Throughout its history, evolutionary biology has relied primarily on comparative studies of living organisms, supplemented whenever possible by data from the fossil record. Such comparative studies have become increasingly powerful with the emergence of molecular data for determining phylogenies and with the increased rigour with which the comparative method has been applied. However, biologists have long been interested in observing the dynamics of evolutionary change more directly. Indeed, Charles Darwin remarked in 1859 that “in looking for the gradations by which an organ in any species has been perfected, we ought to look exclusively to its lineal ancestors; but this is scarcely ever possible, and we are forced in each case to look to species of the same group, that is to the collateral descendants from the same original parent-form”. So, Darwin used a comparative approach by necessity, and admitted how valuable it would be to observe the actual processes of descent with modification and adaptation by natural selection.

Since Darwin’s day, many examples of evolution in action have been studied in nature, ranging from the emergence of antibiotic resistance in bacteria to changes in the beak morphology of Darwin’s finches. Beyond simply observing evolution in action, some biologists sought to carry out experiments that ran for many generations, with controls and replication, to test hypotheses about the evolutionary process. Beginning with T. H. Morgan, and for several subsequent decades, population geneticists that studied fruitflies were the main practitioners of experimental evolution. A few groups used other organisms, including bacteria, in evolution experiments but, except in the fruitfly school, this approach did not take hold. In the case of microorganisms, a rift developed as most microbiologists pursued ever more molecular approaches but largely ignored evolution.

The situation has changed in recent years, with many groups now carrying out evolution experiments on diverse organisms including plants, vertebrates and, especially, microorganisms. On one side, there was a recognition by ecologists, including those that were interested in evolutionary ecology, of the need for rigorous experiments to test hypotheses. On the other side, many microbiologists realized the value of an evolutionary perspective, stimulated by the discovery of the Archaea and the rapidly increasing genomic data. These new attitudes led to the realization that microbes offer powerful systems for experimental evolution (Box 1).

There are several other reasons for which microbial evolution experiments have received increasing attention. By virtue of the control that can be exerted over many variables in a laboratory setting, and the power that is afforded by the direct observation of any
REVIEWS

Box 1 | Advantages of microorganisms for evolution experiments

Microorganisms that have been used in evolution experiments include many bacteria and viruses, as well as unicellular algae and fungi. These organisms are well suited for such experiments for many practical reasons:

• They are easy to propagate and enumerate.
• They reproduce quickly, which allows experiments to run for many generations.
• They allow large populations in small spaces, which facilitates experimental replication.
• They can be stored in suspended animation and later revived, which allows the direct comparison of ancestral and evolved types.
• Many microbes reproduce asexually and the resulting clonality enhances the precision of experimental replication.
• Asexuality also maintains linkage between a genetic marker and the genomic background into which it is placed, which facilitates fitness measurements (BOX 2).
• It is easy to manipulate environmental variables, such as resources, as well as the genetic composition of founding populations.
• There are abundant molecular and genomic data for many species, as well as techniques for their precise genetic analysis and manipulation.

Fitness
The average reproductive success of a genotype in a particular environment. Often expressed relative to another genotype, such as the ancestor in evolution experiments.

Dynamic process, many questions about evolution can be probed with greater rigour than would otherwise be possible. For example, the reproducibility of evolutionary outcomes can be studied in microbial populations that are founded by the same ancestor and placed in identical environments. Although a ‘natural experiment’ that involves the colonization of several neighbouring islands by the same insect species could provide insights into this question, it would be difficult to exclude the possibility that subtle environmental differences promoted divergence, or, alternatively, that mutations that were already present in the source population contributed to parallel responses.

Also, many microbes are of great importance to humans, not only as pathogens but for numerous essential ecosystem services. Therefore, it is crucial to understand the mechanisms and dynamics of microbial evolution. The most important difference between microbial and ‘macrobial’ evolution is probably not organisational size, or even the speed of generations, but is the fact that most microbes can reproduce asexually whereas most plants and animals are sexual. All microorganisms must not be viewed as alike. However, there are grounds for optimism that some generalizations (or at least strong tendencies) do exist because experiments on such different types of microbe as viruses, bacteria and yeast often support broadly similar conclusions. Moreover, many experiments with microbes are designed to test general hypotheses that are derived from evolutionary theory that has been developed for other organisms.

One final reason for the growing interest in microbial evolution experiments is the satisfaction that often comes from observing ‘real time’ a process that is usually inferred indirectly. Indeed, the field of evolutionary biology has long been hounded by sceptics who question, for example, whether evolution can produce adaptation if it depends on random mutations, most of which are deleterious. Of course, such ill-founded criticisms can be dismissed by showing their logical flaws with the appropriate mathematics, but it is also nice to be able to point to experiments that confirm the basic underpinnings of a scientific field.

Most experiments in microbial evolution are conceptually simple. Populations are established (often from single clones), then propagated in a controlled and reproducible environment for many generations. A sample of the ancestral population is stored indefinitely (for example, frozen at –80 °C), as are samples from various time points in the experiment. After a population has been propagated for some time, the ancestral and derived genotypes can be compared with respect to any genetic or phenotypic properties of interest, which provides information on the dynamics of the evolutionary process and the extent of evolutionary change. Importantly, adaptation can be quantified by measuring changes in fitness in the experimental environment, in which fitness reflects the propensity to leave descendants. With microorganisms, fitness can be measured using head-to-head competition between, for example, an evolutionarily derived line and its ancestor that is genetically marked. In brief, the population growth rates that are achieved by each type as they compete for a pool of resources are measured. Different markers can be used to distinguish competitors, such as those that produce visible reactions with dyes, resistance to antibiotics or diagnostic PCR fragments. Provided that the organisms are asexual, as in most microbial evolution experiments, the marker serves as a proxy for the entire genome. By using control assays to measure any effect of the marker, and by replicating competitions, it is possible to reliably quantify evolutionary changes in fitness.

