Genome dynamics during experimental evolution

Jeffrey E. Barrick[1,2] *and Richard E. Lenski*[2,3]

Abstract | Evolutionary changes in organismal traits may occur either gradually or suddenly. However, until recently, there has been little direct information about how phenotypic changes are related to the rate and the nature of the underlying genotypic changes. Technological advances that facilitate whole-genome and whole-population sequencing, coupled with experiments that 'watch' evolution in action, have brought new precision to and insights into studies of mutation rates and genome evolution. In this Review, we discuss the evolutionary forces and ecological processes that govern genome dynamics in various laboratory systems in the context of relevant population genetic theory, and we relate these findings to evolution in natural populations.

Mutation accumulation
A type of evolution experiment in which populations are deliberately forced through a bottleneck of one or a few breeding individuals, which allows non-lethal mutations to accumulate with little or no filtering by natural selection.

[1]*Department of Molecular Biosciences, Institute for Cellular and Molecular Biology, The University of Texas, Austin, Texas 78712, USA.*
[2]*BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, Michigan 48824, USA.*
[3]*Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824, USA.*
e-mails: jbarrick@cm.utexas.edu; lenski@msu.edu
doi:10.1038/nrg3564
Published online
29 October 2013

Evolutionary and ecological questions that could previously be approached using only comparative or theoretical methods are increasingly amenable to direct study. In experimental evolution studies, populations of organisms are maintained in controlled environments in which changes in both genotype and phenotype can be monitored over timescales spanning many tens, hundreds or even thousands of generations[1,2]. Bringing evolution into the laboratory has several advantages, including both the ability to generate a 'fossil' record for later study and the ability to test the predictability of evolution across replicate populations. Studies of microorganisms also benefit from rapid generation times and the viability of frozen organisms, which can be revived either to allow an ancestor to compete head-to-head against its own descendants, or to 'replay' evolution that starts from various past states to investigate whether a particular outcome was contingent on some prior event.

How many, and what types of, genetic changes accumulate in evolving populations over time? The field of population genetics has developed a sophisticated mathematical framework for describing rates of evolutionary change in terms of the fundamental processes of mutation, recombination, genetic drift and natural selection[3]. This theory guides a general understanding of evolutionary regimes and dynamics, but specific outcomes in any given biological system may also crucially depend on the molecular details of a particular genome and on how it encodes metabolic, regulatory and developmental pathways. Both perspectives are necessary for unravelling contentious issues in evolutionary biology that are related to rates of sequence evolution, such as the relative importance of adaptive and non-adaptive processes and whether the predominant tempo of evolutionary change is gradual or episodic.

Recent advances in DNA sequencing technologies have now made it possible to identify genetic changes between ancestral and derived organisms on a whole-genome scale for any species[4,5]. We begin this Review by examining some of the previously hidden details that whole-genome and whole-population sequencing are revealing about evolution in even the simplest laboratory scenarios. We then discuss genetic dynamics in experiments that add back various components of the complexity that is present in the natural world. We primarily focus on asexual microbial systems in which most studies using extensive genome sequencing have been carried out so far. We also discuss multicellular eukaryotes and experiments in which sexual recombination has a role, as genomic data from these systems are increasingly becoming available.

Mutation rates

Most experimental evolution studies begin with clonal or inbred populations of a model organism so that there is a homogenous and well-characterized genetic starting point. Therefore, knowing the rates at which new mutations arise and lead to both genetic and phenotypic diversity in a population is useful for understanding evolutionary dynamics. Mutation accumulation experiments allow one to estimate the intrinsic rates and the effects of new mutations by repeatedly imposing population bottlenecks of one or a few randomly chosen breeding individuals to minimize selection that would otherwise favour

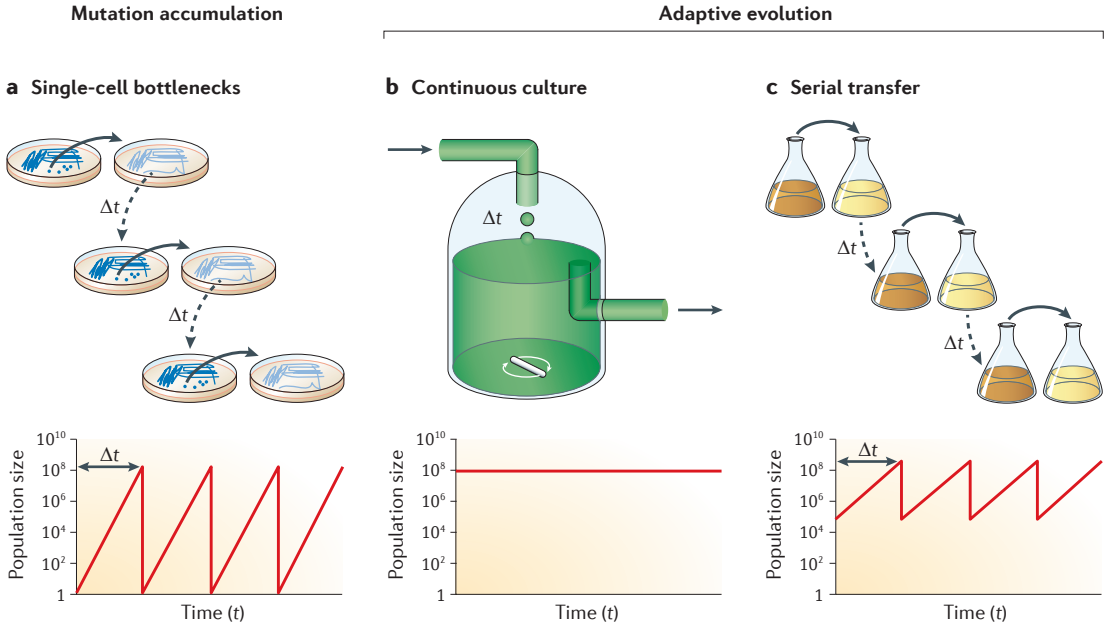


Figure 1 | **Types of evolution experiments.** There are three main ways that populations are propagated in evolution experiments, and they all lead to different types of genetic dynamics. The mechanics of how populations are maintained in each set-up are illustrated for microorganisms (top panels), and representative changes in population sizes over time are also shown for each procedure (bottom panels). Analogous procedures exist for multicellular organisms, although population sizes are generally much smaller. **a** | In mutation accumulation experiments, frequent and deliberate population bottlenecks through one or a few randomly chosen breeding individuals are accomplished by picking colonies of microorganisms that grow from single cells on agar plates. These bottlenecks purge genetic diversity and lead to the fixation of arbitrary mutations without respect to their effects on fitness. **b** | In experiments using continuous culture, populations are maintained in conditions that consist of a constant inflow of nutrients and an outflow of random individuals and waste in a chemostat, which leads to adaptive evolution and genetic diversity in populations that typically maintain a nearly constant size. **c** | In serial transfer experiments, a proportion of the population is periodically transferred to fresh media and allowed to regrow until the limiting nutrient is exhausted. Such batch growth also leads to adaptive evolution because ample genetic diversity is maintained through each transfer. Alternatively, transfers can be made before nutrient depletion, thereby allowing perpetual population growth. A second, cryptic type of population bottleneck occurs during adaptive evolution experiments (parts **b** and **c**) as a consequence of selective sweeps, especially in asexual populations, that drive out competing lineages and thereby reduce genetic diversity.

some variants[6] (FIG. 1a). Under these specific conditions, one can simply count the number of genetic changes that are present in independently evolved genomes after a known number of generations to estimate the spontaneous mutation rate (BOX 1). Recently, classic long-term mutation accumulation studies with model organisms — including *Saccharomyces cerevisiae*[7], *Arabidopsis thaliana*[8], *Drosophila melanogaster*[9] and *Caenorhabditis elegans*[10] — have been revisited using whole-genome sequencing to measure mutation rates. New mutation accumulation studies of microorganisms have also been carried out with the specific aim of estimating mutation rates[11–13].

The overarching conclusion of these experiments is that spontaneous mutation rates are usually very low. Mutation accumulation experiments with bacteria[11–13] and single-celled eukaryotes[7,13] typically find that the rate of single base mutations is of the order of 10^{-10}–10^{-9} per base pair per replication. Given that the typical genome sizes in these organisms are of the order of 10^6–10^7 base pairs, these rates correspond to only one point mutation in every few hundred to several thousand cell divisions, which is in reasonable agreement with earlier estimates for DNA-based microorganisms from reporter-gene assays[14]. Rates of point mutations in multicellular eukaryotes[8–10] are of the order of 0.05–1.0 per generation across the entire protein-coding portions of these genomes[13,15], which is still fairly low given the much longer generation times and the multiple cell divisions in the germ line between generations in these organisms. Some types of mutations, such as insertions and deletions of one or a few bases, typically occur at a lower rate than single base changes but vary more between species and with sequence context[7]. Other types of mutations, such as insertions of mobile DNA elements and large-scale chromosomal rearrangements, are more difficult to identify from short-read DNA sequencing data and have not yet been systematically examined in mutation accumulation experiments.

Mutation rates can change over evolutionary time, so it is instructive to understand how both genetic and environmental factors affect these rates. In particular, hypermutator lineages that have increased mutation rates and highly biased mutational spectra may arise

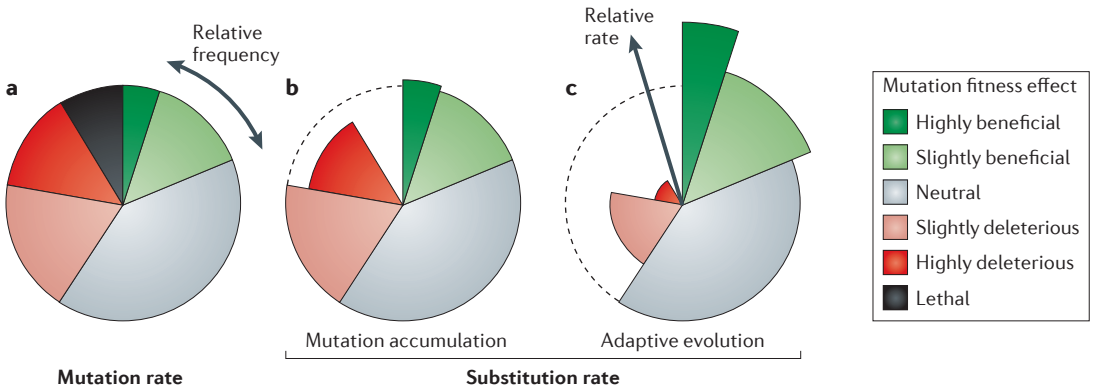
Population bottlenecks
Reductions in population size that typically also reduce genetic diversity. Bottlenecks can be deliberately imposed, such as in a mutation accumulation experiment. Cryptic bottlenecks also arise as a consequence of selective sweeps, especially in asexual populations, that drive out competing lineages and thus reduce genetic diversity.

