Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Statistical questions in experimental evolution

Michael M Desai

Departments of Organismic and Evolutionary Biology and of Physics, and FAS Center for Systems Biology, Harvard University, 435.20 Northwest Labs, 52 Oxford Street, Cambridge, MA 02138, USA
E-mail: mdesai@oeb.harvard.edu

Received 6 April 2012
Accepted 4 June 2012
Published 16 January 2013

Online at stacks.iop.org/JSTAT/2013/P01003
doi:10.1088/1742-5468/2013/01/P01003

Abstract. Recent advances in the mathematical analysis of models describing evolutionary dynamics are rapidly increasing our ability to make precise quantitative predictions. These advances have created a growing need for corresponding improvements in our ability to observe evolutionary dynamics in laboratory evolution experiments. High-throughput experimental methods are particularly crucial, in order to maintain many replicate populations and measure statistical differences in evolutionary outcomes at both phenotypic and genomic levels. In this paper, I describe recent technical developments which have greatly increased the throughput of laboratory evolution experiments, and outline a few promising directions for further improvements. I then highlight a few ways in which these new experimental methods can help to answer simple statistical questions about evolutionary dynamics, and potentially guide future theoretical work.

Keywords: models for evolution (experiment)

J. Stat. Mech. (2013) P01003

1742-5468/13/P01003+20$33.00

Contents

1. Introduction

Experimental evolution is not a new idea. Plant and animal breeders have practiced it, sometimes inadvertently, for thousands of years, and the first explicit attempts to evolve laboratory populations to test evolutionary theory date to contemporaries of Darwin [1]. Since then, many others have carried out evolution experiments in organisms such as flies [2, 3], beetles [4], microbes [5, 6], and viruses [7] (see [8] for a collection of recent reviews). The past two decades have seen an explosive growth in this field, driven primarily by Lenski's long-term evolution experiments in *E. coli*, along with the subsequent work of many of his students and collaborators [5, 9].

Because of the physical demands of the experiments, almost all of these laboratory evolution studies have been limited to a small number of replicate populations (e.g. 12 independent *E. coli* lines in Lenski's long-term experiments). This limit to throughput has colored the types of question the field has addressed. Many experiments are intended to be primarily observational and exploratory: the experimenter challenges a population with some specific selective challenge and watches what happens. In other cases the primary aim is to understand the biological mechanisms underlying adaptation to a specific challenge, such as the response of yeast to nutrient starvation [10], or the evolution of a genetic switch [11]. In still other cases, the main aim is to watch evolutionary dynamics. Experiments in this latter category sometimes attempt to test simple qualitative predictions of evolutionary or ecological theory, such as whether sex increases the rate of adaptation [12, 13], whether increases in mutation rates can be favored in adapting populations [14], or whether certain types of population structure promote the evolution of antibiotic resistance [15].

Despite their often limited statistical power, these existing approaches to experimental evolution continue to be productive. Yet evolution is inherently a random process, so any theoretical prediction or experimental result is inevitably a statistical statement. As recent advances in the analysis of mathematical models of evolutionary dynamics

J. Stat. Mech. (2013) P01003

make ever more precise and quantitative statistical predictions, parallel improvements in experimental evolution have become increasingly important. This calls for high-throughput experimental methods to maintain and observe many independently evolving replicate populations, in order to measure the relative probabilities of different outcomes. Recent technical advances in experimental evolution have made it possible to do exactly this, by dramatically increasing the throughput at which we are able to maintain evolving lines and assay evolutionary outcomes. This has opened new directions for productive interactions between experiments and theory.

In this paper, I describe some of these new developments, and highlight a few important open questions and challenges. I begin in section 2 by describing technical advances which have made it possible to evolve thousands of laboratory populations in parallel. I then describe a few simple statistical questions in evolutionary dynamics that these new high-throughput methods can help us to address.

2. High-throughput methods for experimental evolution in microbes

Recent efforts to increase the throughput of laboratory evolution experiments have focused primarily on two areas: improvements to strain maintenance techniques, and improved methods to assay changes in evolved lines. Here, I focus on methods relevant for microbial populations such as bacteria and yeast. Current work is also making rapid progress along similar lines in improving the throughput of viral evolution experiments (see e.g. [16]), but I do not discuss this work specifically here.

2.1. High-throughput strain maintenance

Experimental evolution of microbial populations is largely an exercise in liquid handling. Other approaches do exist, such as evolution on agar plates [17], but most experiments are carried out in liquid media. Much work has made use of chemostats (figure 1(A)), in which new medium is continuously dripped into a culture vessel while older medium is drained out [18]. These chemostats require regular observation and maintenance. A single researcher can typically maintain at most a few dozen at a time, and avoiding contamination for more than a few months is nearly impossible.

A somewhat higher-throughput alternative to chemostats is batch culture experiments (figure 1(B)). In batch culture, a population is diluted into a fixed volume of new sterile medium on a regular basis, typically at least daily. This is very straightforward, but in practice it is hard to maintain many populations simultaneously by hand over long periods of time. Repeating the same procedure every day (or in some cases several times a day) over periods of months or years makes it easy to lose concentration and make mistakes. Experimental evolution is particularly sensitive to cross-contamination between independent replicates—even a single contaminating cell can be disastrous, and these events can easily go undetected for months. Personally, I have found it impossible to maintain by hand more than about fifty budding yeast lines simultaneously for extended periods, or more than about five hundred lines for a few days.

In the past few years, experiments have begun to achieve much higher throughput in batch culture by using liquid handling robots to transfer cultures. This is an obvious improvement (these robots have been widely available for decades), but special care must

J. Stat. Mech. (2013) P01003

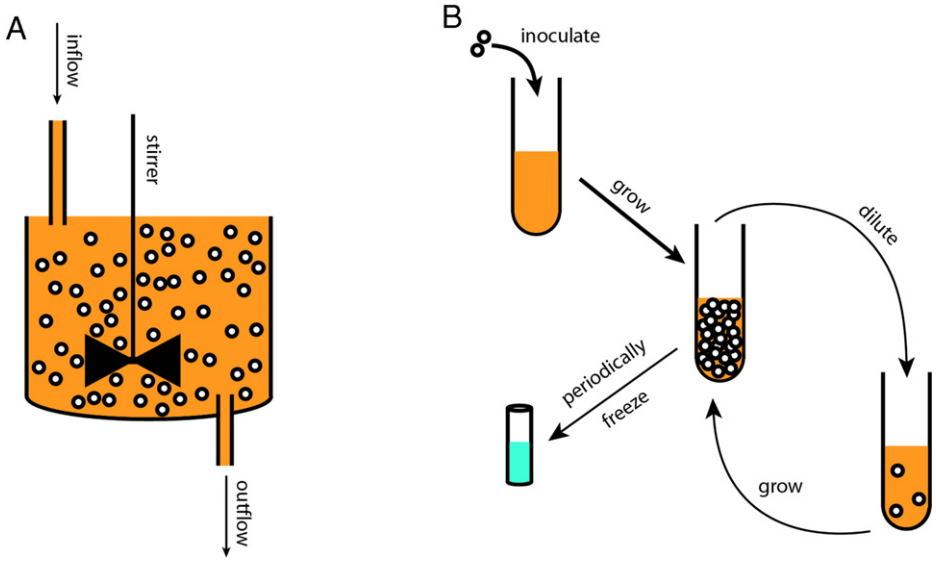

Figure 1. Schematic illustration of microbial evolution experiments in liquid culture. (A) A chemostat. New medium is slowly dripped into a stirred culture vessel, and old medium is slowly removed. The population is maintained in a steady state. Samples of the outflow can be frozen periodically for future analysis. (B) Batch culture. On a regular schedule, a small aliquot of cells is removed from a saturated culture and inoculated into new sterile medium. This new culture will become saturated with cells and the process will be repeated. An aliquot is occasionally removed and frozen for future analysis.

be taken to avoid contamination and cross-contamination of lines. The details necessary to achieve this depend on the organism and the type of robot being used, but over the past decade these difficulties have been fully resolved in a number of systems. As a result, experimental evolution has embraced the liquid handling robot. It is now fairly common to evolve lines in 96- or 384-well microplates, using a robot (often a Biomek or Tecan) to dilute 96 populations at a time (see e.g. [19]–[22]).

Some experimentalists have attempted to fully automate the entire process, using robotic arms to move microplates between incubators, plate readers, and liquid handling systems [23]. In principle this makes it possible to run an experiment for months without any manual input, though in practice this is difficult to achieve. Others use the robots simply for the liquid handling, and move the microplates around by hand. This latter approach allows the highest possible throughput to date: a single researcher can maintain of the order of five thousand lines in parallel indefinitely [20].