Of course, relative fitness depends not only on the genotypes but also on the environment in which it is measured. As discussed later, it is possible to test the specificity of adaptation that occurred in an evolution experiment by measuring fitness in different environments. Unless otherwise specified, however, it should be understood that fitness is measured under conditions that are similar or identical to those that prevailed during an evolution experiment.

In this review of evolution experiments with microorganisms, we examine the dynamics of evolutionary adaptation, the genetic bases of adaptation, tradeoffs and the environmental specificity of adaptation, the origin and consequences of mutators, and the process of drift decay in small populations. Experiments that address complex interactions, including compensatory adaptation, the maintenance of genetic diversity, social conflict and cooperation, the effects of sexual recombination and host–parasite interactions, will be discussed in a second review, to be published in a future issue of Nature Reviews Genetics.

Even with two reviews, there are topics that we cannot cover in the available space. In organizing our subject, we chose to focus on evolution experiments that are open ended and long-term in approach. We do not review selection experiments that targeted specific, often new, metabolic functions, even though this work is of great interest. Reviews of these targeted selection experiments can be found elsewhere, as can reviews of the
Box 2 | Measuring fitness

The fitness of an evolved type is generally expressed relative to its ancestor. Relative fitness is measured by allowing the ancestral and evolved types to compete with one another. Unless otherwise specified, the competition environment is the same as that used for the experimental evolution. The following description presents the protocol used in the long-term serial-transfer experiment with Escherichia coli, but similar procedures are used in experiments with many microorganisms.

The two competitors are grown separately in the competition environment to ensure that they are comparably acclimated to the test conditions. They are then mixed (usually at a 1:1 ratio) and diluted (100-fold in this case) in the competition environment. Initial densities at timepoint \( t = 0 \) are estimated by diluting and spreading the cells on an indicator agar that distinguishes the evolved and ancestral types by colony colour, which differs owing to an engineered marker that is selectively neutral. In this case, red and white colonies correspond to Ara- and Ara+ phenotypes, respectively. After one day (\( t = 1 \)) (corresponding to the serial-transfer cycle in the evolution experiment), final densities are estimated by plating cells, as before, on the indicator agar. The growth rate of each competitor is calculated as the natural logarithm of the ratio of its final density to its initial density (adjusted for dilution during plating). Relative fitness is then defined simply as the ratio of the realized growth rates of the evolved and ancestral types.

The fitness of an evolved type is generally expressed relative to its ancestor. Relative fitness is measured by allowing the ancestral and evolved types to compete with one another. Unless otherwise specified, the competition environment is the same as that used for the experimental evolution. The following description presents the protocol used in the long-term serial-transfer experiment with Escherichia coli, but similar procedures are used in experiments with many microorganisms.

The two competitors are grown separately in the competition environment to ensure that they are comparably acclimated to the test conditions. They are then mixed (usually at a 1:1 ratio) and diluted (100-fold in this case) in the competition environment. Initial densities at timepoint \( t = 0 \) are estimated by diluting and spreading the cells on an indicator agar that distinguishes the evolved and ancestral types by colony colour, which differs owing to an engineered marker that is selectively neutral. In this case, red and white colonies correspond to Ara- and Ara+ phenotypes, respectively. After one day (\( t = 1 \)) (corresponding to the serial-transfer cycle in the evolution experiment), final densities are estimated by plating cells, as before, on the indicator agar. The growth rate of each competitor is calculated as the natural logarithm of the ratio of its final density to its initial density (adjusted for dilution during plating). Relative fitness is then defined simply as the ratio of the realized growth rates of the evolved and ancestral types.

Random drift

The change in frequency of genotypes in a population that is caused by chance differences in survival and reproduction, as opposed to consistent differences in their fitness.

Dynamics of evolutionary adaptation

Adaptation by natural selection occurs through the spread and substitution of mutations that improve the performance of an organism and its reproductive success in its environment. An important focus of evolution experiments using microorganisms has been to investigate the dynamics of this process. Among the questions of general interest are whether genetic adaptation can continue indefinitely even in a constant environment, the magnitude of the contributions of individual mutations to fitness improvement, and the overall reproducibility of evolutionary changes.

One feature that is seen in several experiments, with both bacteria and viruses, is that fitness gains are initially rapid but tend to decelerate over time. Such dynamics indicate that populations, after being placed in a new environment, are evolving from a region of low fitness towards an adaptive peak or plateau (Fig. 1). For example, in a long-term experiment with 12 Escherichia coli populations, the average fitness gain in the first 5,000 generations was approximately tenfold greater than that between 15,000 and 20,000 generations.

Even so, there was some significant improvement over the last interval, which indicated that the populations still had not reached their maximum fitness.

