Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Genomic investigations of evolutionary dynamics and epistasis in microbial evolution experiments

Elizabeth R Jerison[1,2,3] and Michael M Desai[1,2,3]

Microbial evolution experiments enable us to watch adaptation in real time, and to quantify the repeatability and predictability of evolution by comparing identical replicate populations. Further, we can resurrect ancestral types to examine changes over evolutionary time. Until recently, experimental evolution has been limited to measuring phenotypic changes, or to tracking a few genetic markers over time. However, recent advances in sequencing technology now make it possible to extensively sequence clones or whole-population samples from microbial evolution experiments. Here, we review recent work exploiting these techniques to understand the genomic basis of evolutionary change in experimental systems. We first focus on studies that analyze the dynamics of genome evolution in microbial systems. We then survey work that uses observations of sequence evolution to infer aspects of the underlying fitness landscape, concentrating on the epistatic interactions between mutations and the constraints these interactions impose on adaptation.

Addresses
[1] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, United States
[2] Department of Physics, Harvard University, Cambridge, MA 02138, United States
[3] FAS Center for Systems Biology, Harvard University, Cambridge, MA 02138, United States

Corresponding author: Desai, Michael M (mmdesai@fas.harvard.edu)

Current Opinion in Genetics & Development 2015, **35**:33–39

This review comes from a themed issue on **Genomes and evolution**

Edited by **Antonis Rokas** and **Pamela S Soltis**

http://dx.doi.org/10.1016/j.gde.2015.08.008

Introduction

Laboratory evolution experiments complement studies of natural populations, by making it possible to more directly investigate the underlying factors that shape observed patterns of genetic diversity. While these experiments neglect many important complexities of the natural world, they offer a number of advantages. In experimental systems, we can observe evolutionary dynamics in real time, and exploit a 'frozen fossil record' to resurrect ancestral types and directly compare them to their descendants. We can also maintain many populations in parallel, and replay the tape of life thousands of times in identical (or different) conditions. Finally, we can tune key evolutionary parameters such as population sizes and mutation rates, and assess their importance with other parameters held constant.

A large body of work has exploited these advantages to investigate how evolutionary history, chance, and natural selection influence evolutionary outcomes [1,2]. For example, recent work has examined how the distribution of fitness effects of available mutations determines the power of natural selection [3,4]. Additionally, multiple studies have analyzed the role of epistasis in creating 'historical contingency,' where an initial mutation constrains or potentiates future evolution [5•,6,7]. Finally, laboratory evolution has been used to investigate how parameters such as population size affect evolutionary dynamics [8,9]. All of these studies aim to describe how features of the evolutionary landscape combine to determine evolutionary outcomes — a general process that is not confined to laboratory systems. When used for this purpose, laboratory microbial evolution experiments have the potential to be a model system for understanding the structure and diversity of genomes.

Until recently, experimental evolution has been limited primarily to phenotypic measurements. Numerous studies have examined how fitness changes over time, and how the rate of adaptation depends on factors such as the population size, initial genotype, population structure, or environmental conditions. Many studies have also tracked the frequencies of observable markers (e.g. drug resistance or fluorescent reporters) through time [3,4,10,11] to draw inferences [3,12,13] about the evolutionary process. In recent years, however, advances in sequencing technology have made it possible to sequence clones or whole-population samples from hundreds of parallel experimental lines [14–16]. In this review, we summarize recent work that has begun to apply these technical advances to long-term laboratory evolution experiments, to directly observe how microbial genomes evolve in the laboratory. We focus particularly on the influence of epistasis on genome evolution (for a more general and comprehensive review, see [17]).