To achieve still higher replication, alternative approaches will be required. Recent work has begun to exploit the potential of micro- and milli-fluidic devices to generate millions of replicate populations, each maintained in a single small bubble of medium surrounded by oil [24]. These bubbles can be rapidly created in sizes ranging from picoliters to hundreds of nanoliters, depending on the device, and can be merged and split to 'dilute' each population, analogously to batch culture experiments [25, 26]. Numerous technical challenges remain to be fully resolved, but this approach is a promising direction for increasing throughput by several more orders of magnitude. Microfluidic devices which

J. Stat. Mech. (2013) P01003

operate as hundreds or thousands of miniature chemostats on a chip have also been explored in recent years, but these have not yet been exploited for long-term evolution.

2.2. High-throughput phenotypic analysis

Maintaining thousands of replicate populations is pointless unless we can also assay the changes in the evolved lines that result. We can in principle draw useful conclusions from observations of any phenotypic characteristic we choose to measure. There is a wide range of possibilities—for example, Levy and Siegal recently introduced a high-throughput screen for several hundred morphological phenotypes in yeast [27]. Others have focused primarily on quantities expected to be related to fitness, such as growth rate or yield [28].

As the quantity that selection acts on, fitness itself is a special phenotype, and extensive effort has been devoted to measuring it as quickly and accurately as possible. Although there is often controversy about precisely what fitness means (see e.g. [29]), within the context of simple batch culture or chemostat evolution experiments it has a straightforward operational definition: fitness measures the speed with which a given genotype tends to increase or decrease in frequency in competition with other genotypes. Nontransitive or frequency-dependent effects sometimes complicate the picture, but while these effects have been found in a number of specific systems (e.g. [30]–[36]), they are generally the exception rather than the rule, at least to the limits of current experimental resolution.

Since fitness in this context is defined in terms of the outcomes of competition, a natural way to measure it is to allow different strains to compete and measure the change in their ratio. This is typically achieved by labeling a 'reference' strain with a drug resistance or fluorescent marker. The reference strain is then mixed with an evolved line, and the ratio of the two is measured. The lines are then maintained in competition using the same protocol as during the experimental evolution, and after a short period the ratio of the two strains is measured again. The rate of change in the ratio of the two strains provides a measurement of competitive fitness. This procedure is illustrated in figure 2. This process is repeated for each evolved strain, always in competition with the same reference. Assuming that nontransitive interactions are negligible, this provides an estimate of all relative fitnesses.

This pairwise competitive fitness assay has two major bottlenecks: the need to measure the ratio of evolved to reference strain, and the need to evolve the strains in competition. This latter evolution can be carried out using the same high-throughput methods as described above. Since this is only carried out for a short period, it is often possible to achieve of the order of tenfold higher replication than the experimental evolution itself, so multiple replicate measurements or measurements of fitness at several evolutionary timepoints can all be made in parallel. Measurement of the ratio of strains is more difficult to carry out quickly and accurately. Recent experiments have labeled the reference population using a fluorescent protein, and counted the ratio of fluorescent reference to non-fluorescent evolved strains using commercially available 384-well systems designed for fluorescence activated cell sorters (FACS) [20, 21]. Using these FACS methods, it is possible to count of the order of 10^4 cells per population in 384 populations in less than 2 h. This is a dramatic improvement over what is possible with earlier methods based on drug resistance markers.

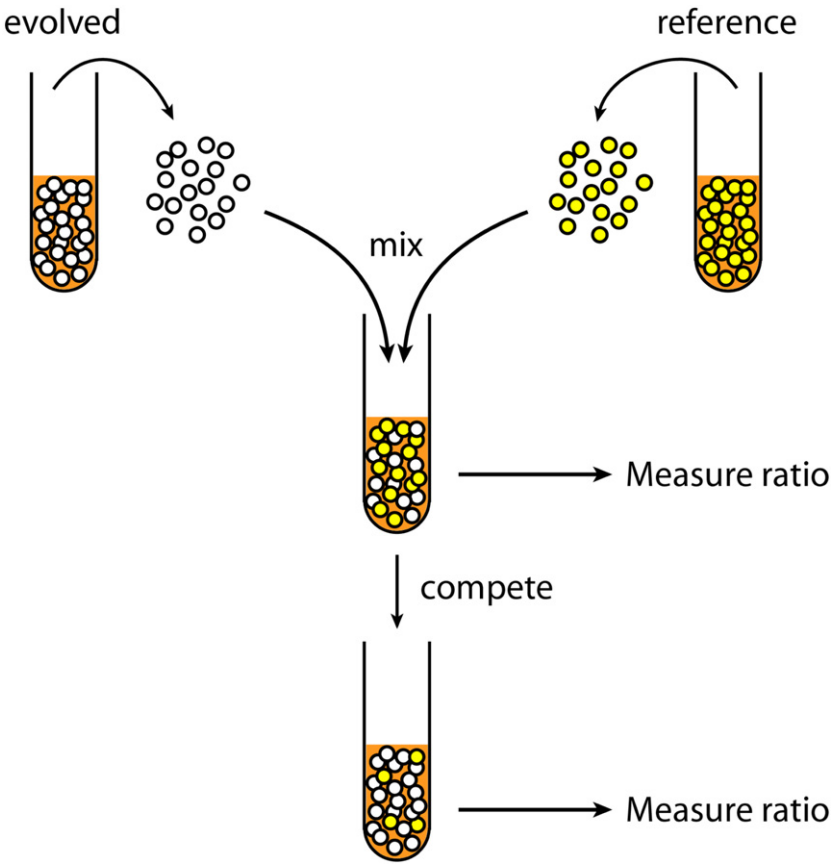

Figure 2. Schematic of a competitive fitness assay. The evolved strain to be assayed is mixed with a labeled reference strain, and the ratio of the two is measured. The strains are allowed to compete for a short period in the same environment in which the evolution was carried out, and the ratio is measured again. The fitness of the evolved strain relative to the reference is then determined from the rate of change in the ratio.

There are numerous potential sources of error in these competitive fitness assays. First, the ratios of reference to evolved cells will not be perfectly accurate, both due to stochastic counting error and due to systematic problems associated with FACS (e.g. dead cells, sticky cells, bubbles, and between-well carry-over). Second, the estimate of the rate of change in the ratios is also sensitive to errors in our estimates of the time in generations of the assay. Third, new mutations and selection on existing variation during the assay will affect our estimate of fitness. Finally, frequency dependence, density dependence, and nontransitive fitness interactions can all confound our measurements.

There is of course a tradeoff between accuracy and throughput. Recently Gallet *et al* [37] conducted a detailed analysis of the various sources of error in these competitive fitness measurements, and showed that it is possible to assay a few bacterial populations simultaneously with errors in absolute fitness of less than 10^{-3}. Using a somewhat less careful but higher-throughput approach, we have recently shown that a single experimenter can carry out of order five thousand replicate competitive fitness assays in one week with errors in absolute fitness of the order of 0.005 [20]. Since even tiny

differences in fitness can have dramatic evolutionary effects, continuing improvements in the throughput and accuracy of these measurements will be an important area of research for some time to come.

Throughput is achieved in pairwise fitness assays by carrying out many such assays in parallel. An alternative approach is to mix together a large number of strains simultaneously in a single competition. To measure the changes in the ratios of these strains, each must have a different marker, typically achieved using DNA barcodes. The frequencies of these barcodes can then be estimated using hybridization to a microarray designed for the purpose, or more recently by deep sequencing of the barcode region. This approach has been used for example to simultaneously assay the fitness of all ∼5000 strains in the yeast deletion collection in a variety of different environmental conditions [38]–[41]. Frequency measurements and hence estimates of fitness using these methods are less accurate (and also more time consuming and expensive) than the FACS-based fluorescence measurements in pairwise assays [42]. However, this is offset by the ability to measure the fitness of many strains simultaneously, and improvements in sequencing technology are rapidly reducing both the costs and the errors associated with this approach. It therefore seems likely that this type of measurement will become a new standard within the next few years.