At first glance, it might seem surprising that the potential for genetic adaptation was not exhausted after thousands of generations in a constant environment. However, several factors contribute to continued adaptation. First, the amount of time that is required for a beneficial mutation to increase in frequency from a single individual to most of the population is inversely proportional to its advantage. Given the population size in the long-term experiment with E. coli, a new mutation that has a 10% advantage would take approximately 250 generations to become the majority status. By comparison, a mutation with only a 0.1% advantage would require 25,000 generations to reach that frequency; so, there had not been enough time for such small beneficial mutations to have been substituted in that experiment. Second, many beneficial mutations are lost by random drift while they are rare. The probability that a beneficial mutation survives extinction by drift is about twice its selective advantage. A mutation with a 10% advantage requires on average five ‘tries’ before it is established, whereas a mutation with only a 0.1% advantage would need ~500 tries to avoid extinction by random drift. Both of these factors imply that adaptation using mutations with progressively smaller benefits can continue indefinitely, albeit ever more slowly, without depleting the supply of useful variants. Third, as discussed later, asexual populations are subject to ‘clonal interference’ that is caused by competition...
among beneficial mutations that occur in different clones. The substitution of small beneficial mutations is especially affected by this phenomenon, further delaying their spread. Finally, it is likely that there are more mutations that confer small advantages than those that provide large benefits. Hence, the supply of small beneficial mutations will not be exhausted as readily as the supply of large beneficial mutations (above and beyond the other dynamical effects).

Over the long term, fitness trajectories might seem smooth and continuous. However, if they are measured with sufficient temporal resolution, then fitness, and traits that are correlated with fitness (such as cell size in E. coli) change with a step-like dynamic. Each step probably corresponds to the spread of a beneficial mutation. The step-like aspect occurs because any new beneficial mutation must increase from a low initial frequency; during its ascendency, it has little effect on mean fitness until it is present in a substantial fraction of the population. Also, whereas sexual reproduction allows two or more beneficial mutations to be substituted simultaneously, adaptation in asexual populations occurs by sequential substitutions that appear as successive steps.

The step-like dynamics of adaptive evolution provide one way of measuring the fitness effects of mutations that are substituted in evolving populations, and several other approaches have also provided such data. These studies support some general points. Evolutionary adaptation in experimental microbial populations typically occurs through the substitution of relatively few mutations that confer large benefits, as opposed to countless mutations with small benefits. The percentage of all mutations that are beneficial is small, but is sufficient to allow adaptation given the supply of mutations even in populations that are relatively small by microbial standards. It is also clear, from both theory and experiment, that the fitness gains of substituted mutations are not representative of the effects of beneficial mutations more generally. Instead, the most extreme beneficial mutations are greatly overrepresented owing to selection, and this bias is especially strong in asexual organisms as a result of clonal interference. As a consequence, it is difficult to test the assertion that the underlying distribution of beneficial mutations has many more with small than with large effects. Indirect support for this conjecture comes from the continued, but much slower, adaptation in the long-term E. coli populations after thousands of generations. One study directly tested this hypothesis by varying the population size of the RNA virus φ6. Small populations tended to improve by more numerous, but smaller, steps than did large populations, which confirmed the preponderance of beneficial mutations with small effects in the underlying distribution.

As was briefly noted previously, an important consequence of asexuality is clonal interference. Clones that carry different beneficial mutations compete with one another and thereby interfere with each other’s spread and substitution in the population. In general, all but one lineage will be excluded by the clone with the most beneficial mutation or combination of mutations (unless several clones partition the environment). Clonal interference has several important implications. First, the probability of substitution of a given beneficial mutation should decline with increasing population size or mutation rate. Second, as population size or mutation rate increases, individual substitutions should entail larger fitness gains. Third, the rate of fitness improvement should show diminishing returns with...
an increasing supply of beneficial mutations that is caused by larger population size or higher mutation rate. Fourth, the rate of spread of a beneficial mutation should be slower than otherwise predicted from its fitness advantage. Fifth, many beneficial mutations should become transiently common but later be excluded by interfering mutations. Sixth, such transient dynamics might give rise to a ‘leapfrog’ event, in which the most common genotype at a given time is genealogically more distantly related to the immediately preceding dominant type than to an earlier dominant type. All of these effects have been reported in experiments with bacteria and viruses.

Although founded by the same clone, and evolving in identical environments, replicate populations often diverge from one another in their relative fitness, demographic components of fitness, morphological features, and performance in other environments. This divergence might indicate that they are approaching different local peaks in the adaptive landscape, especially if fitness differences in the selective environment itself are sustained indefinitely. Experiments that start with different genotypes have also been run to examine how this affects the dynamics and extent of adaptation. In an experiment to investigate the role of historical contingencies, E. coli lines that had diverged greatly in their fitness on maltose as they evolved in glucose for 2,000 generations were then placed in a maltose environment and allowed to evolve for 1,000 generations. Lines that started with the lowest fitness on maltose improved most rapidly, and all the lines tended to converge towards a similar fitness on that sugar. However, convergence does not always occur. Replicate populations that were founded by two different genotypes of virus φ6 consistently evolved to different fitness levels, which indicated that the descendants of one founder might have been ‘trapped’ in the domain of a lower fitness peak. In other words, the mutational pathways that led from one founder to the higher fitness peak might have included maladapted intermediate genotypes that would be disfavored by natural selection.

The genetic bases of evolutionary adaptation

Throughout the history of microbial genetics, most experiments have proceeded by disrupting organismal functions rather than improving them. This approach has been productive in terms of identifying the genes that encode the molecular components that allow various biochemical and physiological functions to occur. However, this focus on defective mutants does not provide much insight into how organismal function can improve. Evolution experiments, by contrast, offer opportunities to study beneficial mutations. Of course, the particular mutations that are beneficial will depend on the genomic and environmental contexts; for example, different genes and pathways might change as E. coli populations adapt to resource abundance versus scarcity. But several more general questions can be posed. What types of molecular event are involved in adaptation? Are point mutations more important than genomic rearrangements, or vice versa? What types of gene are affected? Are most beneficial mutations located in structural genes or in regulatory elements? Does the reproducibility of evolutionary outcomes at the phenotypic level usually indicate parallel mutations, or do similar phenotypic adaptations often arise from different underlying mutations?

