski, P. G. (2011). Competitive dynamics in two species of marine phytoplankton under non-equilibrium conditions. *Mar Ecol Prog Ser*, 429, 19–28.

[19] Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, (pp. 343–366).

[20] Chou, H.-H., Chiu, H.-C., Delaney, N. F., Segrè, D., & Marx, C. J. (2011). Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*, 332(6034), 1190–1192.

[21] Conant, G. C. & Wolfe, K. H. (2007). Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Molecular systems biology*, 3(1).

[22] Craig MacLean, R., Dickson, A., & Bell, G. (2004). Resource competition and adaptive radiation in a microbial microcosm. *Ecology Letters*, 8(1), 38–46.

[23] Cvijović, I., Good, B. H., Jerison, E. R., & Desai, M. M. (2015). Fate of a mutation in a fluctuating environment. *Proceedings of the National Academy of Sciences*, 112(36), E5021–E5028.

[24] Dallinger, W. H. (1887). The president's address. *Journal of the Royal Microscopical Society*, 7(2), 185–199.

[25] De Leenheer, P., Angeli, D., & Sontag, E. D. (2006). Crowding effects promote coexistence in the chemostat. *Journal of Mathematical Analysis and Applications*, 319(1), 48–60.

[26] De Paepe, M., Gaboriau-Routhiau, V., Rainteau, D., Rakotobe, S., Taddei, F., & Cerf-Bensussan, N. (2011). Trade-off between bile resistance and nutritional competence drives escherichia coli diversification in the mouse gut. *PLoS Genet*, 7(6), e1002107.

[27] de Sousa, J. A. M., Campos, P. R., & Gordo, I. (2013). An ABC method for estimating the rate and distribution of effects of beneficial mutations. *Genome Biology and Evolution*, 5(5), 794–806.

[28] de Visser, J. & Rozen, D. (2006). Clonal interference and the periodic selection of new beneficial mutations in Escherichia coli. *Genetics*, 172(4), 2093–2100.

[29] de Visser, J. A., Zeyl, C. W., Gerrish, P. J., Blanchard, J. L., & Lenski, R. E. (1999). Diminishing returns from mutation supply rate in asexual populations. *Science*, 283(5400), 404–406.

[30] Dean, A. M. & Thornton, J. W. (2007). Mechanistic approaches to the study of evolution: the functional synthesis. *Nature Reviews Genetics*, 8(9), 675–688.

[31] Denef, V. J. & Banfield, J. F. (2012). In situ evolutionary rate measurements show ecological success of recently emerged bacterial hybrids. *Science*, 336(6080), 462–466.

[32] Desai, M. & Fisher, D. (2007). Beneficial mutation–selection balance and the effect of linkage on positive selection. *Genetics*, 176(3), 1759–1798.

[33] Desai, M., Fisher, D., & Murray, A. (2007). The speed of evolution and maintenance of variation in asexual populations. *Current Biology*, 17(5), 385–394.

[34] Dykhuizen, D. E. & Dean, A. M. (2009). Experimental evolution from the bottom up. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*, (pp. 67–89).

[35] Dykhuizen, D. E., Dean, A. M., & Hartl, D. L. (1987). Metabolic flux and fitness. *Genetics*, 115(1), 25–31.

[36] Elena, S. F. & Lenski, R. E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, 4(6), 457–469.

[37] Fraenkel, D. G. et al. (2011). *Yeast intermediary metabolism*. Cold Spring Harbor Laboratory Press.

[38] Frenkel, E. M., Good, B. H., & Desai, M. M. (2014). The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations. *Genetics*, 196(4), 1217–1226.

[39] Friesen, M. L., Saxer, G., Travisano, M., & Doebeli, M. (2004). Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in escherichia coli. *Evolution*, 58(2), 245–260.

[40] Gerrish, P. & Lenski, R. (1998). The fate of competing beneficial mutations in an asexual population. *Genetica*, 102, 127–144.

[41] Gerstein, A. C., Lo, D. S., & Otto, S. P. (2012). Parallel genetic changes and nonparallel gene–environment interactions characterize the evolution of drug resistance in yeast. *Genetics*, 192(1), 241–252.

[42] Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., & André, B. (2002). Functional profiling of the saccharomyces cerevisiae genome. *Nature*, 418(6896), 387–391.