2.3. High-throughput views of genetic changes

Recent years have also seen rapid advances in our ability to probe the evolutionary dynamics of the individual mutations underlying adaptation. A common approach has been to track changes in the frequencies of neutral markers—either those that arise naturally during the course of evolution [43, 44] or markers seeded into populations at the start of an experiment [45, 46]. From changes in the frequencies of these neutral markers we can infer the dynamics of the individual mutations which drive these marker shifts, though this inference can be somewhat complicated [47, 48]. We have recently introduced a related system in which a specific class of adaptive mutations alters a fluorescent marker, allowing the frequency of these mutations to be tracked as they arise naturally during evolution (figure 3) [21]. All of these 'marker-divergence' experiments act as a poor-man's sequencing. They provide an incomplete and low-resolution view of the dynamics, but to make up for this the markers are designed to be easy to measure in thousands of parallel lines, allowing us to draw statistical conclusions from the results.

Of course, the declining cost of sequencing is rapidly opening up new possibilities, and numerous studies have begun to use full genome sequencing to track genetic changes through time in evolving populations [10, 46, 49, 50]. However, costs remain a major obstacle, and this will be the case for the foreseeable future. A single lane of Illumina HiSeq can now generate of the order of 30 Gbp of sequence data for less than $2000 (corresponding to about 3000× coverage of a yeast genome). This can be split up using multiplexing however the experimenter desires: to deeply sequence a single entire population, to more shallowly sequence several multiplexed populations, or to sequence many individual clones. However, each additional multiplexed clone or population incurs substantial time and expense in library preparation and barcoding. Further, sequencing errors are a serious problem, requiring significant coverage for sequencing of individual clones and limiting our ability to detect low-frequency mutations using population sequencing even at very high depth.

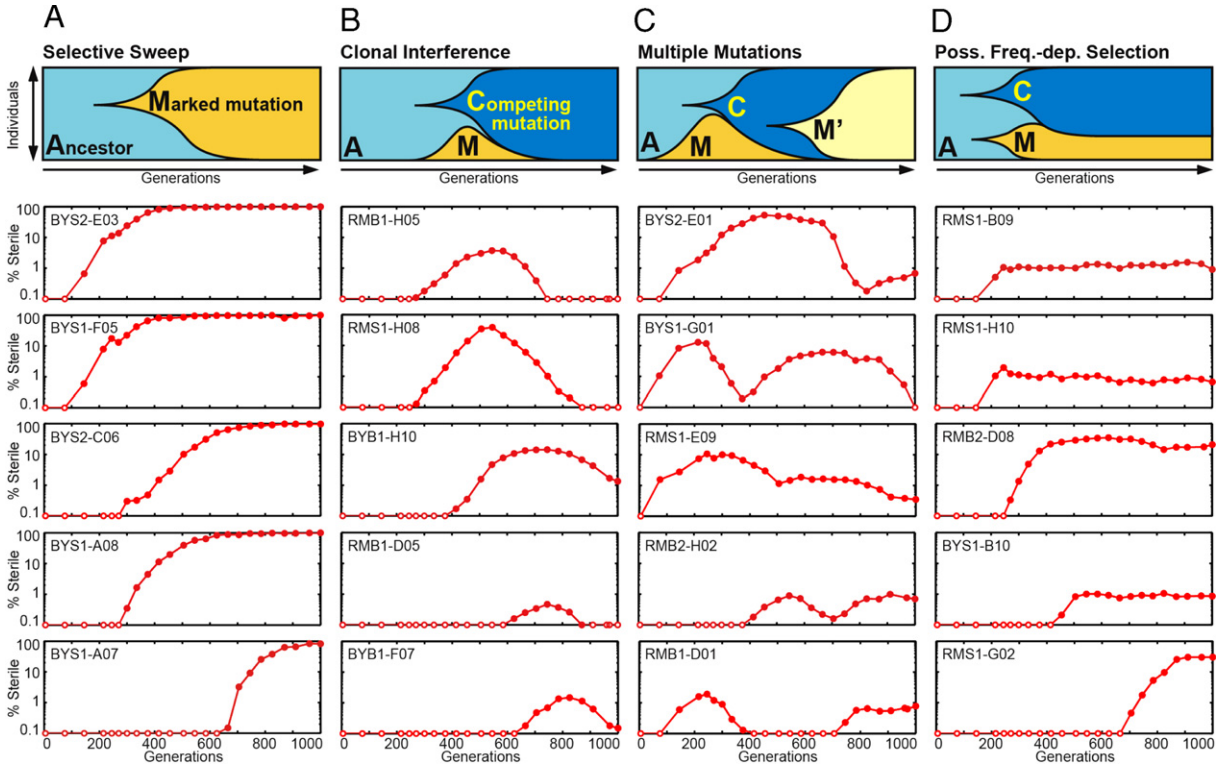

Figure 3. Examples of marker dynamics in a system where a spontaneously occurring beneficial mutation creates a measurable fluorescent marker. The top row illustrates a hypothetical scenario that could produce each of the four observed types of marker dynamic. Below each illustration are five representative examples of each dynamic. For each experimental population, the frequency of sterile mutants is shown as a function of time. Note the logarithmic y-axis. (A) The simplest case is a selective sweep. (B) More commonly, marked beneficial mutations are outcompeted by more fit lineages, a process known as clonal interference. (C) In some cases we observed the reemergence of marked mutants after clonal interference. This could occur either by a second marked mutation (as shown in the illustration at the top) or by an additional beneficial mutation arising in the original marked lineage. (D) Long-term maintenance of marked mutations at a given frequency could indicate the action of frequency-dependent selection, where the difference in fitness between the marked and unmarked subpopulations is a function of the frequency of the subpopulations. The figure and parts of the caption are reproduced from [21].

Methods for handling these problems are developing extremely rapidly, but at present it is difficult to sequence more than several hundred yeast populations in a single study at sufficient depth to detect mutations at frequencies above about 5–10%. It is even more challenging to detect mutations at frequencies as low as 1%. The situation in bacterial populations is only slightly better. Alternatively, a single study might sequence as many as a few hundred individual yeast or bacterial clones [51]. While methods are quickly improving, for the foreseeable future we are unlikely to be able to sequence evolving lines through experimental timecourses at anything approaching the throughput with which we are able to carry out the evolution and assay changes in fitness.

3. Simple statistical questions in experimental evolution

The high-throughput experimental methods described above can be exploited to address more complex questions about evolutionary dynamics than previously possible, or to address simple questions with greater resolution. The former approach often involves problems at the interface between ecology and evolution, such as how spatial structure and social interactions alter evolutionary dynamics (e.g. [19, 22]). Here, I focus primarily on the latter approach: experiments focused on the detailed quantitative analysis of simple evolutionary questions. These experiments aim to use high-throughput experimental evolution to study statistical questions related to recent theoretical advances in a way that was not previously possible.

Experiments in this spirit are often undertaken with the aim of 'testing' theory. The idea is to reproduce the situation envisioned by theory as closely as possible in the laboratory, and compare the results with theoretical predictions. This approach is often unproductive in practice, because any inconsistency between experiments and theory usually only implies that the experiments were not set up carefully enough to accurately represent the situation the theory aimed to describe. In other words, the experiments serve as little more than wet simulations.

To make better use of experiments, it is helpful to separate two basic types of question one can ask about evolutionary dynamics. The first concerns the raw material on which evolution acts. That is, what is the spectrum of possible mutational trajectories available to a population—i.e. the mutations and combinations of mutations that can occur, the rates at which these mutations arise, and their effects on fitness in present and future environments? The second concerns the evolutionary process: given a particular spectrum of possibilities, how does evolution choose among these possibilities, and how does this choice depend on parameters such as population size and structure?

This second question is the focus of the extensive theoretical work on evolutionary dynamics that is the subject of many other articles in this special issue. Yet all of this theory faces a basic question: which scenarios matter? The relevant spectrum of possible mutational trajectories could vary dramatically from system to system. It is possible to imagine a nearly infinite array of situations, without any clear way to determine which effects are most essential. To address this problem, theory must be accompanied by parallel efforts to answer the first question: what is the spectrum of possible mutational trajectories (SPMT) in any given system? This SPMT is a simple but statistical quantity—we want to determine the rates and relative probabilities of the distribution of possible mutational trajectories. By studying this fundamentally empirical question, new high-throughput methods in experimental evolution can help to guide theoretical work.