Mutation rate
The rate at which new genetic mutations spontaneously occur during the replication and transmission of genetic information from parent to offspring.

Substitution rate
The rate at which new mutations accumulate in an evolving lineage over time, which typically depends on both the mutation rate and the effects of natural selection.

Biological fitness
A quantitative measure of the contribution of a specific organism or genotype to future generations owing to differential survival, reproduction or both, that is associated with its phenotype; fitness is often expressed relative to other organisms or genotypes.

Box 1 | **Mutation rates versus substitution rates**



It is important to distinguish between the rate at which spontaneous mutations occur and the rate at which genetic changes accumulate in a surviving lineage. The mutation rate reflects the probability of a change in genome sequence between a parent and its offspring. It is the compound result of unrepaired DNA damage, polymerase errors, intragenomic recombination events, movements of transposable elements and other molecular processes that introduce errors during the transmission of genetic information. However, only those mutations in lineages that persist — typically in the face of selection — contribute to the substitution rate that is measured by whole-genome sequencing. The failure to carefully distinguish between these two types of rates is a persistent cause of confusion and misconceptions about whether mutations are random. In the same vein, the frequency of a mutant allele in a population generally does not equal the rate at which the corresponding mutational event occurs.

Mutations can be broadly categorized as beneficial, neutral, deleterious or lethal with respect to their effects on biological fitness[88]. In various organisms, many or most new mutations are thought to be neutral or nearly neutral, and deleterious mutations greatly outnumber beneficial mutations under most circumstances[89] (see the figure, part **a**). Some mutations may change the magnitude or even the sign of the fitness effects of other mutations — a phenomenon known as epistasis[90]. Nevertheless, each new mutation in an evolving lineage can be classified into one of these categories depending on its fitness effect at the time and in the genetic context in which it appears. In this Review, we discuss two kinds of evolution experiments in which these different categories of mutations make different contributions to the substitution rate.

In mutation accumulation experiments, populations are continually forced through a bottleneck of one or a few breeding individuals, so the probability that any given mutation survives is essentially random and independent of its fitness effect. Thus, all mutations, except lethal or extremely deleterious ones, accumulate at rates that are close to their underlying mutation rates (shown as a dashed arc) in surviving lineages (see the figure, part **b**). The overall number of mutations in these highly unfavourable categories is usually thought to be small, and it is therefore common to equate substitution rates with mutation rates in mutation accumulation experiments, although this will slightly underestimate the true mutation rate. The ultimate mutation accumulation experiment is to sequence large numbers of parents and their offspring to avoid changes in environmental or genetic factors that might affect mutation rates during a longer experiment[91].

By contrast, in adaptive evolution experiments, beneficial mutations typically drive the genetic dynamics. The substitution rate of beneficial mutations exceeds the actual mutation rate (shown as a dashed arc) for this category because lineages with these rare mutations increase in frequency as they outcompete their ancestors and lineages with other mutations (see the figure, part **c**). Conversely, deleterious mutations are under-represented in adaptive evolution experiments because they are usually purged by selection, although slightly deleterious mutations can sometimes hitchhike with beneficial ones. The nature of competition between genetically diverged lineages will affect the extent of mutational diversity in a population, but the expected rate of accumulation of neutral mutations in any surviving lineage will still equal the underlying mutation rate for this category.

when mutations cause a loss of normal DNA repair or proofreading activities[16]. Mutation accumulation lineages that were derived from a *Salmonella enterica* subsp. *enterica* serovar Typhimurium hypermutator strain had a 30-fold increase in point mutation rates compared with a wild-type strain, and 91% of the resulting base substitutions were G:C→T:A transversions — a bias that is consistent with the misincorporation of oxidized guanine bases during DNA replication[11]. Another mutation accumulation study found that an *Escherichia coli* strain that is defective in mismatch repair had a 138-fold increase in mutation rate compared with wild type and had 70% A:T→G:C transitions[12].

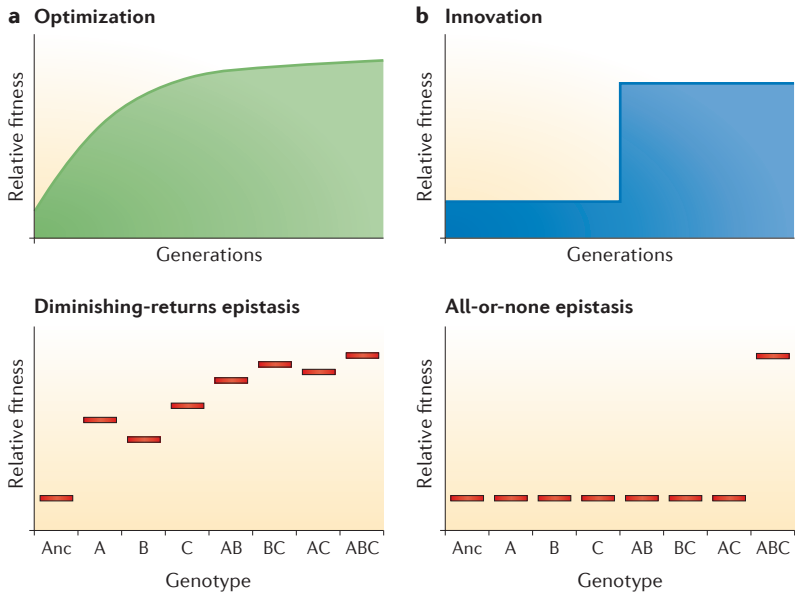
Chemical mutagens are the main environmental factor that has so far been examined with mutation accumulation experiments. At the extreme level of mutation that can be achieved in experimental evolution studies of mutagenized bacteriophage[17,18] — for example, mutating ~1% of the bases in the bacteriophage T7 genome[17] — one can begin to ask questions about what sites in a genome must be unchanged to remain viable. RNA viruses, such as HIV, typically have high mutation rates of the order of one per genome per generation[19], and experiments using mutagens have sought to test whether therapies that increase that rate further could lead to a mutational meltdown and the extinction of viral

Box 2 | **Adaptive evolution: optimization versus innovation**

When new beneficial mutations optimize the overall performance of existing genetic, metabolic and developmental networks during an adaptive evolution experiment, the fitness of organisms tends to gradually improve over time. In some experimental systems, interactions between the fitness effects of these mutations typically show diminishing-returns epistasis[92–94]; that is, each beneficial mutation increases fitness to a lesser extent in the presence of the other beneficial mutations than it would if it appeared alone in the ancestral genetic background (see the figure, part **a**). Typical beneficial mutations in the optimization regime modify gene expression levels, adjust regulatory interactions or alter metabolic fluxes. In some cases, these adjustments may be beneficial simply because they reduce the expression, and hence the energetic costs, of unused functions. The early beneficial mutations of largest effect are often in global control 'hubs' of networks, whereas later mutations often target the 'spokes' of specific pathways[36,85,90,95]. As these networks and pathways become more finely 'tuned' to a particular environment, it becomes more difficult to improve the overall system performance in a single mutational step. This form of gradual evolution is often associated with microevolutionary change.

More sudden and dramatic changes also sometimes occur during evolution experiments, particularly when beneficial mutations produce innovations that allow the organism to occupy a new ecological niche[28] (see the figure, part **b**). This form of adaptive evolution may involve, for example, the generation of a new connection or activity in a cellular network. Innovations may arise from mutations that show all-or-none epistasis[72]; that is, several mutations may first occur that, on their own, have little or no effect on the trait, but this evolved genetic background is required for some 'keystone' mutation to produce the phenotypic novelty. Models of RNA folding and regulatory circuits have been used to investigate the abstract properties of these so-called 'neutral networks' and how they can promote the evolution of novelty[96]. This form of evolution, in which new phenotypes appear suddenly, is sometimes associated with macroevolutionary change, and it can give rise to new opportunities for further adaptation and diversification[97].

Of course, multiple processes may be intertwined and their timescales may overlap. For example, some innovations may only be possible after a period of random exploration by mutation and genetic drift, as assumed in models of neutral networks. Alternatively, an innovation may have been enabled by earlier beneficial mutations that arose during an epoch of optimization for a different function, which is a kind of adaptive pre-adaptation. In this case, the large fitness gain obtained by co-opting these mutations for the innovation may dwarf the fitness gains during the earlier period of optimization. However, note that even during optimization, mutations of large effect can give a step-like appearance to fitness trajectories if they are measured with high resolution on short timescales[98,99].



A, B and C represent mutations that can occur alone or in combination in the genome. Anc, ancestor.

populations in infected individuals[20]. More generally, we anticipate that the sequencing of mutation accumulation lines will be used to enable a more precise version of the Ames test[21] to define the mutagenicity of potential carcinogens in the near future.

It is important to caution that laboratory mutation accumulation experiments may not precisely match the mutation rates or spectra that are found under natural conditions, in which organisms are often nutritionally deprived or otherwise stressed. For example, bacteria in the human gastrointestinal tract are thought to achieve only approximately one generation per day in this complex mixture of nutrients and biotic interactions[22], whereas bacteria in the soil probably undergo prolonged starvation and far fewer generations on average. Similarly, plants may be subject to extreme temperatures, damage from predation and increased ultraviolet radiation exposure in nature[8]. These and other environmental and genetic factors will probably be examined in future mutation accumulation experiments.

Adaptive evolution

In this section, we consider evolutionary dynamics in experiments in which differential survival and reproduction lead to the preferential accumulation of genetic variants that are better adapted to their environment. The simplest adaptive evolution experiments maintain populations that are derived from a single ancestral genotype in a uniform environment, such that selection pressures either remain constant or fluctuate in a controlled way. This situation can be achieved in continuous culture in which both the replenishment of resources and the removal of individuals happen at a constant rate (FIG. 1b) or by periodic serial transfer of a proportion of the population to a new microcosm with fresh resources (FIG. 1c). In adaptive evolution experiments, selection for mutations with beneficial effects drives the evolutionary dynamics (BOX 1). These dynamics are often visualized in terms of successive steps as populations 'climb' ridges and peaks in a fitness landscape[23]. Phenotypic evolution in a population may involve gradual optimization, discontinuous innovation (BOX 2) or perhaps some mixture of the two. For example, a discontinuous change, such as the ability to survive a previously lethal stress or to grow on a new resource, might be followed by a period of gradual refinement of that new ability.