Of course, the fact that one clone is more fit than another in a certain environment tells us nothing about the genetic basis of their difference. To achieve this understanding requires three steps: finding mutations that were substituted, manipulating ancestral and derived alleles to make clones that are isogenic except for known mutations, and measuring the fitness consequences of those mutations in the relevant environment. The second and third steps are needed because neutral and even deleterious mutations might be fixed by random drift or by hitchhiking with beneficial mutations at other loci. Hitchhiking is especially important in asexual populations, in which the entire genome acts as a single linkage group. Sometimes a strong argument can be made that some mutations are beneficial without (or before) performing the last two steps. In particular, if several lines independently substitute the same or similar mutations (and if there is no reason to suspect hypermutability at those sites), then such parallel changes provide compelling evidence that they spread by selection and, hence, were adaptive.

The first challenge is to find mutations that distinguish the evolved and ancestral genotypes. One approach is to sequence as many genes as possible; another is to focus on candidate loci that are identified by parallel phenotypic changes that have evolved, an understanding of selective factors in the environment, or both. For many viruses, it is now practical to sequence the entire genomes of several evolved lines and their ancestor. In viruses generally, and especially those with RNA genomes, mutations can accumulate quickly owing to high per-site mutation rates. An interesting observation from sequencing viral genomes that have been obtained by experimental evolution is the extent of parallel changes at the nucleotide level across replicate lines that evolved in the same environment. These parallel changes are presumably beneficial. In one study, a number of nucleotide substitutions in experimental lines of bacteriophage X174 and S13 recapitulated evolution that also occurred in nature, based on the genome sequences of these closely related viruses. In another experiment, several beneficial mutations were identified in X174, one of which was moved into a different genetic background that represented various intermediate evolutionary stages. The benefit of this mutation was reduced in the later evolutionary stages, which indicated ‘diminishing returns’ epistasis. Such a pattern of gene interaction could contribute to the general pattern of decelerating fitness gains that was noted previously. In the RNA-encoded vesicular stomatitis virus (VSV), some parallel substitutions were synonymous and others were in non-coding regions, which indicated that selection might have been at the level of RNA folding or RNA–protein interactions.
For bacteria, it is not yet practical to obtain entire genome sequences for several evolved lines and their ancestor. However, it is possible to sequence many regions at random to infer the extent of genomic changes and perhaps find mutations for further study. One study, which used the 12 long-term E. coli lines, randomly chose 36 genes and sequenced 500-bp regions in four clones from each line and their ancestor57. Several mutations were found in a few lines that evolved mutator phenotypes (see later), but no mutations were found in any of the eight lines that retained functional DNA repair throughout the 20,000-generation experiment among the 18,374 bp that were sequenced from each clone. Although this study did not find any compelling mutations for further research, the data provide an important baseline against which to compare patterns of change at candidate loci. For example, using genomic arrays to look for parallel changes in gene expression in these lines, several candidate regulatory genes were identified including spoT, which encodes a protein that controls the level of the important effector molecule ppGpp (REF. 58). Sequencing spoT found point mutations that caused amino-acid replacements in 8 of the 12 lines, which indicated much more evolution than in the baseline of random genes. By moving an evolved spoT allele into the ancestral genome and running fitness assays, it was confirmed that the mutation was beneficial in the SERIAL-TRANSFER regime that was used in the long-term evolution experiment56.

In CHEMOSTAT cultures, bacteria are faced with perpetual resource limitation. The transport systems for the limiting resource are probably targets of selection, and the genes that encode those systems are, therefore, candidate loci. Consistent with these expectations, E. coli that evolved in lactose-limited chemostats substituted mutations in ompF that improved permeability across the outer membrane of the nonspecific OmpF porin by reducing channel constriction59. In glucose-limited chemostats, E. coli evolved diverse mutations at several loci that increased glucose permeability through the Lmb porin and the binding protein-dependent transport of glucose across the inner membrane into the cell50–52. Several clones with different alleles at these loci often increased in tandem, probably indicating clonal interference. In another study, E. coli in glucose-limited chemostats, several clones coexisted through a cross-feeding interaction, in which one type secreted acetate that another used as a resource; mutations in regulatory regions upstream of acs, which encodes acetyl-CoA synthetase, were partly responsible for this interaction63,64. Although E. coli in chemostats face limiting resources, their growth rate might still be much faster than they can usually achieve in nature. By allowing E. coli that had been recently taken from nature to evolve for 280 generations in chemostats, it was shown that the bacteria tended to converge on the rapid-growth phenotype of strains with long histories in the laboratory56. Changes in the kinetic properties of the ribosomes, which increased their translational efficiency, were responsible for much of this adaptation.

Bacteria that are kept indefinitely in STATIONARY PHASE experience even more severe limitations than those in chemostats, because no new resources are provided following the initial growth in fresh medium. To survive, cells must scavenge whatever becomes available through excretion or death in an otherwise starving population. After a period of mortality, mutants emerge that can survive and grow, albeit slowly, under such conditions66–68. Null mutations in rpoS, which encodes the σS transcription factor, confer this advantage69. Other mutations, including some that enhance amino-acid catabolism under carbon starvation, also emerge in successive rounds of adaptation to the increasingly dire conditions69–71.