Exactly how we should parameterize the SPMT is a major problem in itself, which I turn to momentarily. Regardless of our choice, however, the details of the SPMT depend on all of the complicated details of the biology of the organisms in question, and their interactions with each other and with the environment. There is no guarantee that any general rules exist. If not, it will be necessary to understand the specific SPMT relevant to any particular situation before making any predictions about how evolution might act. This is a daunting task. Fortunately, although the mechanistic details of the SPMT (e.g. the functional consequences of specific mutations) will likely vary dramatically

J. Stat. Mech. (2013) P01003

between populations, there is reason to hope that there are general rules governing certain statistical aspects of the SPMT. This reason is evolution: the set of genotypes extant in any population is not random, but is the result of a long evolutionary history which has favored certain sorts of changes throughout the past. Thus genotypes that are mutationally close to these existing genotypes may have statistical properties that reflect the general constraints imposed by past evolution. For example, arguments that extant genotypes will tend to be 'robust' and 'evolvable' (see [52] for a recent review) describe two specific aspects of this phenomenon.

To study systematically whether there are general rules governing typical characteristics of the SPMT, we need a sensible and quantitative way to characterize precisely what this spectrum of possibilities is. A number of different approaches are common in the literature. One body of work takes a global approach, focusing on the concept of a general mapping between genotypes and fitnesses, known as the 'fitness landscape' [53]. Other work takes a reductive approach, assuming a specific small set of possible mutational trajectories and asking how evolution chooses between them [21], [54]–[56]. Still others rely on some form of biological intuition, ranging from models based on a detailed understanding of some set of genetic networks [57, 58] to models based on more abstract intuition about phenotypes, such as Fisher's geometric model [59]–[61]. Each of these descriptions has advantages and disadvantages. Since at any given moment evolution can usually only choose among a small number of closely related genotypes, my preference is instead usually to describe the SPMT by the statistical properties of the distribution of fitnesses in the local mutational neighborhood of extant genotypes: the *local statistical structure* of the spectrum of possible mutational trajectories. This characterization is particularly common in the literature on both experimental evolution and population genetics, since it involves simple measurable quantities such as the mutation rates and distribution of fitness effects of new mutations. We can think of it more generally as analogous to a Taylor expansion of the fitness function around a specific point on the hypercube of all possible genotypes, where the distribution of fitness effects of new mutations represents the statistics of the linear Taylor coefficients, and higher order coefficients correspond to more complex epistatic interactions.

Experimental work aimed at measuring aspects of the local statistical structure of the SPMT has already had important effects in uncovering general rules, and in highlighting parameter regimes that were originally thought to be of limited importance, but now appear to be pervasive in microbial evolution. Below, I describe how earlier lower-throughput experiments were able to uncover a surprising general fact about one of the very simplest aspects of the SPMT: the beneficial mutation rate. This has since guided much theoretical work. I then turn to highlighting other aspects of the local statistics of the SPMT where the higher resolution of more recent experimental methods has the potential to uncover other important general properties, and guide future theory into new productive directions.

3.1. Beneficial mutation rates and the importance of genetic variation

Fifteen years ago, adaptation in microbial populations was widely assumed to be limited by the availability of rare beneficial mutations, and laboratory adaptation experiments were interpreted in terms of 'periodic selection' [43], [62]–[66]. In this picture, adaptation

was assumed to consist of a sequence of selective sweeps: a new beneficial mutation would occasionally occur, spread rapidly through the population, and fix; this entire process would then repeat.

This view of adaptation also dominated theoretical work in both evolutionary dynamics and population genetics, leading to extensive analysis based on concepts such as the 'adaptive walk' (see [67] for a review). Of course, theorists have long recognized that in large populations, beneficial mutations might not be rare. Fisher and Muller first described how this leads to complicated 'clonal interference' dynamics where many different beneficial mutations occur and segregate simultaneously in a population, interfering with each other's fixation [59, 68]. These interference effects received attention in the context of the evolutionary advantage of recombination [69]–[73], and in population genetics more broadly [74]. However, these effects were not recognized as widely important throughout microbial adaptation, and much theoretical work focused instead primarily on the mutation-limited 'strong selection weak mutation' regime [67].

This view began to change about fifteen years ago, as Lenski's group found numerous signatures of clonal interference in their long-term experiments [75]. The discovery that mutations which increase mutation rate by one or more orders of magnitude are often selected for in microbial adaptation experiments also suggested that beneficial mutations may occur much more frequently than previously believed [76]–[78]. Before long, a number of further experiments in a variety of different systems confirmed the pervasive importance of these interference effects in laboratory evolution experiments [12, 46], [79]–[83]. These studies suggested that beneficial mutations with effects greater than about 1% occur at strikingly high rates, between $U_b = 10^{-4}$ and 10^{-6} per genome per generation, in a range of bacterial and yeast evolution experiments. These surprisingly high rates were orders of magnitude higher than previously believed, but have been corroborated by many other studies [84]–[87].

These experimental observations have spurred broad interest in the theory of clonal interference, beginning with the work of Gerrish and Lenski [88] and Levine and collaborators [89, 90]. Since then, many authors from the statistical mechanics community have analyzed how these interference effects determine the rate at which a population increases in fitness, the distribution of fitness effects of fixed mutations, the probability of hitchhiking of neutral and deleterious mutations, and other related questions ([91]–[98]; see [99] for a recent review). The interactions between these theoretical developments and experimental work were recently reviewed in [100].

One of the main qualitative implications of this work is that clonal interference maintains substantial genetic variation within populations even in the face of selection, blurring the traditional distinction between adaptation on new mutations and adaptation on standing genetic variation. This observation continues to spur both theoretical and experimental effort, particularly focusing on how interference effects lead to a 'traveling wave' of fitness within the population that maintains substantial variance even as it moves towards higher fitness [81, 92, 93, 97]. We have recently analyzed how this variation in fitness in turn influences the fate of each individual mutation (whose fates must then collectively determine the population-wide variation) [101], building on a 'bubble-surfing' approach introduced by Neher *et al* [102]. We have simultaneously begun to use the high-throughput methods described above to experimentally probe this interaction between genetic variation and the fate of each new mutation [21].

Recent work in population genetics is now suggesting that these interference phenomena are important far more broadly than just microbial evolution experiments, and may be generally important in interpreting patterns of molecular evolution [103]–[106]. These developments have thus far been largely independent of the work described above, but the two areas are beginning to connect. For example, the interaction between genetic variation within the population and the fate of each new mutation is at the heart of much recent work by myself and others that uses structured coalescent approaches to describe patterns of molecular evolution in the presence of strong interference effects [107]–[111].

3.2. The statistics of epistasis

Work on interference effects was inspired by one general observation about a particularly simple aspect of the SPMT, that beneficial mutations are remarkably common in a wide range of adapting microbial populations. To measure the SPMT in more detail, we must begin to exploit the higher resolution offered by more recent high-throughput methods. Extensive recent work along these lines has aimed to characterize the most local aspect of the SPMT, the distribution of fitness effects of new mutations (the 'DFE') [44], [112]–[118]. The DFE, along with the corresponding beneficial and deleterious mutation rates, is clearly crucial to understanding how evolution acts on short timescales [118]–[121]. Theoretical arguments based on extreme value theory suggest that a general principle should exist: the DFE for beneficial mutations should be exponential [112]–[114], [122, 123]. A number of experimental tests have largely supported this theoretical expectation, though the results are not entirely conclusive [44], [115]–[117], [124, 125]. Thus at least this aspect of the local statistical structure of the SPMT does have some general properties which are common to a broad range of experimental systems. This has influenced a number of recent theoretical studies (e.g. [96, 101, 126]).

Evolution on longer timescales can depend on the next most local aspect of the SPMT, the epistatic interactions between mutations (i.e. the effects of combinations of mutations above and beyond the sum of their individual effects, such as when two mutations are individually deleterious but jointly beneficial). These interactions are fundamental to the evolution of complex features involving mutations at several loci [127], such as new receptor–ligand interactions [128, 129], compensated antibiotic resistance [130, 131], multiple mutations required for immune escape [132], or multiple gene knockouts enabling cancer evolution [133]. To systematically measure the statistics of this slightly less local aspect of the SPMT requires even higher experimental resolution than needed to measure the DFE. Nevertheless, recent work has provided some intriguing insights which suggest that epistasis is pervasive throughout microbial evolution. For example, a number of studies have measured epistasis between deleterious mutations [134]–[141] and between single gene deletions in yeast [142]–[145], finding interactions ranging from generally synergistic to generally antagonistic. Several more recent studies have focused instead on epistasis between beneficial mutations, which is more directly relevant to adaptation, finding pervasive strong epistatic interactions between beneficial mutations fixed during the adaptation of a single gene [56, 128, 146, 147]. For example, Dean [146] described strong epistatic interactions between mutations fixed during natural evolution of isopropylmalate dehydrogenase, and Weinreich [56] discovered pervasive epistasis among mutations involved in the evolution of antibiotic resistant *E. coli*. These studies paint a

picture of strong 'sign' epistasis between mutations in a given gene, where each mutation is only beneficial in combination with specific others. This has led to the view that epistasis may substantially slow adaptation, since most potential mutational trajectories from a wildtype to an adapted multiple mutant are not uniformly favored by natural selection. Other recent studies have shown a somewhat different pattern of epistasis between mutations accumulated across the entire genomes of bacteria during experimental evolution [148]–[150]. In these systems, epistasis between beneficial mutations was also pervasive, but showed a general 'diminishing returns' pattern: all the mutations involved would confer a large advantage if they occurred before others, but a smaller advantage if they occurred later. This diminishing returns pattern also suggests that epistasis slows adaptation.