Optimization regime. In the case of asexual organisms, the simplest situation occurs when the rate of appearance of beneficial mutations is low enough relative to both the fitness advantages of new beneficial mutations and the population size, such that there is effectively only one beneficial mutation present at a time. If this mutation survives stochastic loss by genetic drift while it is still rare, then it will begin a selective sweep, whereby its frequency increases until it reaches genetic fixation in the population (that is, the mutant completely replaces its ancestor), before another beneficial mutation becomes established (FIG. 2a). These dynamics have been called periodic selection after classic experiments that inferred sweeps of beneficial mutations in *E. coli* populations[24].

Diminishing-returns epistasis
Interactions among mutations such that the combined effect of the mutations on fitness or on some other trait is less than that expected from their individual contributions.

All-or-none epistasis
Interactions among mutations such that an entire set of mutations is required to confer a fitness advantage or a new trait; no subset that lacks one of these mutations has the advantage or an intermediate form of the relevant trait.

Adaptive evolution
Evolution under conditions in which surviving organisms accumulate genetic changes that lead to a fitness advantage over their progenitors.

Fitness landscape
The visualization of the genotype-to-fitness mapping for an organism in which the height of a position on the map represents the fitness of that genotype and the location is a reduced-dimensional projection of possible genotypes. An evolutionary trajectory of genetic changes can be visualized as a 'walk' and adaptation as a 'climb' in the fitness landscape.

Selective sweep
The increase in the frequency of an advantageous allele in a population as it displaces ancestral and competitor alleles.

Genetic fixation
The point at which an allele has completely displaced ancestral and competitor alleles; that is, the allele is present in every surviving individual in the population.

Periodic selection
The phenomenon whereby successive beneficial mutations completely sweep through an evolving population. Other mutations that are linked, but are not beneficial, can hitchhike with the beneficial driver mutation.

However, genetic dynamics in evolution experiments rarely seem to be in this simple regime. Owing to the many possible routes for adaptation, the rate at which beneficial mutations appear is typically high enough that before one beneficial mutation can sweep to fixation, another appears in a separate lineage (FIG. 2b). In asexual populations, competition between these alternative beneficial mutations means that the rate at which any one of these mutations spreads through the population is slowed because it must displace fitter competitors rather than only its ancestor[25]. This effect is called clonal interference, and the resulting genetic dynamics have been observed in several studies, most notably by deep sequencing entire yeast populations at frequent intervals to follow the frequencies of many new mutations[26].

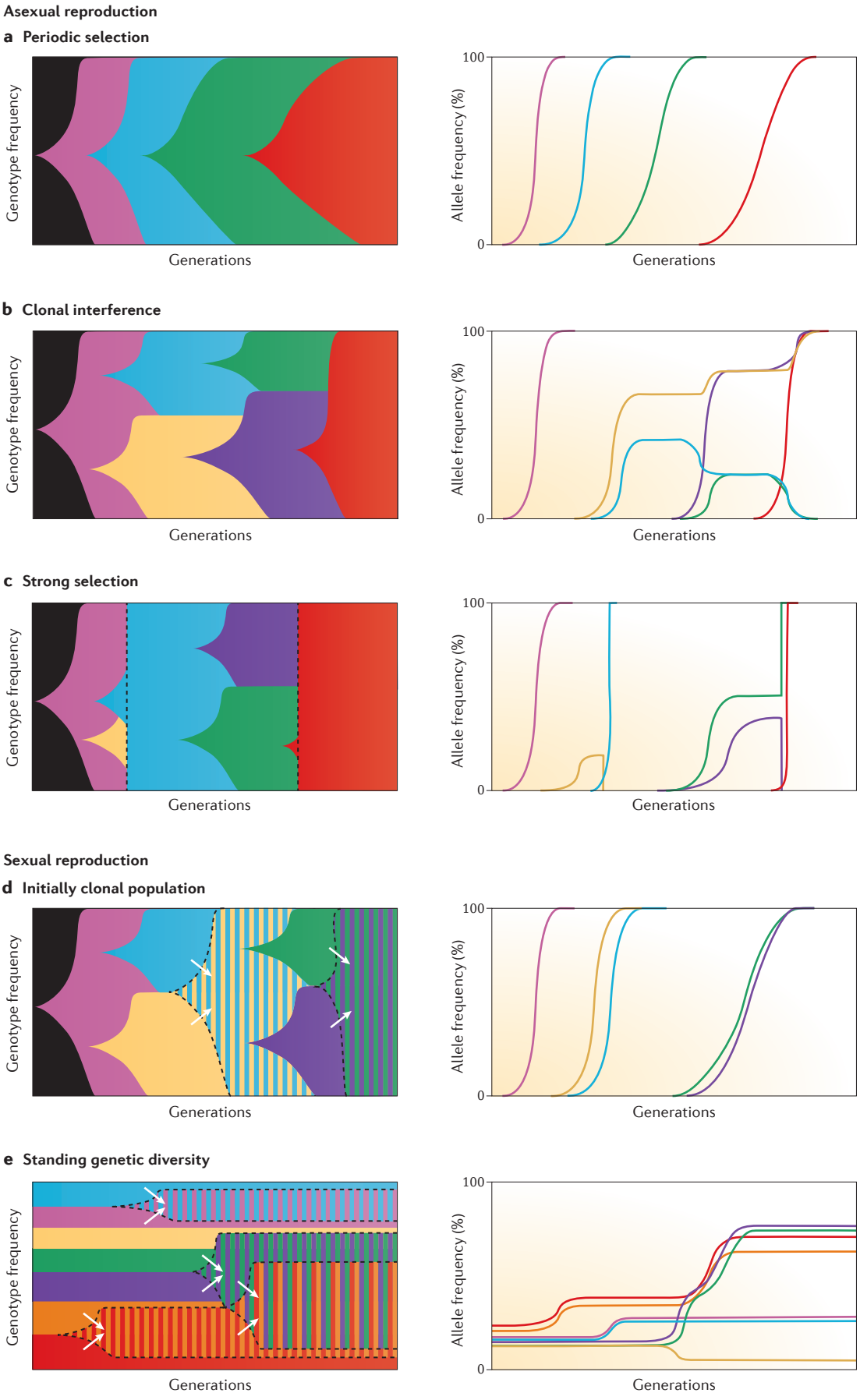
Genetic dynamics become even more complex when considering that neutral and deleterious mutations continually occur alongside the beneficial mutations discussed above. In large populations, deleterious mutations would rarely reach high frequency on their own, and neutral mutations would do so only over very long timescales. However, neutral and even deleterious mutations can rapidly hitchhike to prominence when they occur in the same genome as a beneficial mutation. The interplay of all of the factors discussed above can also give rise to apparently contradictory observations. For example, the rate of genomic change in an *E. coli* population in the long-term evolution experiment (LTEE) was surprisingly constant and clock-like — a signature that is sometimes taken as evidence of neutral evolution — even though most mutations that became fixed in the population were beneficial[27] (BOX 3).

Innovation regime. Some experiments have observed qualitatively new, often 'game-changing' abilities that have the hallmarks of evolutionary innovations[28]. Some innovations may require only a single mutation. For example, whole-genome sequencing of experimentally evolved *Myxococcus xanthus* strains found that 14 mutations were substituted after 1,000 generations in a liquid medium, while social motility and the capability to form fruiting bodies were lost, but only one mutation was involved in the subsequent restoration of those functions[29]. That key mutation did not revert any of the previous mutations but occurred instead in a previously uncharacterized small RNA[30]. The experimentally evolved transition of a chimeric *Ralstonia solanacearum* strain from a plant pathogen into a symbiont that is able to colonize root nodules also required only a single mutation in *hrpG*, which encodes a protein that regulates the expression of several virulence factors[31]. It is also likely that some examples of yeast that evolved a new multicellular 'snowflake' phenotype needed only single adaptive mutations[32].

The evolution of new metabolic capabilities was studied in several early experiments with microorganisms[33]. These organisms often gained the ability to use new compounds as nutrient sources by successive mutations in genes that activated their transcription under new conditions, increased their overall expression levels or altered their substrate specificities. In these examples, the relevant mutations each conferred a direct advantage in terms of using the new resource, but this need not be the case. The LTEE uses a glucose-limited medium, but citrate — another nutrient that *E. coli* generally cannot use — has also been present throughout the experiment. The ability to use this abundant but untapped resource evolved after 30,000 generations (~15 years) and in only one of the 12 replicate populations[34]. This innovation was difficult because it was contingent on one or more earlier 'potentiating' mutations that had to be present in the genetic background for the key 'actualizing' mutation to establish in the population. The key actualizing mutation arose by a chromosomal duplication event that 'rewired' gene expression by placing a new transcriptional promoter upstream of a previously silent citrate transporter gene to give the Cit^+ phenotype (that is, the ability to use citrate as a nutrient source)[35]. The earlier potentiating mutations did not confer any immediate advantage with respect to using citrate, but they may have been beneficial with respect to growth on glucose. If so, it is possible that other populations in the LTEE have also become potentiated and might evolve the Cit^+ phenotype, although no others have done so even after >50,000 generations.

Stressful environments. An important 'dial' that can be adjusted in experimental evolution studies is the strength of selection, particularly if the goal is to improve some phenotypic property. In one limit, selection may be so strong that it becomes a genetic screen in which only rare mutants with extreme, perhaps innovative, phenotypes can survive the stress (FIG. 2c). When selection is less stringent, more genetic diversity can usually be sustained, which allows more opportunities for populations to optimize by exploring alternative mutational paths. Indeed, divergent paths have been described in most evolution experiments with microorganisms, and these dynamics have been reconstructed in certain cases[36]. A potential disadvantage of a weak selection strategy, if the goal is to maximize phenotypic change, is that those mutations that confer the highest tolerance to stresses — such as organic solvent[37], high temperature[38,39], radiation exposure[40] and antibiotic pressures[41] — may not be favoured under this strategy. Conversely, if selection is too strong, then an evolving population might be driven towards a 'quick fix' — a local peak in the fitness landscape — that renders an even better solution less accessible. These concerns can potentially be balanced by constructing connected microenvironments with a gradient of conditions, such that organisms can colonize and exploit new resources in previously inhospitable regions by gaining new mutations that give them greater tolerance[41].