From spoT in the serial-transfer regime58 to rpoS during prolonged starvation66, many of the evolutionarily important mutations are found in global regulatory genes, rather than in genes that might improve single enzymatic steps. This conclusion is also supported by studies that have found, after a few hundred generations, widespread changes in patterns of protein expression in a chemostat-evolved population of E. coli70 and parallel changes in the transcription levels of many genes that are involved in central metabolism in three chemostat-evolved lines of Saccharomyces cerevisiae71. Beside these global regulators, many adaptive mutations involved specific regulatory changes, including the increased glucose transport72 and acetate cross-feeding64 in chemostat-evolved E. coli populations. Evolved changes in gene expression similarly implicate mutations in specific regulatory elements during adaptation by E. coli to high temperature2 and by Candida albicans to the antifungal compound fluconazole72,73. Also, several studies in which bacteria have been exposed to substrates that they cannot normally use have shown that regulatory mutations that cause increased expression of an enzyme with marginal activity on the substrate are important in the early stages of acquiring new catabolic functions74,75. By showing that substantial adaptation can involve a few mutations in regulatory genes, microbial evolution experiments support the famous conjecture by Mary-Claire King and Allan Wilson, based on the high degree of genetic similarity between humans and chimpanzees, that relatively few changes in regulatory genes might be responsible for important phenotypic differences74.

Diverse mutations emerge in evolution experiments, including point mutations, small insertions and deletions that cause frame shifts, and larger rearrangements. These rearrangements usually involve transposable elements that generate insertions, as well as inversions and deletions, through recombination between homologous elements in yeast76,77 and bacteria78–80. These elements are active in starving, as well as growing, populations. Some experiments have found sustained bursts of transposition that lead to an increase in the copy number of particular elements77–79. The underlying causes of these bursts are not well understood, but they might reflect a type of mutator activity (see later). Beyond their inherent interest, transposable elements are useful foci for genetic analyses of experimental lines because the mutations they cause are usually easier to find by molecular
methods than are point mutations. In some cases, different replicate lines have substituted insertions and point mutations in the same gene, which implies that either type of mutation can produce a similar advantage. In other cases, one type of mutation might prevail. For example, the 12 long-term E. coli lines lost the ability to catabolize ribose as a result of various deletions that were all mediated by an IS150 element located just upstream of the rbs operon. Genetic manipulations confirmed that such mutations provided a fitness advantage in the glucose-limited environment in which the deletions had evolved.

**Tradeoffs and the specificity of adaptation**

Sets of related genotypes, populations and species often show tradeoffs in their relative fitness across different environments. Indeed, without tradeoffs, a single type would be expected to prevail across all environments, precluding any comparison among the different types. If individuals usually encounter only one of these environments, then tradeoffs will tend to promote the evolution of specialists (FIG. 2). By contrast, if most individuals encounter a mix of environments, this might favour a generalist type that has the highest average performance, even if it is suboptimal in any constant environment. In principle, several mechanisms can produce tradeoffs. The simplest mechanism is antagonistic pleiotropy (AP), in which a particular mutation that is beneficial in one environment is harmful in the other. A second mechanism is mutation accumulation (MA), in which mutations accumulate by drift in genes whose products of which are not used in one environment but are useful in another. These mutations are, therefore, neutral in the environment in which they were substituted, but deleterious in the other environment. The third mechanism that can produce tradeoffs is the independent adaptation of organisms to alternative environments. If each of two populations substitutes a mutation that is beneficial in one environment and neutral in the other, then each population will be more fit in one environment than the other. A population does not suffer a decline in fitness relative to its progenitor under this third mechanism, unlike the first two. Under all three mechanisms, the net effect is a tradeoff in which different genotypes, populations or species are maximally fit in alternative environments. Although tradeoffs are widespread in nature, the underlying mechanisms are rarely known. Evolution experiments with microorganisms offer the opportunity to distinguish among the mechanisms.

**AP or MA?** There are many compelling examples of AP. Several experiments in which E. coli have evolved resistance to virulent phage show tradeoffs, such that the resistant bacteria are inferior competitors against their ancestors in the absence of the phage. These cases represent AP, and not MA, because the evolving bacteria would still benefit from being efficient competitors for resources; selection for resource use was not eliminated, but selection for resistance was added. The cost of resistance (the magnitude of fitness loss in the phage-free environment) varies for different phages and even among mutations that confer resistance to the same phage. Resistance to phage T4 occurs by mutations that cause defects in the lipopolysaccharide core of the cell envelope, and the cost is greater for mutations that produce defects that are more basal in this structure. Another example of AP comes from E. coli that evolved in lactose-limited chemostats; the same mutations that enhanced the permeability of the nonspecific OprP porin to lactose also increase the susceptibility of the cells to certain antibiotics. In the virus φX174, several mutations that had beneficial effects at high temperatures reduced fitness at lower temperatures.

During a long experiment, many mutations will be substituted in an evolving population. It therefore becomes difficult to test whether the same mutations that produce adaptation to one environment cause tradeoffs in other environments (AP), or whether different mutations produce the direct fitness gain and the correlated losses (MA). Even without testing each mutation, the temporal dynamics of change can give insights into the underlying process. During 20,000 generations on glucose, 12 E. coli populations tended to evolve reduced catabolic function against a battery of other substrates. This decay was fastest early in the experiment, which mirrored the trajectory of fitness gains on glucose, as expected for AP. Also, several populations evolved mutant phenotypes that had ~100-fold higher mutation rates, but showed only slight increases in their rates of catabolic decay, contrary to the expectation of a generalist type.
under MA. Another study with evolving E. coli populations found very different results — the decay rate of unused catabolic functions was much faster in mutator lines than in lines with functional DNA repair, and their rate of decay did not decelerate over time. At a first glance these two studies seem contradictory, but there is a crucial difference that, once understood, explains the different outcomes. In the study in which the effect of MA was dominant, the lines were propagated through single cell bottleneck that amplify the effects of random drift and eliminate the role of selection. Indeed, these lines became much less fit even under the conditions of their evolution, which indicated that there was no adaptation but only decay.