This work makes it clear that epistasis is strong and common during natural and experimental evolution. However, even if epistasis is common between specific pairs of mutations, this does not necessarily imply that it slows (or speeds) adaptation, nor is it clear how epistasis affects the repeatability of evolution at genotypic and phenotypic levels. For example, a given mutational change may have strong epistatic interactions with numerous other mutations, but this merely represents a 'reshuffling' of their fitness consequences, leaving the DFE invariant (figure 4(A)). In this case, epistasis opens up new avenues for adaptation even while it closes others, leading parallel populations down different genotypic but similar phenotypic paths, and despite the fact that epistasis seems to dramatically reduce the accessible evolutionary trajectories, it does not in fact constrain adaptation. Alternatively, each mutation may have some tendency to change the DFE for future mutations (figure 4(B)). A few studies have identified specific examples of situations where particular mutations do in fact affect 'evolvability' and hence help or hinder further adaptation [151]–[154], but it is not clear whether these situations are typical, or unusual special cases.

To guide future theoretical work into the most important directions, there is a need for experimental surveys of the overall statistics of epistasis that are relevant for constraining microbial evolution. Ideally this work will eventually go beyond the recent studies which have probed epistasis between a small and specific set of mutations or found particular examples of differences in evolvability to provide a more comprehensive picture of the overall statistics of epistatic interactions and the degree to which they alter evolutionary dynamics. However, attempts to provide such a high-resolution survey by directly measuring large numbers of individual epistatic interactions are still beyond the reach of even the best experimental methods. Instead, it will be essential to infer key aspects of the overall statistics of epistasis by combining experimental observations of evolutionary dynamics with theoretical predictions for how these dynamics depend on the underlying epistatic interactions.

Since epistasis affects evolutionary dynamics primarily by offering different possible mutational trajectories to different initial genotypes, experimental studies of the interactions between genetic variation and epistasis offer an ideal system for this type of inference. One possible approach is to exploit recent experimental studies of the repeatability of evolution at both genotypic and phenotypic levels between identical replicate lines [10, 51]. However, to infer the statistics of epistasis from this type of experiment, we must first develop a theoretical understanding of how the observed degree of variation and repeatability between replicate lines depends on the interplay

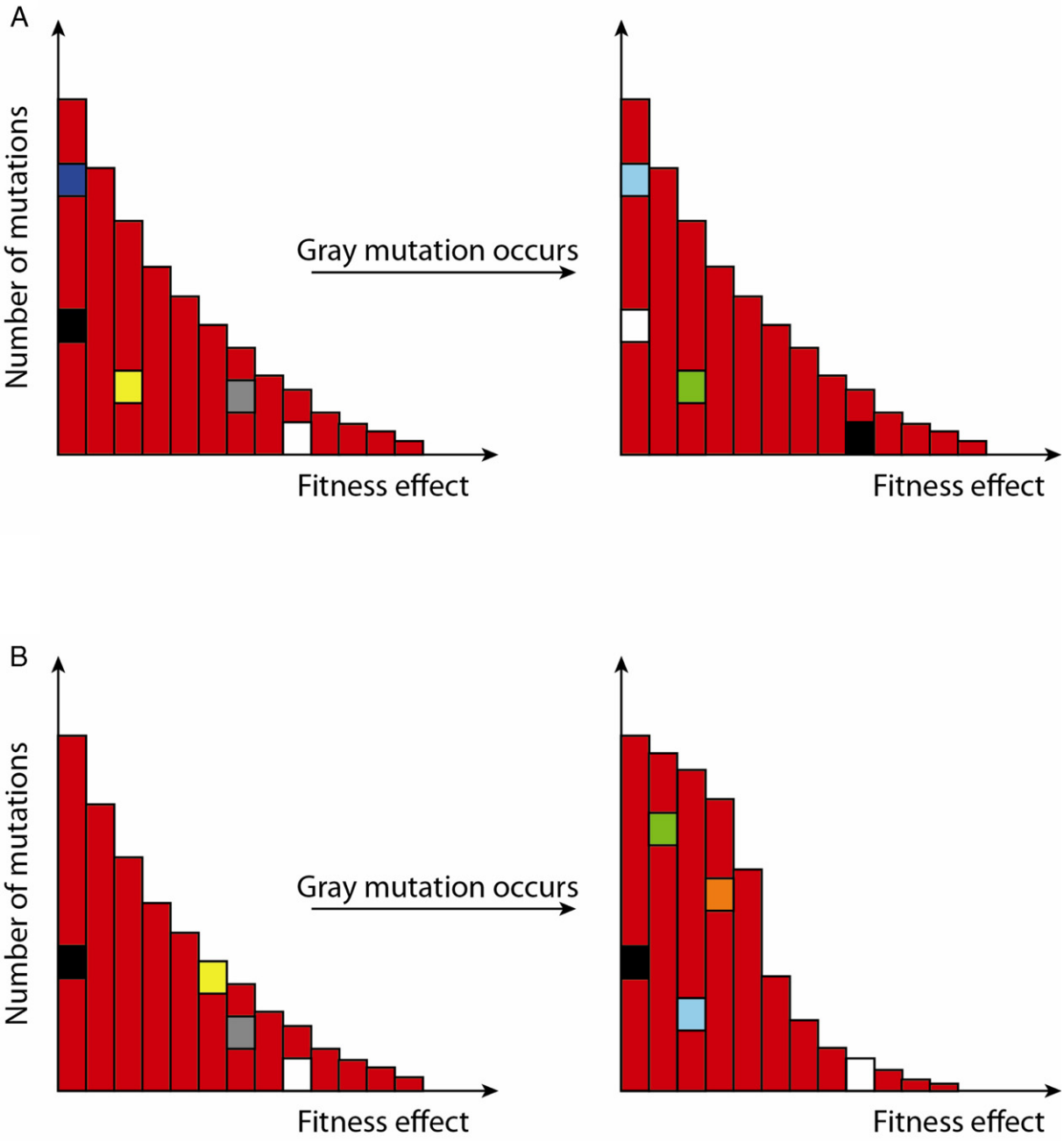

Figure 4. Illustration of how epistasis might alter the statistics of the DFE. Shown at the left is the distribution of fitness effects of new beneficial mutations, with mutations at several specific loci highlighted in different colors. At the right we show a hypothetical way in which the DFE shifts after the gray mutation occurs. These shifts in the DFE are the local statistical effects of epistasis that can affect future evolutionary dynamics. (A) A case where strong epistasis exists between the black and white highlighted mutations, and several previously beneficial mutations (blue and yellow) become deleterious while other previously neutral or deleterious mutations (cyan and green) become beneficial. However, this merely represents a reshuffling that does not alter the DFE and hence will not affect evolutionary dynamics. (B) A case where epistasis alters the DFE.

doi:10.1088/1742-5468/2013/01/P01003

between the statistics of epistasis and inherent evolutionary stochasticity. An alternative approach is to experimentally evolve subdivided populations at a variety of migration rates [20]. Subdivided populations maintain more genetic diversity, and hence explore epistatic interactions more thoroughly, at the cost of reducing the rate of fixation of beneficial mutations. Here too, a better theoretical understanding of how we expect this tradeoff to depend on the interaction between evolutionary randomness and epistasis will be necessary before we can use experiments along these lines to infer aspects of the overall statistics of epistasis. A third approach is to systematically mate individuals within and between evolved populations (e.g. in budding yeast), and measure the distribution of fitnesses of the resulting recombinant offspring. As with the other approaches, theoretical developments will be necessary to properly interpret the resulting data.

3.3. Windows into evolutionary dynamics

In addition to guiding theory by helping us to find important parameter regimes and general rules, higher-throughput methods in experimental evolution can be inspiring in a more open-ended way simply by opening new windows into evolutionary dynamics. For example, we now have the opportunity to watch the frequency dynamics of spontaneously arising mutations, or to survey similarities and differences in gene expression patterns between many identically evolved lines. Watching these new aspects of evolutionary dynamics can inspire theory to tackle new types of questions and goals. My own view is that this makes it a top priority to continue to push the bounds of replication and resolution in experimental evolution, even when it is hard to know or plan in advance what we expect to find.