The effect of clonal interference depends on how closely matched the most beneficial mutation that occurs in a population is to the next-most beneficial mutation. Stressful environments can sometimes help to separate the more beneficial mutations from the less beneficial mutations. For example, a study of bacteriophage ΦX174 manipulated how harsh the environment was by restricting $CaCl_2$ (which is required for the efficient



Clonal interference
Competition between lineages that have different beneficial mutations in asexual populations, which slows the rate at which any particular allele fixes in the population relative to a freely recombining population.

Long-term evolution experiment
(LTEE). An experiment with *Escherichia coli* that has surpassed 25 years and 55,000 generations in duration.

Genetic background
The genotype of an organism; that is, its complete genome sequence or the alleles that distinguish it from other organisms.

Strength of selection
The benefit of accessible beneficial mutations relative to current mean population fitness. Under strong selection, sweeps of new genotypes generally occur more rapidly, and less diversity builds up in a population.

◀ Figure 2 | **Genetic dynamics in evolution experiments.** Five scenarios are illustrated using Muller plots (left-hand panels), which show the frequencies of different genotypes over time as coloured segments[69]. As new mutations appear, they are linked with mutations that previously arose in their predecessors. When there is sexual reproduction, existing mutational variants may also be recombined to produce new genotypes (as indicated by white arrows pointing to multiply shaded regions with dashed boundaries). The frequencies of different alleles (that is, mutational variants) in the population, as would be measured by metagenomic sequencing, are also shown for each scenario (right-hand panels). **a** | During periodic selection, if the rate at which new beneficial mutations appear is low and the fitness benefit of each mutation is large, then only one mutation will usually sweep through the population at a time. These dynamics cause near step-like trajectories for fitness, phenotypic traits and the number of beneficial mutations that accumulate over time. Successive sweeps typically take longer, as the expected marginal benefit of a later mutation decreases if evolution is in the optimization regime. **b** | In clonal interference, if the supply rate of beneficial mutations is higher because either the population size or the overall mutation rate is increased, then multiple beneficial mutations may arise before one of them achieves fixation. In asexual populations, competition between the contending mutations slows their progress towards fixation, which allows time for additional beneficial mutations to occur and gives rise to more complex trajectories for both fitness and mutation number. **c** | If strong selection is periodically imposed, in ways that may even be lethal to most of the population (shown by dashed vertical lines), then only one or a few genotypes may persist, and they can quickly achieve fixation after this selection-induced bottleneck. This scenario can lead to large and sudden changes in a phenotype, such as resistance to an antibiotic or to stress. **d** | The scenario of sexual reproduction in an initially clonal population is shown. As new beneficial mutations arise, they can be recombined into the same genetic background, rather than only competing with one another as in asexual populations. Thus, sexual reproduction may lead to more rapid genetic evolution and adaptation. **e** | In the case of sexual reproduction with standing genetic diversity, shuffling of the genetic diversity that is initially present in a population may generate fitter genotypes at a faster rate than waiting for new beneficial mutations to arise. Even so, if many different combinations of existing alleles give similar benefits, then no one allele will necessarily sweep to fixation on the timescale of the experiment.

attachment of ΦX174 to the *E. coli* host) and monitored genome sequence diversity over time[42]. It showed that clonal interference was more prevalent in benign environments, in which more beneficial mutations of similar effect were evidently available, and that this led to slower overall rates of genetic change.

Second-order selection for evolvability

When there is genetic diversity in evolving populations for long periods of time, there is the opportunity for selection to operate not only on the immediate effects of mutations or new combinations of alleles but also on how those new genotypes differ in their capability to further evolve (that is, their evolvability). In particular, prior mutational steps in a path on the fitness landscape may affect evolvability in at least two main ways. They may alter mutation rates (and/or recombination rates for organisms that are capable of sexual reproduction or horizontal gene transfer), or they may lead to differences in epistatic interactions with potential further mutations (BOX 2).

Mutation rates. It is fairly common for some populations in adaptive evolution experiments to become dominated by hypermutators that have increased mutation rates[43–45]. How do hypermutators invade populations? The immediate effect of an increased mutation rate on fitness is invariably negative, on average, because more mutations are deleterious than beneficial (BOX 1). However, the initial spontaneous mutation rate is typically low, as discussed above, so that the fitness cost of producing mutations at even a 100-fold higher rate is small. Furthermore, new hypermutator variants frequently arise because loss-of-function mutations in many genes give rise to this phenotype. Given the balance between their high rate of occurrence and small fitness costs, hypermutators might exist at frequencies of 10^{-6}–10^{-4} in experimental populations of *E. coli*[46,47]. However, the hypermutator subpopulation has increased evolvability because it has a much higher per-capita chance of producing the next highly beneficial mutation, or multiple beneficial mutations, than a non-mutator competitor (FIG. 3a). Thus, despite their slight fitness costs, hypermutators can sometimes hitchhike to fixation with the beneficial mutations that they generate[48].

In the long term, increased mutation rates are not without evolutionary risk. Opportunities for mutations that greatly improve fitness will eventually run out in the optimization regime. It may then be beneficial to compensate for a hypermutator defect and to become less evolvable in order to lower the genetic load[49] (FIG. 3b). Indeed, both mutation-rate scenarios have been observed in an *E. coli* population in the LTEE[50]. After thousands of generations, a mutation in the nucleoside triphosphate pyrophosphohydrolase gene (*mutT*) that caused a ~150-fold increase in mutation rates spread through the population, presumably by hitchhiking with one or more beneficial mutations. Later, parallel mutations in the adenine DNA glycosylase gene (*mutY*) arose in independent lineages; these mutations approximately halved the mutation rate, apparently by knocking out a mechanism by which misincorporations of oxidized nucleotides during DNA replication (caused by the original *mutT* defect) were misrepaired. The resulting reduction in genetic load was estimated to be ~0.5%. This value seems to be similar in magnitude to other beneficial mutations that drove adaptation late in the LTEE, whereas some beneficial mutations that were substituted early in the experiment had much greater fitness effects[27]. In an experiment with yeast, several populations that started as hypermutators also evolved lower mutation rates and, as a result, reduced genetic loads[51]. In natural populations, comparative evidence indicates that the complete reversion of a hypermutator to an ancestral mutation rate can occur by horizontal gene transfer of an intact gene from a non-mutator[52]. In addition to affecting mutation rates, hypermutators generally change the spectrum of different types of mutations, and those differences might also influence the evolvability of a lineage.

Genetic architecture. As fitness landscapes are complex and may have multiple peaks, some mutational paths may lead to 'dead ends' with no, or at least fewer, opportunities to further improve. In other cases, certain mutations may open up new opportunities for evolution that could not be accessed if other routes were taken. The term genetic architecture refers to how genotypes, and mutations that alter genotypes, map onto

Genetic load
The indirect fitness cost to an organism caused by producing offspring with mutations that either reduce their fitness or are lethal.

Genetic architecture
The properties of an organism, including its metabolic, regulatory and developmental pathways, that determine how new mutations affect phenotypes and fitness.

Box 3 | **Identifying adaptive mutations**

In molecular evolution and comparative genomics studies, the ratio of synonymous to non-synonymous base substitutions (dN/dS) in a protein-coding gene is commonly used to test whether the gene has been subject to positive selection or negative selection (also known as purifying selection). There are rarely enough base substitutions in an evolution experiment to apply this test on a per-gene basis. However, because so few mutations accumulate, it is also highly unlikely that the same gene would change in several independently evolved genomes unless these variants had been enriched by selection. Therefore, such genetic parallelism provides a strong signal that the mutations were beneficial. Depending on the phenotypic effect that is required for adaptation, this parallelism may occur at the level of an individual nucleotide or codon, a specific gene, or some step in a particular metabolic or regulatory pathway[27,63,72,100–102]. More complex patterns of covariation — such as a mutation in either one gene or another, but not both, in multiple independently evolved genomes — can be used to identify new genes that are involved in the same pathway[39,103].

However, hot spots that undergo unusually high rates of spontaneous mutations relative to the rest of a genome can also lead to genetic parallelism for mutations that are only slightly beneficial or not adaptive at all[104–106]. Therefore, the 'gold standard' for establishing whether a particular mutation is adaptive is to use either a genome-editing or a genetic exchange method to make an isogenic construct that differs from another strain by only the single mutation of interest. One can then either test for a change in a trait that is known to be related to fitness or compete the two organisms under the conditions that prevailed during the evolution experiment to determine whether the mutation is beneficial, neutral or deleterious.

In longer evolution experiments, after many mutations have accumulated, one must consider several potential complications when interpreting the results of these measurements. First, some mutations may be adaptive only in the context of the genetic background in which they arose, owing to interactions with mutations that occurred earlier in that lineage[90]. In this case, it might be cleaner to 'deconstruct' the mutation by removing it from an evolved genotype, in which one would then expect fitness to decrease. However, there is, again, the potential for interactions with mutations that evolved after the focal mutation to alter its measured fitness effect. Second, ecological interactions may affect fitness measurements. These often take the form of negative frequency dependence, in which a genotype has an advantage over some competitor only when it is rare in the population[56]. In other cases, non-transitive interactions may arise such that an evolved genotype is more fit than its immediate progenitor, but this genotype is less fit than some earlier ancestor that it never encountered because they were present at different times in the evolving population[107].

changes in phenotypes and fitness; hence, differences in evolvability can result from mutations that change the genetic architecture of an organism (FIG. 3c,d). We have already discussed above how the evolution of citrate use depended on a potentiated genetic background, which could be said to have resulted in greater evolvability[34,35] (FIG. 3c).

A more subtle change in evolvability involved two genotypes that competed early in the history of another *E. coli* population from the LTEE[53]. One of these genotypes prevailed despite having a significantly lower fitness. By replaying evolution many times from these two different starting points, it was demonstrated that the 'eventual winners' reproducibly gained more fitness over time than the 'eventual losers', such that this seemingly unexpected outcome in the original LTEE population was, in fact, the more likely outcome (FIG. 3d). Genome sequencing showed that the eventual winners often underwent subsequent mutations in the *spoT* gene, in which mutations were never observed in the eventual losers. The SpoT protein is a regulator of the stringent response. Reconstruction of this mutation in the two genetic backgrounds showed that it was highly beneficial in the winners but did not significantly affect the fitness of the losers (BOX 3). It remains to be determined exactly why the earlier mutations that distinguished the losers from the winners reduced their evolutionary potential, and how important such epistatic 'cul-de-sacs' are in natural populations.