The evolutionary effects of mutators and very small populations are discussed further in two subsequent sections.

Another evolution experiment with E. coli, this one carried out in germ-free mice, found evidence for MA and showed that adaptive evolution had also occurred. In particular, mutator populations had a fivefold higher load of maladapted auxotrophic mutants than did populations with functional DNA repair. However, this study cannot evaluate the relative contributions of AP and MA to tradeoffs, because it did not test for any signal of AP. More generally, it should be recognized that AP and MA are not mutually exclusive, as both processes can occur in the same population.

An experiment with E. coli that evolved for 2,000 generations under several temperature regimes found a high level of thermal specificity — all the lines improved in fitness relative to their ancestor at the temperatures at which they evolved. However, in many cases, these lines did not lose fitness relative to their ancestor at nearby temperatures. This pattern corresponds to independent adaptation, as defined at the beginning of this section. Across a wider range of temperatures, the situation was more complex. Most lines that evolved at 20 °C lost fitness relative to the ancestor when they competed at 40 °C and above, whereas most lines that evolved at 41.5 °C did not lose fitness at 20 °C and below, which indicated asymmetries in correlated responses to selection in different environments. Moreover, although five lines that adapted to high temperature had no fitness loss relative to the ancestor at low temperature, one line did, which shows that correlated responses can be heterogeneous among replicate lines. Complex patterns of correlated responses are also evident when E. coli evolve on either glucose or maltose, with all other environmental factors held constant. Lines that have evolved on glucose show no improvement, on average, if competed against their ancestor on maltose, and the glucose-adapted lines are highly variable in their fitness on maltose. By contrast, lines that have evolved on maltose show consistent fitness gains on glucose as well as on maltose. This asymmetry indicates that the genetic adaptations to maltose might be a subset of the adaptations to glucose. Several studies have also shown that bacteria that evolved at a particular resource concentration showed greater fitness improvement when tested at that same concentration than at other concentrations.

However, despite this specificity of adaptation with respect to concentration, the overall similarity between the test environments was such that tradeoffs were usually absent and most correlated responses were positive.

**Generalists and specialists.** In Chlamydomonas that has been subjected to alternating light and dark conditions for several hundred generations, generalists evolved that were more fit than their ancestor under both conditions. However, the generalists were not as fit in either constant regime as specialists that had evolved under the corresponding condition. Similar outcomes were obtained with E. coli lines that evolved with either alternating or constant temperatures. These patterns conform to the expectation shown in Fig. 2.

Several experiments with viruses that are able to infect more than one host have found that viruses that evolved on one host became less fit (or at least did not improve) on alternative hosts. These results imply tradeoffs and that viral adaptation is host specific. However, if viral populations evolved on two alternating hosts, they sometimes improved as much on each host as those that had evolved on a single host in apparent contradiction to the evidence for tradeoffs and host specificity. To reconcile these results it is possible to imagine that there are two classes of beneficial mutations. One class is beneficial only on a particular host, and has antagonistic pleiotropic effects during growth on other hosts; mutations in genes that affect interactions with host receptors or other host-specific molecules are candidates for this class. The other class produces beneficial effects on all hosts; mutations in genes that are involved in RNA processing and elongation might be candidates for this class. Even if mutations with host-specific benefits were more common than the generally beneficial mutations, the latter class would be differentially enriched in viral populations that evolved on alternating host types. Further studies to identify and characterize both the direct and correlated fitness effects of individual mutations in viruses that evolved on single and alternating hosts would allow a test of these ideas. Such work is also relevant for developing attenuated vaccines and managing parae virulence.

**Emergence and consequences of mutators.** A ‘mutator’ is a genotype that has an increased mutation rate throughout its genome owing to a mutation that disrupts some aspect of DNA replication or repair. This effect can be large; in bacteria that have become defective in the methyl-directed mismatch-repair pathway, for example, the genomic mutation rate is increased by the order of 100-fold (Ref. 201). Genotypes with enhanced activity of a transposon might also behave like mutators, because the genome-wide rate of insertions is increased. The past few years have seen growing interest in mutators from various perspectives, including theoretical modelling, and surveys of natural populations as well as evolution experiments. The experiments aim to understand how mutators
Role of mutators in generating variation. Both mutator and normal (DNA-repair proficient) bacteria produce more deleterious than beneficial mutations. On a per capita basis, mutators produce more of both types. During adaptation to a new environment, mutators might promote faster adaptation by producing more beneficial mutations than do normal bacteria. However, this advantage is offset by the greater load of deleterious mutations that mutators produce. Figure modified with permission from REF. 103.

reach high frequencies in populations given their increased genetic load, and the evolutionary consequences of their spread. Towards these goals, some experiments showed the spontaneous emergence of mutators in populations that were founded by non-mutators with functional DNA repair 106–108, whereas others deliberately introduced mutators and documented the effects on rates of adaptive evolution 106–107,109.