Acknowledgments

I thank Daniel Fisher and Sergey Kryazhimskiy for useful conversations. This work was supported by the James S McDonnell Foundation, the Harvard Milton Fund, and the Alfred P Sloan Foundation.

References

[1] Dallinger W H, *President's address*, 1887 *J. R. Microscopical Soc.* **10** 185
[2] Dobzhansky T and Pavlovsky O, *An experimental study of interaction between genetic drift and natural selection*, 1957 *Evolution* **11** 311
[3] Rose M R, *Artificial selection on a fitness-component in drosophila melanogaster*, 1984 *Evolution* **38** 516
[4] Wade M J and Goodnight C J, *Wright's shifting balance theory: an experimental study*, 1991 *Science* **253** 1015
[5] Elena S and Lenski R E, *Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation*, 2003 *Nature Rev. Genet.* **4** 457
[6] Buckling A, MacLean R C, Brockhurst M A and Colegrave N, *The Beagle in a bottle*, 2009 *Nature* **457** 824
[7] Elena S F and Sanjuán R, *Virus evolution: insights from an experimental approach*, 2007 *Ann. Rev. Ecol, Evol. Syst.* **38** 27
[8] Garland J R and Rose M R (ed), 2009 *Experimental Evolution* (Berkeley, CA: UC Press)
[9] Lenski R, Rose M, Simpson S and Tadler S, *Long-term experimental evolution in Escherichia coli I. Adaptation and divergence during 2,000 generations*, 1991 *Am. Nat.* **138** 1315
[10] Gresham D, Desai M M, Tucker C M, Jenq H T, Pai D A, Ward A, DeSevo C G, Botstein D and Dunham M J, *The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast*, 2008 *PLoS Genet.* **4** e1000303

[11] Refardt D and Rainey P B, *Tuning a genetic switch: experimental evolution and natural variation of prophage induction*, 2010 *Evolution* **64** 1086
[12] Colegrave N, *Sex releases the speed limit on evolution*, 2002 *Nature* **420** 664
[13] Goddard M R, Godfray H C J and Burt A, *Sex increases the efficacy of natural selection in experimental yeast populations*, 2005 *Nature* **434** 636
[14] Cox E C and Gibson T C, *Selection for high mutation rates in chemostats*, 1974 *Genetics* **77** 169
[15] Perron G G, Gonzalez A and Buckling A, *Source–sink dynamics shape the evolution of antibiotic resistance and its pleiotropic fitness cost*, 2007 *Proc. R. Soc.* B **274** 2351
[16] Turner P E, Draghi J A and Wilpiszeski R, *High-throughput analysis of growth differences among phage strains*, 2012 *J. Microbiol. Methods* **88** 117
[17] Kirill S K, Melanie J I M, Nilay K, Andrew W M, Oskar H and David R N, *Selective sweeps in growing microbial colonies*, 2012 *Phys. Biol.* **9** 026008
[18] Novick A and Szilard L, *Description of the chemostat*, 1950 *Science* **112** 715
[19] Bell G and Gonzalez A, *Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration*, 2011 *Science* **332** 1327
[20] Kryazhimskiy S K, Rice D P and Desai M M, *Population subdivision and adaptation in asexual populations of saccharomyces cerevisiae*, 2012 *Evolution* **66** 1931
[21] Lang G I, Botstein D and Desai M M, *Genetic variation and the fate of beneficial mutations in asexual populations*, 2011 *Genetics* **188** 647
[22] Kerr B, Neuhauser C, Bohannan B J M and Dean A M, *Local migration promotes competitive restraint in a host–pathogen 'tragedy of the commons'*, 2006 *Nature* **442** 75
[23] Hegreness M, Shoresh N, Damian D, Hartl D and Kishony R, *Accelerated evolution of resistance in multidrug environments*, 2008 *Proc. Nat. Acad. Sci.* **105** 13977
[24] Agresti J J, Antipov E, Abate A R, Ahn K, Rowat A C, Baret J-C, Marquez M, Klibanov A M, Griffiths A D and Weitz D A, *Ultrahigh-throughput screening in drop-based microfluidics for directed evolution*, 2010 *Proc. Nat. Acad. Sci.* **107** 4004
[25] Guo M T, Rotem A, Heyman J A and Weitz D A, *Droplet microfluidics for high-throughput biological assays*, 2012 *Lab Chip* **12** 2146
[26] Baraban L, Bertholle F, Salverda M L M, Bremond N, Panizza P, Baudry J, de Visser J A G M and Bibette J, *Millifluidic droplet analyser for microbiology*, 2011 *Lab Chip* **11** 4057
[27] Levy S F and Siegal M L, *Network hubs buffer environmental variation in saccharomyces cerevisiae*, 2008 *PLoS Biol.* **6** e264
[28] Bell G, *Experimental genomics of fitness in yeast*, 2010 *Proc. R. Soc.* B **277** 1459
[29] Orr H A, *Fitness and its role in evolutionary genetics*, 2009 *Nat. Rev. Genet.* **10** 531
[30] Kerr B, Riley M A, Feldman M W and Bohannan B J M, *Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors*, 2002 *Nature* **418** 171
[31] Rainey P B and Travisano M, *Adaptive radiation in a heterogeneous environment*, 1998 *Nature* **394** 69
[32] Treves D S, Manning S and Adams J, *Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of Escherichia coli*, 1998 *Mol. Biol. Evolution* **15** 789
[33] Turner P E, Souza V and Lenski R E, *Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes*, 1996 *Ecology* **77** 2119
[34] Rozen D E and Lenski R E, *Long-term experimental evolution in Escherichia coli. VIII. Dynamics of a balanced polymorphism*, 2000 *Am. Nat.* **155** 24
[35] Friesen M L, Saxer G, Travisano M and Doebeli M, *Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in Escherichia coli*, 2004 *Evolution* **58** 245
[36] Levin B R, *Coexistence of two asexual strains on a single resource*, 1972 *Science* **175** 1272
[37] Gallet R, Cooper T F, Elena S F and Lenormand T, *Measuring selection coefficients below 10-3: method, questions and prospects*, 2011 *Genetics* **190** 175
[38] Smith A M, Heisler L E, Mellor J, Kaper F, Thompson M J, Chee M, Roth F P, Giaever G and Nislow C, *Quantitative phenotyping via deep barcode sequencing*, 2009 *Genome Res.* **19** 1836
[39] Giaever G *et al*, *Functional profiling of the saccharomyces cerevisiae genome*, 2002 *Nature* **418** 387
[40] Hillenmeyer M E *et al*, *The chemical genomic portrait of yeast: uncovering a phenotype for all genes*, 2008 *Science* **320** 362
[41] Steinmetz L M *et al*, *Systematic screen for human disease genes in yeast*, 2002 *Nat. Genet.* **31** 400
[42] Manna F, Gallet R, Martin G and Lenormand T, *The high-throughput yeast deletion fitness data and the theories of dominance*, 2012 *J. Evol. Biol.* **25** 892
[43] Pacquin C and Adams J, *Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations*, 1983 *Nature* **302** 495