Eco-evolutionary dynamics

In *The Origin of Species*[54], Darwin memorably envisioned a "tangled bank" in which organisms were "dependent upon each other in so complex a manner" as an outcome of natural selection. To this point, we have generally ignored ecological interactions beyond 'scramble' competition for limiting resources. Even in simple laboratory environments, evolution can lead to niche construction[55] that enables diverged lineages of organisms to stably coexist for long periods of time. Other experiments have examined evolution in environments with multiple resources or multiple interacting species. In both cases, metagenomic sequencing of DNA that was isolated from whole-population or whole-community samples, rather than from individual clones, is yielding new insights into the dynamic interactions between distinct ecotypes.

Multiple nutrients. The well-shaken flask environment of the LTEE nominally has a single niche, with a low concentration of glucose that limits growth. However, one of these *E. coli* populations gave rise to two distinct ecotypes, first noticed as small and large colony morphotypes after 6,000 generations, and these types coexisted for at least 30,000 generations[56]. The two types show negative frequency dependence, such that each type has a fitness advantage and can invade the other type when it is rare in the population[57]. In this case, the balance that leads to stable coexistence results from the large type growing faster on glucose and the small type having better growth on metabolic by-products[58]. The genetic and physiological bases of these differences are subject to ongoing investigation[57,58].

Metagenomic sequencing of another population in the LTEE revealed more transient diversification[59]. Mutations in genes that are related to acetate use repeatedly arose, but they never persisted for more than a few thousand generations or reached high frequencies. Acetate is a by-product excreted by *E. coli* during growth on glucose. As glucose runs out, acetate is taken up and used by cells. This cross-feeding opportunity for acetate 'specialists' suggests that other populations in the LTEE might also be on the cusp of evolving more complex ecologies.

In both of these cases from the LTEE, continued evolution after diversification feeds back on the ecological interactions and the stability of such interactions. In the case of the small and large polymorphism, the large types continually encroached on the niche occupied by the small type; if the small type had stopped evolving then it would have been driven to extinction[57]. In the case of acetate use, descendants that were derived from the main population apparently displaced the acetate specialists multiple times but repeatedly gave rise to new mutants that later reinvaded this niche[59]. Understanding

Isogenic construct
An organism, produced in the laboratory using various genetic tools, that has defined genetic differences from a reference organism. It is used to study the effects on fitness and on other phenotypic traits of single mutations or combinations of mutations.

Niche construction
The production of a new resource or other ecological opportunity that is caused by the actions or evolution of organisms.

Metagenomic sequencing
The sequencing of DNA fragments that are randomly derived from a population containing a mixture of many genotypes.

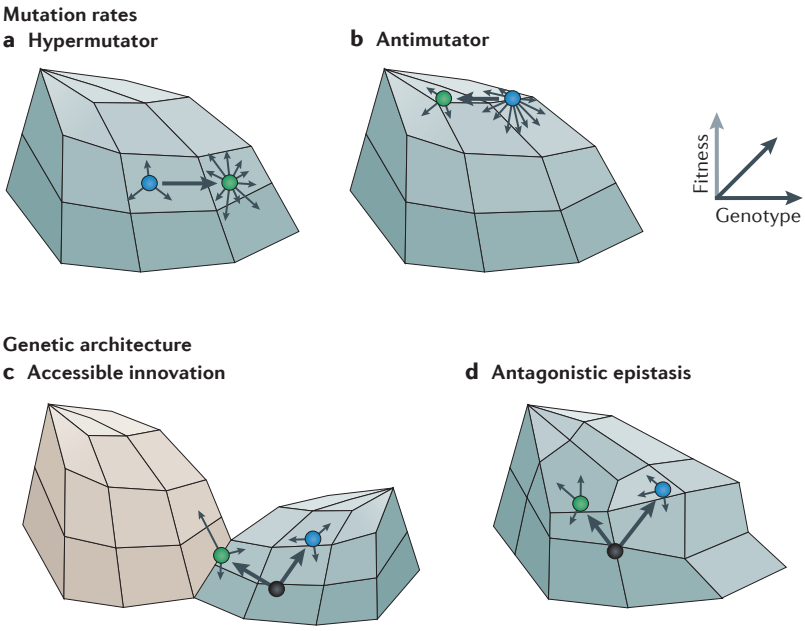


Figure 3 | **Second-order selection for evolvability.** The success of a new mutation or a new combination of alleles may depend on its effect on either the rate or the fitness benefits of subsequent mutations, in addition to its immediate effect on fitness. Several scenarios are illustrated as alternative mutational paths in fitness landscapes. Genotypes are represented by circles; thick arrows represent initial mutations that generate a new genotype, and thin arrows represent subsequent mutational paths that are available to the genotypes; the mutation rate of a genotype is reflected by the number of thin arrows projecting from it. **a** | A fitness landscape favouring a hypermutator is shown. From a progenitor with a low ancestral mutation rate (blue circle), a variant that causes an increased mutation rate (green circle) can sometimes take over an asexual population because it has a higher per-capita probability of generating beneficial mutations. Access to these opportunities may outweigh the immediate fitness cost of an increased genetic load. **b** | A fitness landscape favouring an antimutator is shown. In the longer term, as a genotype (blue circle) evolves to approach a local optimum (the peak of the illustrated fitness landscape) and there are fewer beneficial mutations available, a genotype with a lower mutation rate (green circle) may evolve and be favoured because it has a reduced genetic load. **c** | A fitness landscape promoting an accessible innovation is shown. Starting from the same progenitor genotype (black circle), two mutants may have different probabilities of eventual success owing to differences in their evolvability. In this case, one mutation (green circle) makes it possible for a subsequent mutation to 'invade' an open niche (beige landscape), whereas the other mutation (blue circle) does not. **d** | A fitness landscape with antagonistic epistasis is shown. Even in the same niche, one beneficial mutation (blue circle) may constrain opportunities for further fitness gain more than another (green circle) because of antagonistic epistatic interactions. In essence, some beneficial mutations may lead to 'cul-de-sacs' in the fitness landscape, which allows other beneficial mutations that do not limit further adaptation to prevail, provided that they coexist for enough time. If evolution can be 'replayed' many times starting with the two different genotypes, then an over-representation or an under-representation of mutations in specific genes would provide a signature of such epistatic effects.

Negative frequency dependence
An allele (or a trait) that undergoes a decline in fitness as it becomes more common in a population. If the allele confers an advantage when it is rare but is disadvantageous when it is common, then a genetic polymorphism is stably maintained.

the effects of newly evolved ecologies on evolutionary dynamics is an area that is worthy of both theoretical and empirical investigation.

Complex ecologies also evolve in closed cultures in which the nutrients are exhausted and are not renewed. In these cultures, the viable cell population declines over time, as starvation takes its toll. However, not all genotypes die at the same rate, and the survivors are enriched for growth advantage in stationary phase (GASP) mutants[60,61] that exploit by-products of dead and dying cells. Mutations that confer a GASP phenotype have been identified in *rpoS*, which encodes the alternative 'starvation' σ-factor, and in other genes that encode high-affinity amino acid transporters; all of these mutations increase the ability of the cell to obtain and use amino acids for carbon and energy[61]. In addition, genomic analyses reveal frequent gene-amplification variants among the survivors[61]. It seems that copy-number variants are generated at a high rate during starvation, and those that result in extra copies of genes that encode products which prove useful during starvation can then proliferate.

Substantial diversity also evolves in glucose-limited continuous-culture chemostats[62,63], even without the feast–famine seasonality in resource abundance that occurs in serial transfer studies. Metagenomic sequencing of populations that were propagated in chemostats at two different dilution rates found more genetic and phenotypic diversity in populations that were evolved at slow dilution rates than in those that were evolved at high dilution rates[64], which possibly indicates a higher mutation rate at the lower growth rate, more distinct strategies for dealing with this challenge or some combination of the two. Diversification in this system is driven by regulatory mutations that lead to different balances in the trade-off between stress resistance and nutrient use, including *rpoS* mutations[64].

Diversity can be encouraged experimentally by providing a mixture of nutrients to create multiple niches. Evolution experiments in which *E. coli* populations are serially propagated in a mixture of glucose and acetate reliably give rise to two strategies: glucose specialist 'slow switchers' that slowly shift from using glucose to acetate and 'fast switchers' that use acetate earlier than the slow switchers[65]. Metagenomic sequencing of many populations showed clear molecular signatures of parallel evolution in each ecotype and suggests that some mutations in one ecotype forced evolution in the other ecotype to follow certain mutational pathways[66].

Spatial gradients. An alternative approach to generating multiple niches is to establish spatial heterogeneity that results in distinct microenvironments[67,68]. *Pseudomonas fluorescens* rapidly evolves into three ecotypes — recognizable by their distinct colony morphologies — that populate different physical zones within unshaken flasks owing to heterogeneity in the availability of oxygen[68]. Diversification also occurred when *Burkholderia cenocepacia* evolved under daily selection for both the ability to disperse and the ability to then colonize a new surface as a biofilm[69]. In this case, three distinct colony morphotypes also rapidly emerged. However, metagenomic sequencing revealed more complex genetic dynamics that were not apparent from the time course of phenotypic diversification. Evolved genotypes that have a 'studded' colony morphotype gave rise to new versions of all three ecotypes, which then drove lineages that had previously exploited other niches to extinction. In essence, the studded type seems to be a 'source' population that continuously generates new variants which periodically displace the current lineages in secondary niches that become evolutionary 'sinks' (FIG. 4a).