The dynamics of adaptation

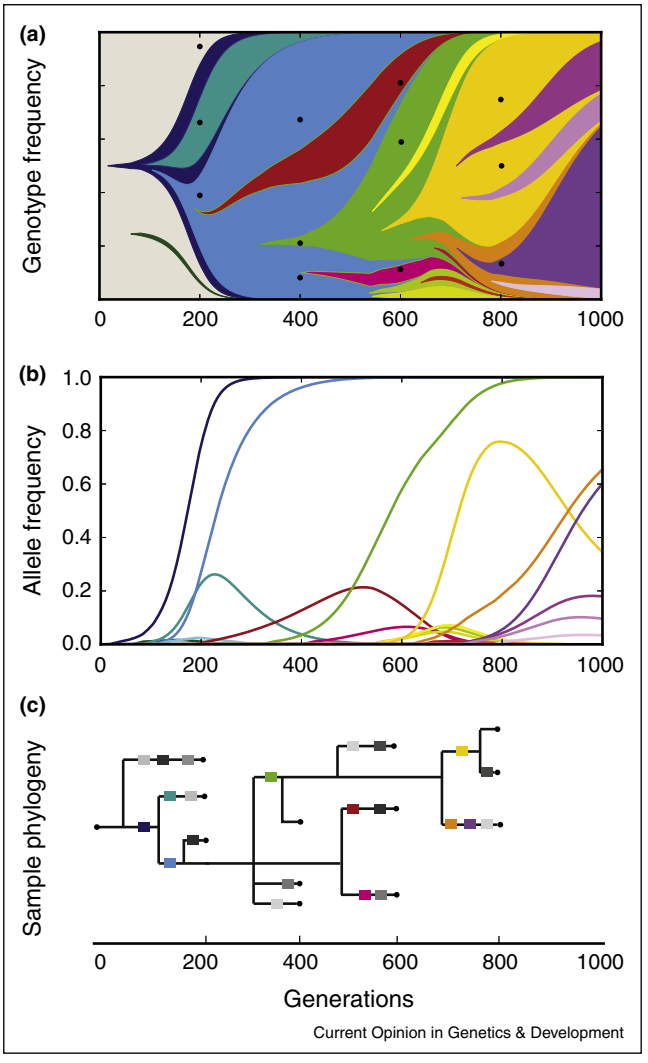
The population genetic forces of mutation, selection, and drift govern the dynamics of genome evolution, determining which mutations will survive competition to fix in

the population, and the signatures this process will leave on patterns of genetic diversity. An extensive population genetics literature has explored these connections theoretically. However, until recently it has been impossible to directly observe these dynamics in laboratory populations. Empirical studies of this type are necessary to determine which theoretically proposed regimes evolving populations actually experience. For example, is adaptation typically mutation limited, and hence in a regime where beneficial mutations fix independently, on their own merits? Or do beneficial mutations typically arise in multiple different lineages, resulting in a 'clonal interference' regime where many mutations compete for dominance? Studies that directly observe genome evolution in real time can distinguish between these and other possibilities. They can also point the way to complications that theory has overlooked, but that may be essential to the evolutionary process.

The first experiments to directly observe genomic evolution were in bacteriophage, where several studies Sanger-sequenced individual phage clones at multiple timepoints in several replicate populations to describe patterns of parallel and convergent evolution [18–23]. Next-generation sequencing now makes similar studies possible in microbial populations. Recent studies have exploited both whole-population ('metagenomic') sequencing of samples isolated at multiple timepoints from one or more evolving lines, and also sequencing of clones, to understand genomic evolution in these populations. These approaches offer somewhat different perspectives on the dynamics of molecular evolution (Figure 1).

The first major study to examine the dynamics of genome evolution in microbial laboratory populations focused on a single line of a long-term evolution experiment in *Escherichia coli*. By analyzing a clone sampled from each of five timepoints during 20,000 generations of adaptation, Barrick *et al.* [24••] showed that mutations continue to accumulate steadily through time despite a dramatic slowdown in the rate at which fitness increases. This divergence between phenotypic and sequence-level evolution points to the potential importance of epistasis in shaping adaptation [25–27]. More recent studies have sequenced clones isolated from multiple timepoints in several replicate populations [28,29]. This work demonstrates that adaptation is typically not mutation-limited: instead, there is clonal interference between competing beneficial mutations. Clonal interference affects which mutations fix in the population, and hence influences both molecular diversity and the dynamics of adaptation. In particular, the competition among beneficial mutations ensures that many are wasted, reducing the efficiency of selection and the predictability of molecular evolution. Clonal interference also affects the efficacy of other evolutionary processes, such as indirect selection on mutation rates [30].

Figure 1



Schematic of the evolutionary dynamics in a microbial evolution experiment. **(a)** Muller diagram depicting the frequencies of each genotype in the population over time. Only lineages that reach substantial frequency are shown (many lower-frequency lineages will typically also exist). **(b)** Allele frequencies in the population from (a), as they would be measured using whole-population metagenomic sequencing. This strategy reveals the dynamics of major alleles, but low-frequency mutations are undetectable. This metagenomic data also yields incomplete haplotype information: it is not always clear which mutations arise on which genetic backgrounds. **(c)** A phylogenetic tree built from clones that could be sampled from the population in (a; black dots). Colored boxes show the major mutations pictured in (a) and (b); grey boxes show 'private' mutations shared only by this clone and close relatives. This clone sequencing approach can be used to measure mutation rates and genetic diversity statistics such as heterozygosity, but provides limited information about allele frequencies over time.

Sequencing whole-population samples through time in multiple replicate populations offers an alternative view of the dynamics of adaptation [31••,32,33]. For example, one large-scale study in budding yeast highlighted the