[44] Imhof M and Schlotterer C, *Fitness effects of advantageous mutations in evolving Escherichia coli populations*, 2001 *Proc. Nat. Acad. Sci. USA* **98** 1113
[45] Hegreness M, Shoresh N, Hartl D and Kishony R, *An equivalence principle for the incorporation of favorable mutations in asexual populations*, 2006 *Science* **311** 1615
[46] Kao K C and Sherlock G, *Molecular characterization of clonal interference during adaptive evolution in asexual populations of saccharomyces cerevisiae*, 2008 *Nat. Genet.* **40** 1499
[47] Zhang W, Sehgal V, Dinh D M, Azevedo R B R, Cooper T F and Azencott R, *Estimation of the rate and effect of new beneficial mutations in asexual populations*, 2012 *Theor. Popul. Biol.* **81** 168
[48] Illingworth C J R and Mustonen V, *A method to infer positive selection from marker dynamics in an asexual population*, 2012 *Bioinformatics* **28** 831
[49] Barrick J E, Yu D S, Yoon S H, Jeong H, Oh T K, Schneider D, Lenski R E and Kim J F, *Genome evolution and adaptation in a long-term experiment with Escherichia coli*, 2009 *Nature* **461** 1243
[50] Barrick J E and Lenski R E, *Genome-wide mutational diversity in an evolving population of Escherichia coli*, 2009 *Cold Spring Harbor Symp. Quant. Biol.* **74** 119
[51] Tenaillon O, Rodríguez-Verdugo A, Gaut R L, McDonald P, Bennett A F, Long A D and Gaut B S, *The molecular diversity of adaptive convergence*, 2012 *Science* **335** 457
[52] Masel J and Trotter M V, *Robustness and evolvability*, 2010 *Trends Genet.* **26** 406
[53] Wright S, *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*, 1932 *Proc. 6th Intl Congr. Genetics* **1** 356
[54] Weissman D B, Desai M M, Fisher D S and Feldman M W, *The rate at which asexual populations cross fitness valleys*, 2009 *Theor. Popul. Biol.* **75** 286
[55] Michor F, Iwasa Y and Nowak M A, *Dynamics of cancer progression*, 2004 *Nature Rev. Cancer* **4** 197
[56] Weinreich D, Delaney N, Depristo M and Hartl D, *Darwinian evolution can follow only very few mutational paths to fitter proteins*, 2006 *Science* **312** 111
[57] Segre D, DeLuna A, Church G M and Kishony R, *Modular epistasis in yeast metabolism*, 2005 *Nat. Genet.* **37** 77
[58] Szathmary E, *Do deleterious mutations act synergistically? metabolic control theory provides a partial answer*, 1993 *Genetics* **133** 127
[59] Fisher R A, 1930 *The Genetical Theory of Natural Selection* (Oxford: Oxford University Press)
[60] Martin G and Lenormand T, *A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species*, 2006 *Evolution* **60** 893
[61] Martin G, Elena S F and Lenormand T, *Distributions of epistasis in microbes fit predictions from a fitness landscape model*, 2007 *Nat. Genet.* **39** 555
[62] Atwood K C, Schneider L K and Rryan F J, *Selective mechanisms in bacteria*, 1951 *Cold Spring Harbor Symp. Quant. Biol.* **16** 345
[63] Atwood K C, Schneider L K and Ryan F J, *Periodic selection in Escherichia coli*, 1951 *Proc. Nat. Acad. Sci.* **37** 146
[64] Novick A C and Szilard L, *Experiments with the chemostat on spontaneous mutations of bacteria*, 1950 *Proc. Nat. Acad. Sci. USA* **36** 708
[65] Koch A L, *The pertinence of the periodic selection phenomenon to prokaryote evolution*, 1974 *Genetics* **77** 127
[66] Levin B R, *Periodic selection, infectious gene exchange, and the genetic structure of E. coli populations*, 1981 *Genetics* **99** 1
[67] Orr H A, *The genetic theory of adaptation: a brief history*, 2005 *Nat. Rev. Genet.* **6** 119
[68] Muller H, *Some genetic aspects of sex*, 1932 *Am. Nat.* **66** 118
[69] Crow J F and Kimura M, *Evolution in sexual and asexual populations*, 1965 *Am. Nat.* **909** 439
[70] Maynard-Smith J, *Evolution in sexual and asexual populations*, 1968 *Am. Nat.* **102** 469
[71] Maynard Smith J, *What use is sex?*, 1971 *J. Theor. Biol.* **30** 319
[72] Bodmer W F, *The evolutionary significance of recombination in prokaryotes*, 1970 *Prokaryotic and Eukaryotic Cells* ed HP Charles and B C J G Knight (Cambridge: Cambridge University Press)
[73] Felsenstein J, *The evolutionary advantage of recombination*, 1974 *Genetics* **78** 737
[74] Hill W G and Robertson A, *The effect of linkage on limits to artificial selection*, 1966 *Genet. Res.* **8** 269
[75] de Visser J A, Zeyl C W, Gerrish P J, Blanchard J L and Lenski R E, *Diminishing returns from mutation supply rate in asexual populations*, 1999 *Science* **283** 404
[76] Denamur E and Matic I, *Evolution of mutation rates in bacteria*, 2006 *Mol. Microbiol.* **60** 820
[77] Sniegowski P D, Gerrish P J and Lenski R E, *Evolution of high mutation rates in experimental populations of E. coli*, 1997 *Nature* **387** 703

J. Stat. Mech. (2013) P01003

[78] Thompson D A, Desai M M and Murray A W, *Ploidy controls the success of mutators and nature of mutations during budding yeast evolution*, 2006 *Curr. Biol.* **16** 1581
[79] Miralles R, Gerrish P J, Moya A and Elena S F, *Clonal interference and the evolution of RNA viruses*, 1999 *Science* **285** 1745
[80] Miralles R, Moya A and Elena S F, *Diminishing returns of population size in the rate of RNA virus adaptation*, 2000 *J. Virology* **74** 3566
[81] Desai M M, Fisher D S and Murray A W, *The speed of evolution and maintenance of variation in asexual populations*, 2007 *Curr. Biol.* **17** 385
[82] Bollback J P and Huelsenbeck J P, *Clonal interference is alleviated by high mutation rates in large populations*, 2007 *Mol. Biol. Evol.* **24** 1397
[83] Betancourt A J, *Genomewide patterns of substitution in adaptively evolving populations of the RNA bacteriophage MS2*, 2009 *Genetics* **181** 1535
[84] Joseph S B and Hall D W, *Spontaneous mutations in diploid saccharomyces cerevisiae: more beneficial than expected*, 2004 *Genetics* **168** 1817
[85] Hall D W and Joseph S B, *A high frequency of beneficial mutations across multiple fitness components in saccharomyces cerevisiae*, 2010 *Genetics* **185** 1397
[86] Perfeito L, Fernandes L, Mota C and Gordo I, *Adaptive mutations in bacteria: high rate and small effects*, 2007 *Science* **317** 813
[87] Wloch D M, Szafranieca K, Bortsb R H and Korona R, *Direct estimate of the mutation rate and the distribution of fitness effects in the yeast saccharomyces cerevisiae*, 2001 *Genetics* **159** 441
[88] Gerrish P and Lenski R, *The fate of competing beneficial mutations in an asexual population*, 1998 *Genetica* **102/103** 127
[89] Kessler D, Levine H, Ridgway D and Tsimring L, *Evolution on a smooth landscape*, 1997 *J. Stat. Phys.* **87** 519
[90] Tsimring L S, Levine H and Kessler D A, *RNA virus evolution via a fitness-space model*, 1996 *Phys. Rev. Lett.* **76** 4440
[91] Desai M M and Fisher D S, *Beneficial mutation-selection balance and the effect of linkage on positive selection*, 2007 *Genetics* **176** 1759
[92] Rouzine I, Wakeley J and Coffin J, *The solitary wave of asexual evolution*, 2003 *Proc. Nat. Acad. Sci.* **100** 587
[93] Rouzine I, Brunet E and Wilke C, *The traveling-wave approach to asexual evolution: Muller's ratchet and the speed of adaptation*, 2008 *Theor. Popul. Biol.* **73** 24
[94] Park S C and Krug J, *Clonal interference in large populations*, 2007 *Proc. Nat. Acad. Sci.* **104** 18135
[95] Ridgway D, Levine H and Kessler D, *Evolution on a smooth landscape: the role of bias*, 1998 *J. Stat. Phys.* **90** 191
[96] Schiffels S, Szöllösi G J, Mustonen V and Lässig M, *Emergent neutrality in adaptive asexual evolution*, 2011 *Genetics* **189** 1361
[97] Hallatschek O, *The noisy edge of traveling waves*, 2011 *Proc. Nat. Acad. Sci.* **108** 1783
[98] Cohen E, Kessler D A and Levine H, *Front propagation up a reaction rate gradient*, 2005 *Phys. Rev.* E **72** 066126
[99] Park S-C, Simon D and Krug J, *The speed of evolution in large asexual populations*, 2010 *J. Stat. Phys.* **138** 381
[100] Sniegowski P D and Gerrish P J, *Beneficial mutations and the dynamics of adaptation in asexual populations*, 2010 *Phil. Trans. R. Soc.* B **365** 1255
[101] Good B H, Rouzine I M, Balick D J, Hallatschek O and Desai M M, *The rate of adaptation and the distribution of fixed beneficial mutations in asexual populations*, 2012 *Proc. Nat. Acad. Sci.* **109** 4950
[102] Neher R A, Shraiman B I and Fisher D S, *Rate of adaptation in large sexual populations*, 2010 *Genetics* **184** 467
[103] Hahn M W, *Toward a selection theory of molecular evolution*, 2008 *Evolution* **62** 255
[104] Karasov T, Messer P W and Petrov D A, *Evidence that adaptation in Drosophila is not limited by mutation at single sites*, 2010 *PLoS Genet.* **6** e1000924
[105] Sella G, Petrov D A, Przeworski M and Andolfatto P, *Pervasive natural selection in the Drosophila genome?*, 2009 *PLoS Genet.* **5** e1000495
[106] Pritchard J K, Pickrell J K and Coop G, *The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation*, 2010 *Curr. Biol.* **20** R208
[107] Barton N H and Etheridge A M, *The relation between reproductive value and genetic contribution*, 2011 *Genetics* **188** 953