[43] Gifford, D. R. & MacLean, R. C. (2013). Evolutionary reversals of antibiotic resistance in experimental populations of Pseudomonas aeruginosa. *Evolution*, 67(10), 2973–2981.

[44] Gillespie, J. (1983). A simple stochastic gene substitution model. *Theoretical Population Biology*, 23(2), 202–215.

[45] Giraud, A., Matic, I., Tenaillon, O., Clara, A., Radman, M., Fons, M., & Taddei, F. (2001). Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science*, 291(5513), 2606–2608.

[46] Good, B. H. & Desai, M. M. (2013). Fluctuations in fitness distributions and the effects of weak linked selection on sequence evolution. *Theoretical Population Biology*.

[47] Good, B. H. & Desai, M. M. (2014a). Deleterious passengers in adapting populations. *Genetics*, 198(3), 1183–1208.

[48] Good, B. H. & Desai, M. M. (2014b). The impact of macroscopic epistasis on long-term evolutionary dynamics. *Genetics*, (pp. genetics–114).

[49] Good, B. H., Rouzine, I. M., Balick, D. J., Hallatschek, O., & Desai, M. M. (2012). Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. *Proceedings of the National Academy of Sciences*.

[50] Grover, J. (1997). *Resource competition*, volume 19. Springer.

[51] Habets, M. G., Czaran, T., Hoekstra, R. F., & de Visser, J. A. G. (2007). Spatial structure inhibits the rate of invasion of beneficial mutations in asexual populations. *Proceedings of the Royal Society B: Biological Sciences*, 274(1622), 2139–2143.

[52] Hardin, G. et al. (1960). The competitive exclusion principle. *science*, 131(3409), 1292–1297.

[53] Hazen, K. & Hazen, B. (1987). A polystyrene microsphere assay for detecting surface hydrophobicity variations within candida albicans populations. *Journal of microbiological methods*, 6(5), 289–299.

[54] Hegreness, M., Shoresh, N., Hartl, D., & Kishony, R. (2006). An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science*, 311(5767), 1615–1617.

[55] Helling, R., Kinney, T., & Adams, J. (1981). The maintenance of plasmid-containing organisms in populations of Escherichia coli. *Journal of General Microbiology*, 123(1), 129–141.

[56] Helling, R. B., Vargas, C. N., & Adams, J. (1987). Evolution of escherichia coli during growth in a constant environment. *Genetics*, 116(3), 349–358.

[57] Herron, M. D. & Doebeli, M. (2011). Adaptive diversification of a plastic trait in a predictably fluctuating environment. *Journal of theoretical biology*, 285(1), 58–68.

[58] Herron, M. D. & Doebeli, M. (2013). Parallel evolutionary dynamics of adaptive diversification in escherichia coli. *PLoS Biol*, 11(2), e1001490.

[59] Holt, R. D. (2001). Species coexistence. *Encycl. Biodiv*, 5, 413–426.

[60] Huisman, J. & Weissing, F. J. (1999). Biodiversity of plankton by species oscillations and chaos. *Nature*, 402(6760), 407–410.

[61] Hutchinson, G. E. (1961). The paradox of the plankton. *The American Naturalist*, 95(882), 137–145.

[62] Illingworth, C. & Mustonen, V. (2012). A method to infer positive selection from marker dynamics in an asexual population. *Bioinformatics*, 28(6), 831–837.

[63] Imhof, M. & Schlötterer, C. (2001). Fitness effects of advantageous mutations in evolving Escherichia coli populations. *Proceedings of the National Academy of Sciences*, 98(3), 1113–1117.

[64] Jansen, M. L., Diderich, J. A., Mashego, M., Hassane, A., de Winde, J. H., Daran-Lapujade, P., & Pronk, J. T. (2005). Prolonged selection in aerobic, glucose-limited chemostat cultures of Saccharomyces cerevisiae causes a partial loss of glycolytic capacity. *Microbiology*, 151(5), 1657–1669.

[65] Jerison, E. R. & Desai, M. M. (2015). Genomic investigations of evolutionary dynamics and epistasis in microbial evolution experiments. *Current opinion in genetics & development*, 35, 33–39.

[66] Joseph, S. B. & Hall, D. W. (2004). Spontaneous mutations in diploid Saccharomyces cerevisiae more beneficial than expected. *Genetics*, 168(4), 1817–1825.