At first glance, it is tempting to directly link the emergence and consequences of mutators by assuming that they reach high frequency because they accelerate adaptive evolution. Although this view might have some merit, it also presents several difficulties that demand a closer look. First, beneficial mutations are much less common than are deleterious mutations, and mutators suffer from a higher load of deleterious mutations (FIG. 3). The cost of an elevated mutation rate in terms of a higher load, coupled with the potential benefit of increased evolvability, represents a tradeoff. Such a cost also explains why most organisms retain DNA-repair functions. Second, the potential advantage of a high mutation rate in terms of accelerating adaptive evolution is limited in large asexual populations 103,104. In large populations, beneficial mutations occur even at low mutation rates, and asexuality gives rise to clonal interference which impedes the substitution of many beneficial mutations. Third, the view that mutators can become common by promoting adaptive evolution is teleological, or goal directed; that is, mutators might accelerate adaptive evolution once they become sufficiently common in a population to be an important contributor to the supply of beneficial mutations, but that cannot explain how they become common enough to have this effect.

This third point was shown by mixing different initial ratios of mutator and normal (repair proficient) bacteria, and propagating the mixtures until a beneficial mutation emerged in one clone that took over the population 104. When the initial ratio of mutator to normal cells was low (for example, 1:1000), the mutators tended to decline slowly in the short term owing to their greater load of deleterious mutations, and then much more quickly after the normal clone acquired a beneficial mutation. Although each mutator cell had a higher per capita probability of acquiring the first beneficial mutation than a normal cell, the combined probability of the mutator clone was lower than that of the normal clone owing to the difference in their total numbers. By contrast, if the initial ratio of mutator to normal cells was increased sufficiently (for example, 1:10), then the mutator clone prevailed, following a slight initial decline, because it was more likely to acquire the first beneficial mutation. A series of such experiments showed a threshold ratio below which the normal cells generally prevailed and above which the mutators usually won. This threshold is understood by realizing that the production of beneficial mutations depends on the product Nμμ, in which N is the cell number and μ, is the beneficial mutation rate. If this product is greater for the mutator clone than for the normal clone, a mutator will probably produce the first beneficial mutation, and vice versa. This experiment shows the difficulty of understanding how a mutator can spread after it first emerges in a population.

Nonetheless, mutators can become common and take over a population. In the long-term experiment with E. coli, 3 of the 12 populations spontaneously evolved into mutators in 10,000 generations 102 and a fourth by 20,000 generations 104. The mutations that produced these mutator phenotypes provided no direct competitive advantage, but spread by hitchhiking with beneficial mutations elsewhere in the genome 104. But how did a mutator clone become common enough to produce a beneficial mutation? The answer lies in understanding that the critical frequency of mutators, explained previously, represents a stochastic and not a deterministic threshold; in other words, there is a certain probability that a mutator, although its frequency is below the threshold, will produce the next beneficial mutation. Also, although mutators are purged by selection owing to their load of deleterious mutations, new mutators are constantly generated by mutations in genes that encode DNA-repair functions, giving rise to a quasi-equilibrium frequency of mutators. Although this frequency is below the threshold, each beneficial substitution in an evolving population provides another opportunity for a mutator to produce that mutation and hitchhike with it. If 10 beneficial substitutions occurred in every population, then each population had 10 chances to be converted to a mutator. Although the odds were against such conversion in any single case, the 12 populations collectively had 120 opportunities for conversions. Calculations using rough estimates of the relevant parameters (for example, the rate of mutation to mutator status) support this model 104,105. Following a conversion in which a mutator is substituted, the process is unlikely to be reversed until a population has become so well adapted to its present environment that the best mutation is one that reduces genetic load by restoring the lost repair function. However, a mutator might rise transiently to high frequency, then be eliminated, if the non-mutator type produces an even more beneficial mutation than that produced by the mutator 103,104.
In the long-term E. coli experiment, populations that became mutators showed fitness gains that were only slightly, if at all, greater than those that remained DNA-repair proficient. This finding indicates that evolvability per se might not have increased, despite the evidence that mutators spread by hitchhiking with beneficial mutations. This paradox highlights an important distinction between the causes and consequences of mutators. Theory indicates that the extent to which a mutator will accelerate adaptive evolution depends crucially on population size. When populations are moderate in size, a mutator can accelerate adaptive evolution by shortening the waiting time for a beneficial mutation to emerge. However, as population size becomes very large, even a low mutation rate suffices to generate beneficial mutations. Assuming that the population is asexual, clonal interference means that only one beneficial mutation can be substituted at a time. Hence, there is a ‘speed limit’ on the rate of adaptation in asexual populations. These predictions of this model were supported by experiments with E. coli and VSV in which population size and mutation rate were manipulated and the effects on the rate of adaptation were measured.

Let us summarize the present understanding of the emergence and consequences of mutators. Rapid adaptive evolution, as often occurs when a population encounters a new environment, causes many beneficial mutations to be substituted, and every such substitution provides another opportunity for a mutator to emerge. Whether mutators appreciably accelerate adaptive evolution depends on population size. In populations of moderate size, mutators might accelerate adaptation by reducing the waiting time for beneficial mutations. But in large populations, mutators might not accelerate adaptive evolution because the supply rate of beneficial mutations is not limiting. Two other factors also become important under certain conditions. First, as discussed in the section on tradeoffs, the accumulation of mutations in genes that experience relaxed selection will cause more rapid fitness loss by mutators if they later encounter environments in which those genes are important. Second, as discussed in the next section, mutators are deleterious in small populations that fight a losing battle between random mutation and drift, on one side, and selection to remove deleterious mutations, on the other.

Drift and decay in very small populations
So far, we have focused on adaptive evolution, which can occur because selection finds and amplifies rare beneficial mutations by means of differential survival and reproduction. But selection loses its discriminating power in very small populations in which success becomes a matter of chance. At the extreme limit, at which a population has only a single individual, there is no variation on which selection can act. Mutations do not stop, however, and are substituted at random. Over time, the average number of mutations in the genome increases (FIG. 4), and fitness declines because many more mutations are harmful than are beneficial. So, how and why would one study this process of genetic decay?