J. Stat. Mech. (2013) P01003

[108] Desai M M, Nicolaisen L E, Walczak A M and Plotkin J B, *The structure of allelic diversity in the presence of purifying selection*, 2012 *Theor. Population Biol.* **81** 144
[109] Walczak A M, Nicolaisen L E, Plotkin J B and Desai M M, *The structure of genealogies in the presence of purifying selection: A fitness-class coalescent*, 2012 *Genetics* **190** 753
[110] Seger J, Smith W A, Perry J J, Hunn J, Kaliszewska Z A, Sala L L, Pozzi L, Rowntree V J and Adler F R, *Gene genealogies strongly distorted by weakly interfering mutations in constant environments*, 2010 *Genetics* **184** 529
[111] O'Fallon B D, Seger J and Adler F R, *A continuous-state coalescent and the impact of weak selection on the structure of gene genealogies*, 2010 *Mol. Biol. Evolution* **27** 1162
[112] Orr H A, *The distribution of fitness effects among beneficial mutations*, 2003 *Genetics* **163** 1519
[113] Gillespie J H, *Molecular evolution over the mutational landscape*, 1984 *Evolution* **38** 1116
[114] Beisel C J, Rokyta D R, Wichman H A and Joyce P, *Testing the extreme value domain of attraction for distributions of beneficial fitness effects*, 2007 *Genetics* **176** 2441
[115] Kassen R and Bataillon T, *Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria*, 2006 *Nat. Genet.* **38** 484
[116] Sanjuan R, Moya A and Elena S, *The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus*, 2004 *Proc. Nat. Acad. Sci. USA* **101** 8396
[117] Cowperthwaite M C, Bull J J and Meyers L A, *Distributions of beneficial fitness effects in RNA*, 2005 *Genetics* **170** 1449
[118] Betancourt A J and Bollback J P, *Fitness effects of beneficial mutations: the mutational landscape model in experimental evolution*, 2006 *Curr. Opin. Genet. Dev.* **16** 618
[119] Rokyta D R, *An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus*, 2005 *Nat. Genet.* **37** 441
[120] Pinkel D, *Analytical description of mutational effects in competing asexual populations*, 2007 *Genetics* **177** 2135
[121] Barrett R D H, M'Gonigle L K and Otto S P, *The distribution of beneficial mutant effects under strong selection*, 2006 *Genetics* **174** 2071
[122] Orr H A, *The population genetics of adaptation: the adaptation of DNA sequences*, 2002 *Evolution* **56** 1317
[123] Martin G and Lenormand T, *The distribution of beneficial and fixed mutation fitness effects close to an optimum*, 2008 *Genetics* **179** 907
[124] MacLean R C and Buckling A, *The distribution of fitness effects of beneficial mutations in Pseudomonas aeruginosa*, 2009 *PLoS Genet.* **5** e1000406
[125] Barrett R D H, Craig MacLean R and Bell G, *Mutations of intermediate effect are responsible for adaptation in evolving Pseudomonas fluorescens populations*, 2006 *Biol. Lett.* **2** 236
[126] Fogle C A, Nagle J L and Desai M M, *Clonal interference, multiple mutations and adaptation in large asexual populations*, 2008 *Genetics* **180** 2163
[127] Blount Z D, Borland C Z and Lenski R E, *Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli*, 2008 *Proc. Nat. Acad. Sci.* **105** 7899
[128] Bridgham J T, Carroll S M and Thornton J W, *Evolution of hormone-receptor complexity by molecular exploitation*, 2006 *Science* **312** 97
[129] Goh C-S, Bogan A A, Joachimiak M, Walther D and Cohen F E, *Co-evolution of proteins with their interaction partners*, 2000 *J. Mol. Biol.* **299** 283
[130] Schulz zur Wiesch P, Engelstadter J and Bonhoeffer S, *Compensation of fitness costs and reversibility of antibiotic resistance mutations*, 2010 *Antimicrob. Agents Chemother* **54** 2085
[131] Levin B R, Perrot V and Walker N, *Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria*, 2000 *Genetics* **154** 985
[132] McDonald B A and Linde C, *Pathogen population genetics, evolutionary potential, and durable resistance*, 2002 *Ann. Rev. Phytopathol.* **40** 349
[133] Knudson A G, *Two genetic hits (more or less) to cancer*, 2001 *Nat. Rev. Cancer* **1** 157
[134] DeVisser J, Hoekstra R and Van den Ende H, *Test of interaction between genetic markers that affect fitness in aspergillis niger*, 1997 *Evolution* **51** 1499
[135] DeVisser J, Hoekstra R and Van den Ende H, *An experimental test for synergistic epistasis and its application to chlamydomonas*, 1997 *Genetics* **145** 815
[136] DeVisser J, Hoekstra R and Van den Ende H, *The effect of sex and deleterious mutations on fitness in chlamydomonas*, 1996 *Proc. R. Soc.* B **263** 193

J. Stat. Mech. (2013) P01003

[137] Elena S and Lenski R, *Test of synergistic interactions among deleterious mutations in bacteria*, 1997 *Nature* **390** 395
[138] Elena S F, *Little evidence for synergism among deleterious mutations in a nonsegmented RNA virus*, 1999 *J. Mol. Evolution* **49** 703
[139] Whitlock M and Bourguet D, *Factors affecting the genetic load in Drosophila: synergistic epistasis and correlations among fitness components*, 2000 *Evolution* **54** 1654
[140] Wloch D, Borts R and Korona R, *Epistatic interactions of spontaneous mutations in haploid strains of the yeast saccharomyces cervisiae*, 2001 *J. Evol. Biol.* **14** 310
[141] Burch C and Chao L, *Epistasis and its relationship to canalization in the RNA virus phi6*, 2004 *Genetics* **167** 559
[142] Typas A *et al*, *High-throughput, quantitative analyses of genetic interactions in E. coli*, 2008 *Nat. Meth.* **5** 781
[143] Collins S R, *Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map*, 2007 *Nature* **466** 806
[144] Collins S R, Schuldiner M, Krogan N J and Weissman J S, *A strategy for extracting and analyzing large-scale quantitative epistatic interaction data*, 2006 *Genome Biol.* **7** R63
[145] Costanzo M *et al*, *The genetic landscape of a cell*, 2010 *Science* **327** 425
[146] Lunzer M, Miller S, Felsheim R and Dean A, *The biochemical architecture of an ancient adaptive landscape*, 2005 *Science* **310** 499
[147] Bloom J D, Gong L I and Baltimore D, *Permissive secondary mutations enable the evolution of influenza oseltamivir resistance*, 2010 *Science* **328** 1272
[148] Chou H-H, Chiu H-C, Delaney N F, Segrè D and Marx C J, *Diminishing returns epistasis among beneficial mutations decelerates adaptation*, 2011 *Science* **332** 1190
[149] Khan A I, Dinh D M, Schneider D, Lenski R E and Cooper T F, *Negative epistasis between beneficial mutations in an evolving bacterial population*, 2011 *Science* **332** 1193
[150] MacLean R C, Perron G G and Gardner A, *Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in pseudomonas aeruginosa*, 2010 *Genetics* **186** 1345
[151] Woods R J, Barrick J E, Cooper T F, Shrestha U, Kauth M R and Lenski R E, *Second-order selection for evolvability in a large escherichia coli population*, 2011 *Science* **331** 1433
[152] Burch C L and Chao L, *Evolvability of an RNA virus is determined by its mutational neighbourhood*, 2000 *Nature* **406** 625
[153] Barrick J E, Kauth M R, Strelioff C C and Lenski R E, *Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects*, 2010 *Mol. Biol. Evolution* **27** 1338
[154] McBride R, Ogbunugafor C B and Turner P, *Robustness promotes evolvability of thermotolerance in an RNA virus*, 2008 *BMC Evol. Biol.* **8** 231

J. Stat. Mech. (2013) P01003