[67] Kao, K. & Sherlock, G. (2008). Molecular characterization of clonal interference during adaptive evolution in asexual populations of Saccharomyces cerevisiae. *Nature Genetics*, 40(12), 1499–1504.

[68] Kassen, R. (2009). Toward a general theory of adaptive radiation: insights from microbial experimental evolution. *Ann N Y Acad Sci*, 1168, 3–22.

[69] Kassen, R. & Bataillon, T. (2006). Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. *Nature Genetics*, 38(4), 484–488.

[70] Keightley, P. D. & Eyre-Walker, A. (2010). What can we learn about the distribution of fitness effects of new mutations from DNA sequence data? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1544), 1187–1193.

[71] Kellis, M., Birren, B. W., & Lander, E. S. (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast saccharomyces cerevisiae. *Nature*, 428(6983), 617–624.

[72] Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., & Cooper, T. F. (2011). Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, 332(6034), 1193–1196.

[73] Kim, W., Racimo, F., Schluter, J., Levy, S. B., & Foster, K. R. (2014). Importance of positioning for microbial evolution. *Proceedings of the National Academy of Sciences*, 111(16), E1639–47.

[74] Klotz, S., Drutz, D., & Zajic, J. (1985). Factors governing adherence of candida species to plastic surfaces. *Infection and immunity*, 50(1), 97–101.

[75] Korolev, K. S., Müller, M. J., Karahan, N., Murray, A. W., Hallatschek, O., & Nelson, D. R. (2012). Selective sweeps in growing microbial colonies. *Physical biology*, 9(2), 026008.

[76] Kosheleva, K. & Desai, M. M. (2013). The dynamics of genetic draft in rapidly adapting populations. *Genetics*, 195, 1007–1025.

[77] Koza, A., Moshynets, O., Otten, W., & Spiers, A. J. (2011). Environmental modification and niche construction: developing o2 gradients drive the evolution of the wrinkly spreader. *The ISME journal*, 5(4), 665–673.

[78] Kryazhimskiy, S., Rice, D. P., & Desai, M. M. (2012). Population subdivision and adaptation in asexual populations of saccharomyces cerevisiae. *Evolution*.

[79] Kryazhimskiy, S., Rice, D. P., Jerison, E. R., & Desai, M. M. (2014). Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science*, 344(6191), 1519–1522.

[80] Kvitek, D. J. & Sherlock, G. (2013). Whole genome, whole population sequencing reveals that loss of signaling networks is the major adaptive strategy in a constant environment. *PLoS Genetics*, 9(11), e1003972.

[81] Lang, G., Botstein, D., & Desai, M. (2011). Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics*, 188(3), 647–661.

[82] Lang, G. I. & Murray, A. W. (2008). Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae. *Genetics*, 178(1), 67–82.

[83] Lang, G. I., Rice, D. P., Hickman, M. J., Sodergren, E., Weinstock, G. M., Botstein, D., & Desai, M. M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*, 500(7464), 571–574.

[84] Larsen, D. H. & Dimmick, R. (1964). Attachment and growth of bacteria on surfaces of continuous-culture vessels. *Journal of bacteriology*, 88(5), 1380–1387.

[85] Le Gac, M., Plucain, J., Hindré, T., Lenski, R. E., & Schneider, D. (2012). Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with escherichia coli. *Proceedings of the National Academy of Sciences*, 109(24), 9487–9492.

[86] Lee, M.-C. & Marx, C. J. (2013). Synchronous waves of failed soft sweeps in the laboratory: Remarkably rampant clonal interference of alleles at a single locus. *Genetics*, 193(3), 943–952.

[87] Lenski, R. E., Rose, M. R., Simpson, S. C., & Tadler, S. C. (1991). Long-term experimental evolution in escherichia coli. i. adaptation and divergence during 2,000 generations. *American naturalist*, (pp. 1315–1341).

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

[89] Levin, B. R. (1972). Coexistence of two asexual strains on a single resource. *Science*, 175(4027), 1272–1274.

[90] Levin, S. A. (1970). Community equilibria and stability, and an extension of the competitive exclusion principle. *American Naturalist*, (pp. 413–423).