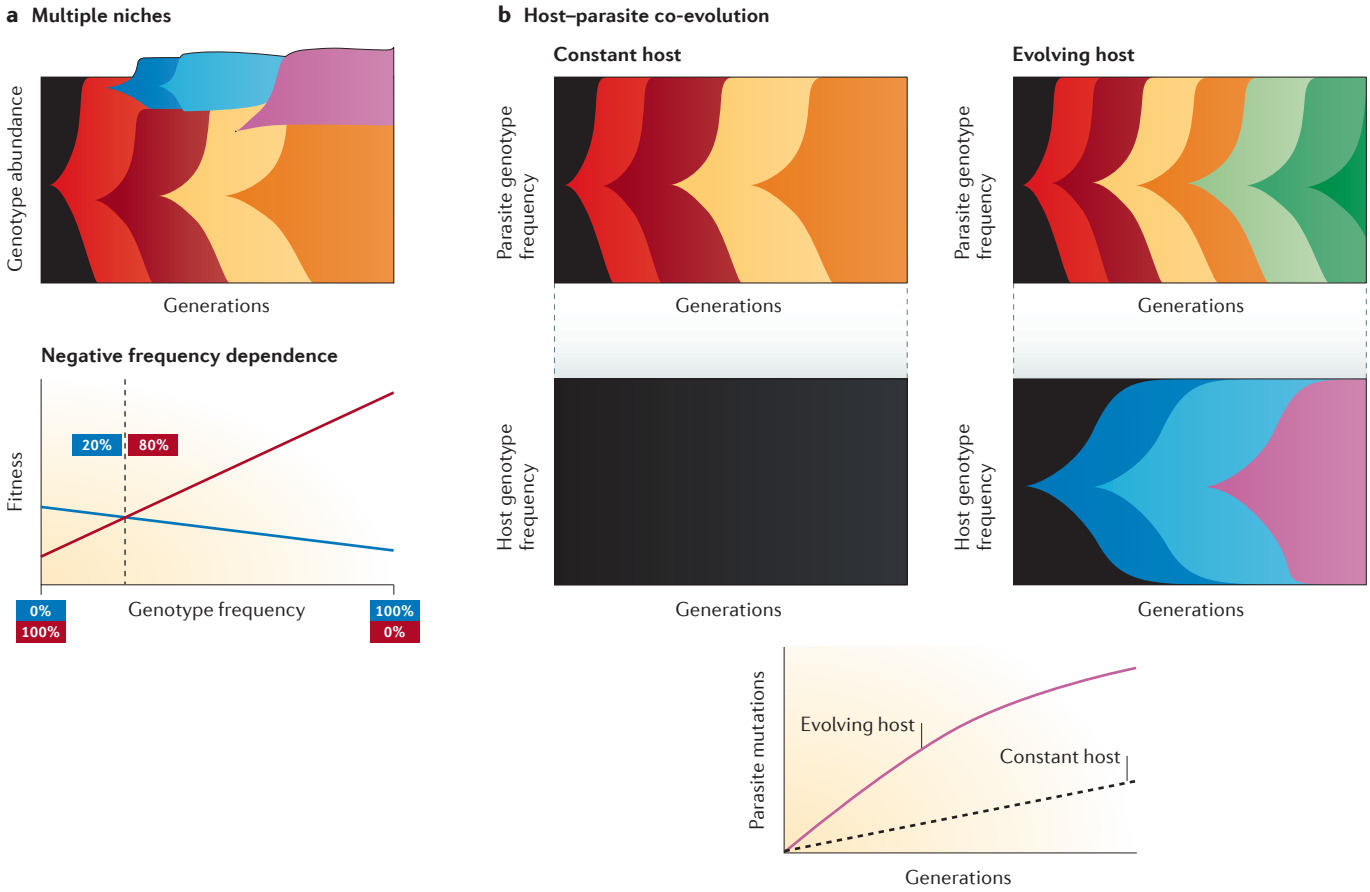


Figure 4 | **Ecological and co-evolutionary dynamics.** Examples of genetic diversification and dynamics that are driven by ecological interactions within and between species are shown. **a** | Multiple niches stabilize genetic diversity that evolves within one species. A new lineage colonizes an open niche and increases the total population size, as shown here when the red ecotype using the primary niche gives rise to the blue ecotype that expands into an open niche. Negative frequency dependence in the fitness of these ecotypes allows their coexistence. In some cases, one lineage may be a source population that can evolve to recolonize the other niche and displace the lineage that previously occupied that niche, as shown here when the yellow ecotype from the primary niche gives rise to the pink ecotype that invades the second niche, displacing the previous occupant. **b** | In host–parasite co-evolution experiments, either hosts and parasites can be allowed to co-evolve over time or one partner — in this case, the host — can be kept unchanged by continually replenishing its population from a non-evolving stock. Genetic and phenotypic evolution typically occurs at higher rates when the partners co-evolve in response to one another — a phenomenon that is often referred to as 'Red Queen' dynamics.

Co-evolution. Theory predicts, and comparative studies support, 'Red Queen' dynamics during host–parasite co-evolution, in which the rate of genome evolution is accelerated, especially in the genes that encode parasite-invasion or host-protection factors. Studies of *P. fluorescens* and its phage Φ2 have been used to examine the genome-wide dynamics of allele frequencies in replicate phage populations[70]. The experimental set-up allowed the investigation of an interesting contrast: in one treatment the host co-evolved with the parasite, whereas in another the parasite evolved while the host genotype was kept constant by repeatedly reviving the ancestral strain from a frozen stock (FIG. 4b). As predicted, the phage genome evolved faster in the co-evolution treatment than when the host was not allowed to co-evolve. By continually providing new opportunities for adaptive mutations of large effect to arise, co-evolution also leads to second-order selection for hypermutators in this system[71].

Another study of host–pathogen co-evolution, using *E. coli* and a virulent (that is, exclusively lytic) variant of phage λ, integrates many of the concepts discussed above[72]. In some replicate populations, the phage evolved the ability to infect hosts through a new cell-surface receptor. This innovation was contingent on the bacterial population evolving along a certain mutational pathway — one that initially reduced, but did not eliminate, expression of the original receptor; it was also important that later host mutations did not result in the loss of the channel that the phage uses to cross the inner membrane of the cell. Moreover, the innovation in the phage required several prior mutations in the gene that encodes the tail fibre of the phage; these prior mutations apparently spread because they improved the ability of the

phage to infect through the original receptor. In short, the complex interplay between ecological interactions and genetic contingency determined the evolutionary outcomes in these co-evolving populations.

Sexual reproduction

Whole-genome studies of the genetic dynamics of adaptive evolution in multicellular animals and plants face different challenges compared with those in microorganisms. Most multicellular organisms can or must sexually reproduce and are diploid, so they have two different copies of each chromosome. Recombination, usually during meiosis, in which DNA sequences are exchanged between homologous chromosomes, breaks the linkage between mutations and their genetic background and produces new combinations of alleles. Many classical population genetics models assume recombination, and certain inferential procedures (for example, distinguishing beneficial driver mutations from linked hitchhikers) may prove to be easier in sexual systems.

In populations in which adaptation is driven by new beneficial mutations, recombination that brings those mutations together into the same genome may outpace the *de novo* appearance of successive beneficial mutations in any one lineage[73]. Perhaps at least partly for this reason, even bacteria have mechanisms — including conjugation, transduction and transformation — that allow parasexual recombination of alleles between lineages. In experimental *E. coli* populations in which new beneficial mutations drive adaptation, adding recombination has been shown to alleviate clonal interference and to accelerate adaptation under some circumstances[74] (FIG. 2d). However, most experiments with multicellular organisms do not (and often cannot) begin with a clonal population and then wait for new beneficial mutations to arise. Instead, they usually begin with substantial genetic diversity, such as the diversity that typically exists as standing variation in natural populations. Under these conditions, the initial genotypic diversity and the new types that are generated by recombining alleles from across the genome are the primary sources of the genetic variation available for adaptation, at least in the short term (FIG. 2e).

The effects of these issues on the tempo and mode of genome evolution are just beginning to be explored in experiments with sexually reproducing multicellular organisms. Whole-genome sequencing of outbred *D. melanogaster* populations that were selected for accelerated development (that is, a shorter time from egg to adult) over 600 generations found that no alleles, neither those that were initially present nor those that arose *de novo*, had swept to fixation[75]. However, various parallel changes occurred in both the distribution of allele frequencies and the levels of homozygosity across each chromosome, which indicates that selection had repeatedly enriched certain variants. These 'soft' sweeps of existing alleles may have been incomplete because the experiment was too short or, alternatively, because different allelic combinations resulted in similar levels of phenotypic improvement. By analogy to clonal interference, the second possibility suggests a sort of 'sexual interference', in which alternative allelic combinations that produce similar benefits impede any given sweep to fixation and thereby maintain genetic diversity for longer than would otherwise be expected.

In another study, inbred lines of *D. melanogaster* were pooled, and their offspring evolved in increasingly hypoxic environments for >200 generations[76]. Individual flies from end-point populations could survive at low oxygen levels that were lethal to their ancestors. Genomic sequencing of the evolved populations found numerous apparently complete fixations of alleles, as indicated by the depletion of genetic diversity in some chromosomal regions. There are several potential explanations for the difference between this study and the one on accelerated development. One possibility is that the stronger selection for survival may have caused more extreme population bottlenecks. Another possibility is that hypoxia tolerance may involve fewer genetic loci than developmental time. A third possible explanation is that this experiment began with a set of inbred lines, whereas the experiment on developmental time used outbred lines that presumably harboured much more initial genetic diversity.

In an even longer experiment with *D. melanogaster*, flies were propagated for >50 years in complete darkness, and they have also been studied by whole-genome sequencing[77]. However, it is difficult to draw conclusions from these data because the experiment suffered an extinction of control lineages and because historical DNA samples are not available. In future studies, the planned preservation of time series of samples for genomic analysis should provide additional insights into the tempo and mode of genetic change in both animal and plant populations.

Perspectives

This Review ends by briefly commenting on other systems in which whole-genome sequencing is likely to be applied to understand evolutionary dynamics in the near future. First, more genetics, systems biology and synthetic biology studies may be unwitting evolution experiments than is commonly appreciated. In microbiology, there is a growing realization that strains that were previously believed to be isogenic are not — additional mutations beyond those that were deliberately introduced and studied have accumulated over their history[78,79]. Furthermore, one often introduces single genetic changes, or defined combinations of changes, in a reference genome in these types of studies. This genetic manipulation may be accomplished in various ways: by spontaneous mutation, by the use of a mutagen, by some means of genetic exchange, by genome editing technologies or by some combination of these strategies. These manipulations may, either occasionally or typically, cause unintended mutations in addition to the desired changes[80–82]. These considerations will also apply to the synthesis and large-scale editing of genomes[83,84], in which mutations may occur and perhaps even be favoured during these iterative processes. Such secondary mutations may need to be removed to accurately infer the effects of the intended genetic manipulations.