It would be difficult to follow a single cell, and remove one of two daughter cells when it divides, for many generations. However, it is possible to achieve a similar effect by periodically plating a population of cells (or viruses), randomly choosing a single colony (or plaque) and repeating the process indefinitely (FIG. 4). Each colony contains millions of individuals, but they are all derived from a single individual. Hence, this procedure is a simple way to subject a population to extreme bottlenecks. Although the population grows and mutations occur in the intervening generations, variation is eliminated at every bottleneck, such that random mutation and drift dominate the evolutionary process. By measuring the rate of fitness loss as well as the variation among replicate lines in the extent of their loss, it is possible to estimate the genomic rate of deleterious mutations and their average effect. The shape of the decay trajectories might also provide information on epistatic interactions between mutations. Besides genetic interest in these quantities, they are important for understanding the evolution of sex, and the survival and evolution of small populations.

Experiments with repeated single-individual bottlenecks have been carried out with several RNA viruses, a retrovirus, bacteria and yeast. In all these studies fitness declined, but the estimated mutational parameters differed greatly. The spontaneous deleterious mutation rate in the RNA-encoded VSV was estimated as ~1 per genome per generation, whereas for E. coli, the corresponding rate was only ~2×10−4 (Ref. 128). Estimates of the average fitness effects of deleterious mutations also vary, with the largest values in yeast and the lowest in VSV. However, it is unclear whether the differences in average effect are biologically meaningful or, instead, might reflect a statistical problem of estimating the mutation rate and average effect from the same data.

It is possible that RNA viruses, by virtue of their high mutation rates, have been selected to minimize the harmful effects of mutation by moving into regions of sequence space in which many mutations have less impact on performance. This process has been shown experimentally with self-replicating computer programs that can mutate and evolve. There are also other ways to reduce the harmful effects of mutations. For example, molecular chaperones help mediate the proper folding of proteins, perhaps including those that might otherwise fold improperly owing to mutations. Their elevated expression has been proposed to increase the mutational robustness of endosymbiotic bacteria, which face severe bottlenecks and, therefore, the potential for decay by random drift and mutation.
Conclusions

The experimental approach to studying evolution, especially using microorganisms, has greatly expanded over the past decade, with a wealth of studies that address a wide range of issues. We have focused here on those studies that inform our understanding of the temporal dynamics of evolutionary adaptation, the genetic changes that underlie adaptation, the causes and generality of tradeoffs during adaptation, the emergence and effects of mutator genotypes, and the process of genetic decay in very small populations. With respect to these issues, microbial evolution experiments support the following generalizations. First, populations adapt rapidly when they are introduced into new environments. However, they might continue to improve indefinitely, albeit slowly, even in a constant environment because beneficial mutations with ever smaller effects become increasingly accessible to selection. Second, genetic comparisons of ancestral and evolved organisms provide striking examples of parallel molecular evolution in replicate populations, including cases of adaptive mutations in genes that encode important global regulators. In contrast to these genetic targets of selection, most genes do not change, even over thousands of generations. Third, genetic adaptation to one environment is often, but not always, associated with fitness loss in other environments. Antagonistic pleiotropy is responsible for many of these tradeoffs, although elevated mutation rates can also reduce ecological breadth owing to mutation accumulation in genes that are under relaxed selection. Fourth, rapidly evolving asexual populations provide repeated opportunities for hypermutable genotypes to spread, along with the beneficial mutations that they occasionally generate. However, the emergence of such mutators does not always substantially accelerate adaptive evolution. Finally, in very small populations, the random processes of mutation and drift overwhelm the capacity of natural selection to retain well-adapted genotypes. Antagonistic pleiotropy is responsible for many of these tradeoffs, although elevated mutation rates can also reduce ecological breadth owing to mutation accumulation in genes that are under relaxed selection.

In the future, we look forward to the increased integration of genetic and phenotypic analyses, as well as to the improved temporal resolution of evolutionary dynamics. As an example, how is it possible to reconcile the evidence for the phenotypic divergence of replicate populations with the findings of parallel molecular changes? We see three possibilities. According to the first, different mutations in the genes that show parallel adaptation have heterogeneous pleiotropic effects on correlated traits. The second explanation proposes that although many genes undergo parallel evolution, it is unique adaptive mutations at other loci that are responsible for this phenotypic divergence. The third hypothesis is that mutations in genes that are under relaxed selection, which do not contribute to adaptation during the evolution experiment, cause the phenotypic divergence.

We also anticipate growing interest in evolution experiments that are carried out in silico, including simulation models that are based on microorganisms and studies that use abstract digital organisms. These
experiments will allow more detailed analyses of complex phenomena and, in turn, might stimulate increasingly sophisticated evolution experiments with real organisms. Already, many microbial evolution experiments have begun to explore the interesting dynamics that can emerge from complex interactions among several components of a larger system. Our second article for Nature Reviews Genetics will, therefore, focus on evolution experiments that address some of the fascinating interactions between multiple mutations in the same genome, between different genotypes in a population, and between interacting microbial species.


113. Acknowledgements We thank the reviewers for comments, and N. Haja for the photographic drawings used in Box 2. S.F.E. is supported at present by the Spanish Consejo Superior de Investigaciones Científicas. R.E.L. is funded by the National Science Foundation and the National Aeronautics and Space Administration.

Online links

DATABASES

The following terms in this article are linked online to: BioCyc: http://biocyc.org/.TeX?ac=Gnel.OLmpf.impl|p|007.FURTHER INFORMATION