[91] Levy, S. F., Blundell, J. R., Venkataram, S., Petrov, D. A., Fisher, D. S., & Sherlock, G. (2015). Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature*.

[92] Ley, R. E., Peterson, D. A., & Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, 124(4), 837–848.

[93] Livingston, G., Matias, M., Calcagno, V., Barbera, C., Combe, M., Leibold, M. A., & Mouquet, N. (2012). Competition-colonization dynamics in experimental bacterial metacommunities. *Nat Commun*, 3, 1234.

[94] Luria, S. E. & Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28(6), 491.

[95] Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C. R., Dopman, E. B., Dickinson, W. J., Okamoto, K., Kulkarni, S., Hartl, D. L., et al. (2008). A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceedings of the National Academy of Sciences*, 105(27), 9272–9277.

[96] MacLean, R. & Buckling, A. (2009). The distribution of fitness effects of beneficial mutations in Pseudomonas aeruginosa. *PLoS Genetics*, 5(3), e1000406.

[97] MacLean, R. C. (2005). Adaptive radiation in microbial microcosms. *J Evol Biol*, 18(6), 1376–86.

[98] Maddamsetti, R., Lenski, R. E., & Barrick, J. E. (2015). Adaptation, clonal interference, and frequency-dependent interactions in a long-term evolution experiment with escherichia coli. *Genetics*, 200(2), 619–631.

[99] McDonald, M., Cooper, T., Beaumont, H., & Rainey, P. (2011). The distribution of fitness effects of new beneficial mutations in Pseudomonas fluorescens. *Biology letters*, 7(1), 98–100.

[100] McDonald, M. J., Rice, D. P., & Desai, M. M. (2016). Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature*, 531(7593), 233–236.

[101] Miralles, R., Gerrish, P. J., Moya, A., & Elena, S. F. (1999). Clonal interference and the evolution of RNA viruses. *Science*, 285(5434), 1745–1747.

[102] Neher, R. A. & Shraiman, B. I. (2011). Genetic draft and quasi-neutrality in large facultatively sexual populations. *Genetics*, 188(4), 975–996.

[103] Novick, A. & Szilard, L. (1950). Experiments with the chemostat on spontaneous mutations of bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 36(12), 708.

[104] Orr, H. (2003). The distribution of fitness effects among beneficial mutations. *Genetics*, 163(4), 1519–1526.

[105] Paquin, C. & Adams, J. (1983). Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. *Nature*, 302, 495–500.

[106] Pedrós-Alió, C. (2006). Marine microbial diversity: can it be determined? *Trends in microbiology*, 14(6), 257–263.

[107] Perfeito, L., Fernandes, L., Mota, C., & Gordo, I. (2007). Adaptive mutations in bacteria: high rate and small effects. *Science*, 317(5839), 813–815.

[108] Pinkel, D. (2007). Analytical description of mutational effects in competing asexual populations. *Genetics*, 177(4), 2135–2149.

[109] Poltak, S. R. & Cooper, V. S. (2010). Ecological succession in long-term experimentally evolved biofilms produces synergistic communities. *ISME*, 5(3), 369–378.

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

[111] Rokyta, D., Beisel, C., Joyce, P., Ferris, M., Burch, C., & Wichman, H. (2008). Beneficial fitness effects are not exponential for two viruses. *Journal of Molecular Evolution*, 67(4), 368–376.

[112] Rosenzweig, R. F., Sharp, R., Treves, D. S., & Adams, J. (1994). Microbial evolution in a simple unstructured environment: genetic differentiation in escherichia coli. *Genetics*, 137(4), 903–917.

[113] Rozen, D. E., De Visser, J., & Gerrish, P. J. (2002). Fitness effects of fixed beneficial mutations in microbial populations. *Current biology*, 12(12), 1040–1045.

[114] Rozen, D. E., Philippe, N., Arjan de Visser, J., Lenski, R. E., & Schneider, D. (2009). Death and cannibalism in a seasonal environment facilitate bacterial coexistence. *Ecology letters*, 12(1), 34–44.

[115] Sanjuán, R., Moya, A., & Elena, S. (2004). The distribution of fitness effects caused by single-nucleotide substitutions in an rna virus. *Proceedings of the National Academy of Sciences of the United States of America*, 101(22), 8396–8401.