At the same time, it will be very interesting to see what similarities and differences emerge between the dynamics of genome evolution in laboratory experiments and in natural settings, including medically relevant settings, such as during microbial infections and tumour progression. For example, one can sequence genomes of bacteria that are sampled at multiple points over the course of chronic infections, including samples stored in the past. This approach has recently been applied to *Pseudomonas aeruginosa* that was sampled over the course of multiple decades as the bacteria evolved in the airways of individuals with cystic fibrosis[85]. It has also been used to identify adaptive mutations that arose during a local outbreak of *Burkholderia dolosa* among people with cystic fibrosis[86]. To do so, whole-genome sequences were first used to reconstruct the transmission history between host individuals, and analyses were then carried out to identify genes in the pathogen that underwent parallel changes that implied adaptation to the host environment[86]. The identified genes include some that are related to therapeutic interventions (for example, antibiotic resistance genes) and host immune responses (for example, genes encoding cell-surface antigens), as well as other genes that were not previously known to be important for these infections. Elucidating genome dynamics has similarly proved to be crucial for understanding many observations regarding the progression of neoplastic tumours[87]. These types of studies will undoubtedly lead to important advances in identifying specific genes and mutations that contribute to disease and resistance to treatment. Future studies might also reveal the extent to which ecological interactions and differences in evolvability in these genetically diverse cell populations affect disease outcomes.

1. Garland, T. & Rose, M. R. (eds) *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments* (Univ. of California Press, 2009).
2. Kawecki, T. J. *et al.* Experimental evolution. *Trends Ecol. Evol.* **27**, 547–560 (2012).
3. Hartl, D. L. & Clark, A. G. *Principles of Population Genetics* (Sinauer Associates, Inc., 2007).
4. Mardis, E. R. Next-generation DNA sequencing methods. *Annu. Rev. Genom. Hum. Genet.* **9**, 387–402 (2008).
5. Schadt, E. E., Turner, S. & Kasarskis, A. A window into third-generation sequencing. *Hum. Mol. Genet.* **19**, R227–R240 (2010).
6. Halligan, D. L. & Keightley, P. D. Spontaneous mutation accumulation studies in evolutionary genetics. *Annu. Rev. Ecol. Evol. Syst.* **40**, 151–172 (2009).
7. Lynch, M. *et al.* A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl Acad. Sci. USA* **105**, 9272–9277 (2008).
8. Ossowski, S. *et al.* The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* **327**, 92–94 (2010).
9. Keightley, P. D. *et al.* Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Res.* **19**, 1195–1201 (2009).
10. Denver, D. R. *et al.* A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proc. Natl Acad. Sci. USA* **106**, 16310–16314 (2009).
11. Lind, P. A. & Andersson, D. I. Whole-genome mutational biases in bacteria. *Proc. Natl Acad. Sci. USA* **105**, 17878–17883 (2008).
12. Lee, H., Popodi, E., Tang, H. & Foster, P. L. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc. Natl Acad. Sci. USA* **109**, E2774–E2783 (2012).
13. Sung, W., Ackerman, M. S., Miller, S. F., Doak, T. G. & Lynch, M. Drift-barrier hypothesis and mutation-rate evolution. *Proc. Natl Acad. Sci. USA* **109**, 18488–18492 (2012).
14. Drake, J. W. Spontaneous mutation. *Annu. Rev. Genet.* **25**, 125–146 (1991).
15. Lynch, M. *The Origins of Genome Architecture* (Sinauer Associates, Inc., 2007).
16. Friedberg, E. C. *et al. DNA Repair and Mutagenesis* (American Society for Microbiology Press, 2006).
17. Bull, J. J., Badgett, M. R., Rokyta, D. & Molineux, I. J. Experimental evolution yields hundreds of mutations in a functional viral genome. *J. Mol. Evol.* **57**, 241–248 (2003).
18. Domingo-Calap, P., Cuevas, J. M. & Sanjuán, R. The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genet.* **5**, e1000742 (2009).
19. Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. Rates of spontaneous mutation. *Genetics* **148**, 1667–1686 (1998).
20. Springman, R., Keller, T., Molineux, I. J. & Bull, J. J. Evolution at a high imposed mutation rate: adaptation obscures the load in phage T7. *Genetics* **184**, 221–232 (2010).
21. Ames, B. N., Durston, W. E., Yamasaki, E. & Lee, F. D. Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. *Proc. Natl Acad. Sci. USA* **70**, 2281–2285 (1973).
22. Savageau, M. A. *Escherichia coli* habitats, cell types, and molecular mechanisms of gene control. *Am. Nat.* **122**, 732–744 (1983).
23. Orr, H. A. Fitness and its role in evolutionary genetics. *Nature Rev. Genet.* **10**, 531–539 (2009).
24. Atwood, K. C., Schneider, L. K. & Ryan, F. J. Periodic selection in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **37**, 146–155 (1951).
This study is a classic early demonstration of adaptive evolution in experimental populations of bacteria.
25. Fogle, C. A., Nagle, J. L. & Desai, M. M. Clonal interference, multiple mutations and adaptation in large asexual populations. *Genetics* **180**, 2163–2173 (2008).
26. Lang, G. I. *et al.* Pervasive genetic hitchhiking and clonal interference in 40 evolving yeast populations. *Nature* **500**, 571–574 (2013).
This paper presents the most detailed analysis so far of the dynamics of mutations in asexual populations, including the effects of clonal interference, by metagenomic sequencing.
27. Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
This paper describes the first application of whole-genome sequencing to the LTEE and includes discussions of genetic parallelism, changes in mutation rates and evolution in the optimization regime.
28. Wagner, A. *The Origins of Evolutionary Innovations* (Oxford Univ. Press, 2011).
29. Velicer, G. J. *et al.* Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. *Proc. Natl Acad. Sci. USA* **103**, 8107–8112 (2006).
This study is the first to use whole-genome sequencing to discover the genetic basis of an innovative change in the lifestyle of an organism that occurred during an evolution experiment.
30. Yu, Y.-T. N., Yuan, X. & Velicer, G. J. Adaptive evolution of an sRNA that controls *Myxococcus* development. *Science* **328**, 993 (2010).
31. Marchetti, M. *et al.* Experimental evolution of a plant pathogen into a legume symbiont. *PLoS Biol.* **8**, e1000280 (2010).
32. Ratcliff, W. C., Denison, R. F., Borrello, M. & Travisano, M. Experimental evolution of multicellularity. *Proc. Natl Acad. Sci. USA* **109**, 1595–1600 (2012).
33. Mortlock, R. P. (ed.) *Microorganisms as Model Systems for Studying Evolution* (Plenum Press, 1984).
34. Blount, Z. D., Borland, C. Z. & Lenski, R. E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **105**, 7899–7906 (2008).
35. Blount, Z. D., Barrick, J. E., Davidson, C. J. & Lenski, R. E. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* **489**, 513–518 (2012).
This paper describes the genetic basis of the evolution of citrate use in the LTEE, including the potentiation, actualization and refinement stages of this innovation.
36. Conrad, T. M., Lewis, N. E. & Palsson, B. O. Microbial laboratory evolution in the era of genome-scale science. *Mol. Syst. Biol.* **7**, 509 (2011).
37. Minty, J. J. *et al.* Evolution combined with genomic study elucidates genetic bases of isobutanol tolerance in *Escherichia coli*. *Microb. Cell Fact.* **10**, 18 (2011).
38. Blaby, I. K. *et al.* Experimental evolution of a facultative thermophile from a mesophilic ancestor. *Appl. Environ. Microbiol.* **78**, 144–155 (2012).
39. Tenaillon, O. *et al.* The molecular diversity of adaptive convergence. *Science* **335**, 457–461 (2012).
This study uses whole-genome sequencing of a large number of independently evolved populations to examine the diversity of alternative genetic pathways that lead to improved fitness at high temperature.
40. Harris, D. R. *et al.* Directed evolution of ionizing radiation resistance in *Escherichia coli*. *J. Bacteriol.* **191**, 5240–5252 (2009).
41. Zhang, Q. *et al.* Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Science* **333**, 1764–1767 (2011).
This study shows that migration between populations living in environments with different selection strengths can speed up adaptation.
42. Pepin, K. M. & Wichman, H. A. Experimental evolution and genome sequencing reveal variation in levels of clonal interference in large populations of bacteriophage ΦX174. *BMC Evol. Biol.* **8**, 85 (2008).
43. Sniegowski, P. D., Gerrish, P. J. & Lenski, R. E. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387**, 703–705 (1997).
44. Brown, C. T. *et al.* Whole genome sequencing and phenotypic analysis of mutations found in *Bacillus subtilis* following evolution under relaxed selection for sporulation. *Appl. Env. Microbiol.* **77**, 6867–6877 (2011).
45. Aguilar, C. *et al.* Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system. *BMC Genomics* **13**, 385 (2012).
46. Mao, E. F., Lane, L., Lee, J. & Miller, J. H. Proliferation of mutators in a cell population. *J. Bact.* **179**, 417–422 (1997).
47. Lenski, R. E. Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*. *Plant Breed. Rev.* **24**, 225–265 (2004).
48. Desai, M. M. & Fisher, D. S. The balance between mutators and nonmutators in asexual populations. *Genetics* **188**, 997–1014 (2011).
49. Sniegowski, P. D., Gerrish, P. J., Johnson, T. & Shaver, A. The evolution of mutation rates: separating causes from consequences. *BioEssays* **22**, 1057–1066 (2000).