[116] Schiffels, S., Szöllősi, G. J., Mustonen, V., & Lässig, M. (2011). Emergent neutrality in adaptive asexual evolution. *Genetics*, 189(4), 1361–1375.

[117] Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford University Press.

[118] Segrè, A. V., Murray, A. W., & Leu, J.-Y. (2006). High-resolution mutation mapping reveals parallel experimental evolution in yeast. *PLoS biology*, 4(8), e256.

[119] Sliwa, P. & Korona, R. (2005). Loss of dispensable genes is not adaptive in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 102(49), 17670–17674.

[120] Sniegowski, P. & Gerrish, P. (2010). Beneficial mutations and the dynamics of adaptation in asexual populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1544), 1255–1263.

[121] Sommer, U. (1989). The role of competition for resources in phytoplankton succession. In *Plankton ecology* (pp. 57–106). Springer.

[122] Stewart, F. M. & Levin, B. R. (1973). Partitioning of resources and the outcome of interspecific competition: a model and some general considerations. *American Naturalist*, (pp. 171–198).

[123] Tilman, D. (1982). *Resource competition and community structure*. Princeton University Press.

[124] Toprak, E., Veres, A., Michel, J.-B., Chait, R., Hartl, D. L., & Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nature genetics*, 44(1), 101–105.

[125] Traverse, C. C., Mayo-Smith, L. M., Poltak, S. R., & Cooper, V. S. (2013). Tangled bank of experimentally evolved burkholderia biofilms reflects selection during chronic infections. *Proceedings of the National Academy of Sciences*, 110(3), E250–E259.

[126] Turner, P. E., Souza, V., & Lenski, R. E. (1996). Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes. *Ecology*, 77(7), 2119–2129.

[127] Vance, R. R. (1985). The stable coexistence of two competitors for one resource. *American Naturalist*, (pp. 72–86).

[128] Vandenbosch, D., De Canck, E., Dhondt, I., Rigole, P., Nelis, H. J., & Coenye, T. (2013). Genomewide screening for genes involved in biofilm formation and miconazole susceptibility in saccharomyces cerevisiae. *FEMS yeast research*, 13(8), 720–730.

[129] Verstrepen, K. J., Reynolds, T. B., & Fink, G. R. (2004). Origins of variation in the fungal cell surface. *Nat Rev Microbiol*, 2(7), 533–40.

[130] Wahl, L. M., Gerrish, P. J., & Saika-Voivod, I. (2002). Evaluating the impact of population bottlenecks in experimental evolution. *Genetics*, 162(2), 961–971.

[131] Waite, A. J. & Shou, W. (2012). Adaptation to a new environment allows cooperators to purge cheaters stochastically. *Proceedings of the National Academy of Sciences*, 109(47), 19079–19086.

[132] Wei, Y., Wang, X., Liu, J., Nememan, I., Singh, A. H., Weiss, H., & Levin, B. R. (2011). The population dynamics of bacteria in physically structured habitats and the adaptive virtue of random motility. *Proceedings of the National Academy of Sciences*, 108(10), 4047–4052.

[133] Wenger, J. W., Piotrowski, J., Nagarajan, S., Chiotti, K., Sherlock, G., & Rosenzweig, F. (2011). Hunger artists: yeast adapted to carbon limitation show trade-offs under carbon sufficiency. *PLoS Genetics*, 7(8), e1002202.

[134] Wiser, M. J., Ribeck, N., & Lenski, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science*, 342(6164), 1364–1367.

[135] Wolfe, K. H., Shields, D. C., et al. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature*, 387(6634), 708–712.

[136] Xavier, J. B. & Foster, K. R. (2007). Cooperation and conflict in microbial biofilms. *Proceedings of the National Academy of Sciences*, 104(3), 876–81.

[137] Yona, A. H., Manor, Y. S., Herbst, R. H., Romano, G. H., Mitchell, A., Kupiec, M., Pilpel, Y., & Dahan, O. (2012). Chromosomal duplication is a transient evolutionary solution to stress. *Proceedings of the National Academy of Sciences*, 109(51), 21010–21015.

[138] Zhang, W., Sehgal, V., Dinh, D. M., Azevedo, R. B., Cooper, T. F., & Azencott, R. (2012). Estimation of the rate and effect of new beneficial mutations in asexual populations. *Theoretical population biology*, 81(2), 168–178.