50. Wielgoss, S. *et al.* Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc. Natl Acad. Sci. USA* **110**, 222–227 (2013).
51. McDonald, M. J., Hsieh, Y.-Y., Yu, Y.-H., Chang, S.-L. & Leu, J.-Y. The evolution of low mutation rates in experimental mutator populations of *Saccharomyces cerevisiae. Curr. Biol.* **22**, 1235–1240 (2012).
52. Denamur, E. *et al.* Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell* **103**, 711–721 (2000).
53. Woods, R. J. *et al.* Second-order selection for evolvability in a large *Escherichia coli* population. *Science* **331**, 1433–1436 (2011).
This study uses replay experiments to demonstrate that antagonistic epistasis can lead some genotypes towards an adaptive cul-de-sac and eventual extinction.
54. Darwin, C. *The Origin of Species by Means of Natural Selection* (John Murray, 1859).
55. Laland, K. N., Odling-Smee, F. J. & Feldman, M. W. Evolutionary consequences of niche construction and their implications for ecology. *Proc. Natl Acad. Sci. USA* **96**, 10242–10247 (1999).
56. Rozen, D. E., Schneider, D. & Lenski, R. E. Long-term experimental evolution in *Escherichia coli.* XIII. Phylogenetic history of a balanced polymorphism. *J. Mol. Evol.* **61**, 171–180 (2005).
57. Le Gac, M., Plucain, J., Hindré, T., Lenski, R. E. & Schneider, D. Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli. Proc. Natl Acad. Sci. USA* **109**, 9487–9492 (2012).
58. Rozen, D. E., Philippe, N., Arjan De Visser, J., Lenski, R. E. & Schneider, D. Death and cannibalism in a seasonal environment facilitate bacterial coexistence. *Ecol. Lett.* **12**, 34–44 (2009).
59. Barrick, J. E. & Lenski, R. E. Genome-wide mutational diversity in an evolving population of *Escherichia coli. Cold Spring Harbor Symp. Quant. Biol.* **74**, 119–129 (2009).
60. Zambrano, M. M., Siegele, D. A., Almirón, M., Tormo, A. & Kolter, R. Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. *Science* **259**, 1757–1760 (1993).
61. Finkel, S. E. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nature Rev. Microbiol.* **4**, 113–120 (2006).
62. Treves, D. S., Manning, S. & Adams, J. Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli. Mol. Biol. Evol.* **15**, 789–797 (1998).
63. Maharjan, R., Seeto, S., Notley-McRobb, L. & Ferenci, T. Clonal adaptive radiation in a constant environment. *Science* **313**, 514–517 (2006).
64. Maharjan, R. P. *et al.* The multiplicity of divergence mechanisms in a single evolving population. *Genome Biol.* **13**, R41 (2012).
This paper uses whole-genome sequencing to uncover extensive genetic diversity in populations that are evolving in chemostats.
65. Saxer, G., Doebeli, M. & Travisano, M. The repeatability of adaptive radiation during long-term experimental evolution of *Escherichia coli* in a multiple nutrient environment. *PLoS ONE* **5**, e14184 (2010).
66. Herron, M. D. & Doebeli, M. Parallel evolutionary dynamics of adaptive diversification in *Escherichia coli. PLoS Biol.* **11**, e1001490 (2013).
This study examines whole-population genetic dynamics in a system that diversified into two ecotypes which coexisted owing to different strategies for nutrient use.
67. Korona, R., Nakatsu, C. H., Forney, L. J. & Lenski, R. E. Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. *Proc. Natl Acad. Sci. USA* **91**, 9037–9041 (1994).
68. Rainey, P. B. & Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* **394**, 69–72 (1998).
69. Traverse, C. C., Mayo-Smith, L. M., Poltak, S. R. & Cooper, V. S. Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proc. Natl Acad. Sci. USA* **110**, E250–E259 (2013).
This study characterizes the mutational pathways by which ecotypes from a primary niche evolved to displace ecotypes that were established in other niches.
70. Paterson, S. *et al.* Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278 (2010).
This study uses whole-genome sequencing of a bacteriophage to observe Red Queen dynamics.
71. Pal, C., Maciá, M. D., Oliver, A., Schachar, I. & Buckling, A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature* **450**, 1079–1081 (2007).
72. Meyer, J. R. *et al.* Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **335**, 428–432 (2012).
This paper characterizes multistep genetic pathways that lead to an innovative host–receptor shift in bacteriophage populations and shows that the shift is contingent on the evolutionary trajectory of the bacterial host populations.
73. Muller, H. Genetic aspects of sex. *Am. Nat.* **66**, 118–138 (1932).
74. Cooper, T. F. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli. PLoS Biol.* **5**, e225 (2007).
This study uses knowledge of the genetic targets of selection to show that plasmid-mediated recombination accelerates adaptive evolution by overcoming clonal interference.
75. Burke, M. K. *et al.* Genome-wide analysis of a long-term evolution experiment with *Drosophila. Nature* **467**, 587–590 (2010).
This paper presents one of the first analyses based on whole-genome sequencing to characterize experimentally evolved populations of a sexual multicellular organism.
76. Zhou, D. *et al.* Experimental selection of hypoxia-tolerant *Drosophila melanogaster. Proc. Natl Acad. Sci. USA* **108**, 2349–2354 (2011).
This paper uses extensive sequencing to identify targets of selection in another set of experimentally evolved populations of a sexual multicellular organism.
77. Izutsu, M. *et al.* Genome features of "Dark-fly", a *Drosophila* line reared long-term in a dark environment. *PLoS ONE* **7**, e33288 (2012).
78. Ferenci, T. *et al.* Genomic sequencing reveals regulatory mutations and recombinational events in the widely used MC4100 lineage of *Escherichia coli* K-12. *J. Bacteriol.* **191**, 4025–4029 (2009).
79. Nahku, R. *et al.* Stock culture heterogeneity rather than new mutational variation complicates short-term cell physiology studies of *Escherichia coli* K-12 MG1655 in continuous culture. *Microbiology* **157**, 2604–2610 (2011).
80. Hobman, J. L. *et al.* Comparative genomic hybridization detects secondary chromosomal deletions in *Escherichia coli* K-12 MG1655 mutants and highlights instability in the *flhDC* region. *J. Bacteriol.* **189**, 8786–8792 (2007).
81. Pósfai, G. *et al.* Emergent properties of reduced-genome *Escherichia coli. Science* **312**, 1044–1046 (2006).
82. Studier, F. W., Daegelen, P., Lenski, R. E., Maslov, S. & Kim, J. F. Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3) and comparison of the *E. coli* B and K-12 genomes. *J. Mol. Biol.* **394**, 653–680 (2009).
83. Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52–56 (2010).
84. Isaacs, F. J. *et al.* Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
85. Yang, L. *et al.* Evolutionary dynamics of bacteria in a human host environment. *Proc. Natl Acad. Sci. USA* **108**, 7481–7486 (2011).
86. Lieberman, T. D. *et al.* Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature Genet.* **43**, 1275–1280 (2011).
This study uses whole-genome sequencing to investigate adaptation of a bacterial pathogen during a multi-decade outbreak in a human population; although this is not based on an experiment, it relies on inferential approaches similar to those used in experimental evolution.
87. Sprouffske, K., Merlo, L. M. F., Gerrish, P. J., Maley, C. C. & Sniegowski, P. D. Cancer in light of experimental evolution. *Curr. Biol.* **22**, R762–R771 (2012).
88. Loewe, L. & Hill, W. G. The population genetics of mutations: good, bad and indifferent. *Phil. Trans. R. Soc.* **365**, 1153–1167 (2010).
89. Eyre-Walker, A. & Keightley, P. D. The distribution of fitness effects of new mutations. *Nature Rev. Genet.* **8**, 610–618 (2007).
90. De Visser, J. A. G. M., Cooper, T. F. & Elena, S. F. The causes of epistasis. *Phil. Trans. R. Soc.* **278**, 3617–3624 (2011).
91. Kondrashov, F. A. & Kondrashov, A. S. Measurements of spontaneous rates of mutations in the recent past and the near future. *Phil. Trans. R. Soc.* **365**, 1169–1176 (2010).
92. Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E. & Cooper, T. F. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* **332**, 1193–1196 (2011).
93. Chou, H.-H., Chiu, H.-C., Delaney, N. F., Segrè, D. & Marx, C. J. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* **332**, 1190–1192 (2011).
94. Kvitek, D. J. & Sherlock, G. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet.* **7**, e1002056 (2011).
95. Philippe, N., Crozat, E., Lenski, R. E. & Schneider, D. Evolution of global regulatory networks during a long-term experiment with *Escherichia coli. BioEssays* **29**, 846–860 (2007).
96. Wagner, A. Neutralism and selectionism: a network-based reconciliation. *Nature Rev. Genet.* **9**, 965–974 (2008).
97. Heard, S. & Hauser, D. Key evolutionary innovations and their ecological mechanisms. *Histor. Biol.* **10**, 151–173 (1995).
98. Lenski, R. E. & Travisano, M. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc. Natl Acad. Sci. USA* **91**, 6808–6814 (1994).
99. Carroll, S. M. & Marx, C. J. Evolution after introduction of a novel metabolic pathway consistently leads to restoration of wild-type physiology. *PLoS Genet.* **9**, e1003427 (2013).
100. Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. Different trajectories of parallel evolution during viral adaptation. *Science* **285**, 422–424 (1999).
This paper is one of the first to use whole-genome sequencing to examine the nature of genetic changes during an adaptive evolution experiment.
101. Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A. & Lenski, R. E. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli. Proc. Natl Acad. Sci. USA* **103**, 9107–9112 (2006).
102. Bull, J. J. *et al.* Exceptional convergent evolution in a virus. *Genetics* **147**, 1497–1507 (1997).
103. Crozat, E. *et al.* Parallel genetic and phenotypic evolution of DNA superhelicity in experimental populations of *Escherichia coli. Mol. Biol. Evol.* **27**, 2113–2128 (2010).
104. Jerome, J. P. *et al.* Standing genetic variation in contingency loci drives the rapid adaptation of *Campylobacter jejuni* to a novel host. *PLoS ONE* **6**, e16399 (2011).
105. Lee, M.-C. & Marx, C. J. Repeated, selection-driven genome reduction of accessory genes in experimental populations. *PLoS Genet.* **8**, e1002651 (2012).
106. Cooper, V. S., Schneider, D., Blot, M. & Lenski, R. E. Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *Escherichia coli* B. *J. Bacteriol.* **183**, 2834–2841 (2001).
107. Paquin, C. E. & Adams, J. Relative fitness can decrease in evolving asexual populations of *S. cerevisiae. Nature* **306**, 368–370 (1983).

Acknowledgements
The authors thank the reviewers for their suggestions. They acknowledge support from the US National Institutes of Health (R00-GM087550 to J.E.B.), the US National Science Foundation (NSF; DEB-1019989 to R.E.L.) and the BEACON Center for the Study of Evolution in Action (NSF Cooperative Agreement DBI-0939454 to J.E.B. and R.E.L.).

Competing interests statement
The authors declare no competing interests.

FURTHER INFORMATION
LTEE: http://myxo.css.msu.edu/ecoli/
The US National Science Foundation BEACON Center for the Study of Evolution in Action: http://beacon-center.org/
ALL LINKS ARE ACTIVE IN THE ONLINE PDF