APPLICATIONS OF NEXT-GENERATION SEQUENCING

Elucidating the molecular architecture of adaptation via evolve and resequence experiments

Anthony Long[1], Gianni Liti[2], Andrej Luptak[3] and Olivier Tenaillon[4]

Abstract | Evolve and resequence (E&R) experiments use experimental evolution to adapt populations to a novel environment, then next-generation sequencing to analyse genetic changes. They enable molecular evolution to be monitored in real time on a genome-wide scale. Here, we review the field of E&R experiments across diverse systems, ranging from simple non-living RNA to bacteria, yeast and the complex multicellular organism *Drosophila melanogaster*. We explore how different evolutionary outcomes in these systems are largely consistent with common population genetics principles. Differences in outcomes across systems are largely explained by different starting population sizes, levels of pre-existing genetic variation, recombination rates and adaptive landscapes. We highlight emerging themes and inconsistencies that future experiments must address.

[1]*Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697, USA.*
[2]*Institute for Research on Cancer and Ageing of Nice (IRCAN), Centre national de la recherche scientifique (CNRS) Unité Mixte de Recherche (UMR) 7284 — Institut National de la Santé et de la Recherche Médicale (INSERM) U1081 — Université de Nice Sophia Antipolis, 06107 Nice, France.*
[3]*Department of Pharmaceutical Sciences, University of California, Irvine, California 92697, USA.*
[4]*Institut National de la Santé et de la Recherche Médicale (INSERM), Université Denis Diderot Paris 7, Unité Mixte de Recherche (UMR) 1137, Infection, Antimicrobials, Modelling, Evolution (IAME), 75205 Paris, France.*
Correspondence to A.L.
e-mail: tdlong@uci.edu
doi:10.1038/nrg3937
Published online
8 September 2015

The incredible diversity of life results from adaptation in response to a changing environment. Our understanding of how adaptation occurs at the molecular level is surprisingly rudimentary and is derived mostly from comparisons within and between species. Identifying the causative beneficial mutations that give rise to species differences from such genomic comparisons, not to mention inferring their effects on selection and dynamics, remains a challenge. As a result, the adaptive landscape on which organisms evolve is still largely uncharacterized, and debates persist as to how rates, effects and interactions among beneficial mutations and their environments determine allele frequency change[1–4].

Over the past century, different fields of biology have converged towards the use of evolution in the laboratory to study the process of adaptation. Microbiologists[5,6], geneticists[7,8], biochemists[9–11] and population geneticists[12,13] have found that experimental evolution, in which replicate populations of diverse model organisms are allowed to adapt to novel but controlled laboratory environments (FIG. 1), could shed new light on the biological processes they studied. For a long time, the different fields focused on the dynamics of adaptation and/or on the phenotypic or physiological effects of adaptation on organisms, rather than on the underlying genetic changes, which were not easily accessible. However, with the rise of sequencing technologies, the evolve and resequence (E&R) approach[14] can uncover the molecular determinants of adaptation in many different systems. The sequencing of evolved RNA molecules[15] and viruses[16] was followed by that of bacteria[17–20], yeast[21] and *Drosophila melanogaster*[22], such that it is now possible to compare E&R experiments across systems. Indeed, the E&R approach affords the possibility of watching evolution occurring in real time and on a genome-wide scale, and fitting dynamic models to the resulting data[23,24]. Highly replicated E&R experiments, especially those using temporal sampling, now allow for several long-standing questions on the nature and dynamics of molecular evolution to be addressed across models. What is the role of standing genetic variation versus newly arising mutations in contributing to selection response? How reproducible is evolution at the molecular level[25]? Do the selection coefficients (represented by the variable *s*) of beneficial alleles change during evolutionary time as a function of distance to a new fitness optimum[1]? What is the role of protein-coding versus regulatory variation in adaptation[26]? E&R experiments might begin to resolve these debates, as the trajectories of loci contributing to adaptation can be observed in real time.

E&R experiments are currently being performed in various vastly different systems. The four main systems we present in this Review are: *in vitro* evolution of libraries of oligonucleotides (RNA or DNA) selected under defined chemical conditions; asexually evolving bacteria or yeast with selection initiated from an initially isogenic

founder strain; asexually or semi-sexually evolving yeast with selection initiated from a synthetic founder population derived by intercrossing a small number of diverse natural strains; and obligate sexual *D. melanogaster* with selection initiated from an outbred population of hundreds of individuals sampled from the wild and established in the laboratory (TABLE 1). Evolution of proteins[27–29] or systematic evolution of ligands by exponential enrichment (SELEX)-type experiments used to study transcription-factor-binding sites[30,31] are examples of other fields of *in vitro* evolution that have benefited from next-generation sequencing (NGS), but we do not focus on these systems here. The systems we review here are extremely diverse in various ways: the appearance on Earth of the equivalent of these organisms spans billions of years (in the case of *in vitro* evolution, the 'organisms' are not even living); genome sizes span seven orders of magnitude (from ~50 nucleotides to ~10^8 bp); and reproduction and propagation range from fully asexual to fully sexual. Interestingly, the field of *in vitro* E&R experiments has developed almost completely independently of *in vivo* work, despite the underlying evolutionary principles being identical. The outcome of E&R experiments across systems seems to be different. *In vitro* evolution experiments typically result in multiple evolved solutions, asexual microbial systems seem to exhibit fixation of a small number of new mutations with divergence of solutions across replicates, and sexual systems initiated from population samples seem to exhibit a highly polygenic response with a high degree of parallel evolution across replicates. These observations create a theoretical dilemma: is the molecular basis of adaptation different in different systems, or are observations generally consistent with one another when different aspects of experimental design are taken into consideration? All of these systems evolve according to the laws of natural selection (BOX 1); thus, the framework of population genetics should capture the details of their evolution. Nevertheless, population genetics is predictive only when applied to a given adaptive landscape (BOX 2), which presumably differs across systems.

In this Review, we aim to reconcile observations from E&R experiments across these disparate systems. We first describe the differences between systems in terms of experimental setting and population genetics parameters and explain how these may affect the dynamics of adaptation. We then review what we have learned using E&R methods across this diverse set of systems. We focus on the population dynamics of beneficial alleles, the molecular bases of adaptation, inferences that can be drawn from parallel adaptation in replicate evolving populations and the role of epistasis in adaptation. Finally, we attempt to identify experiments that can potentially clarify remaining discontinuities between the systems.

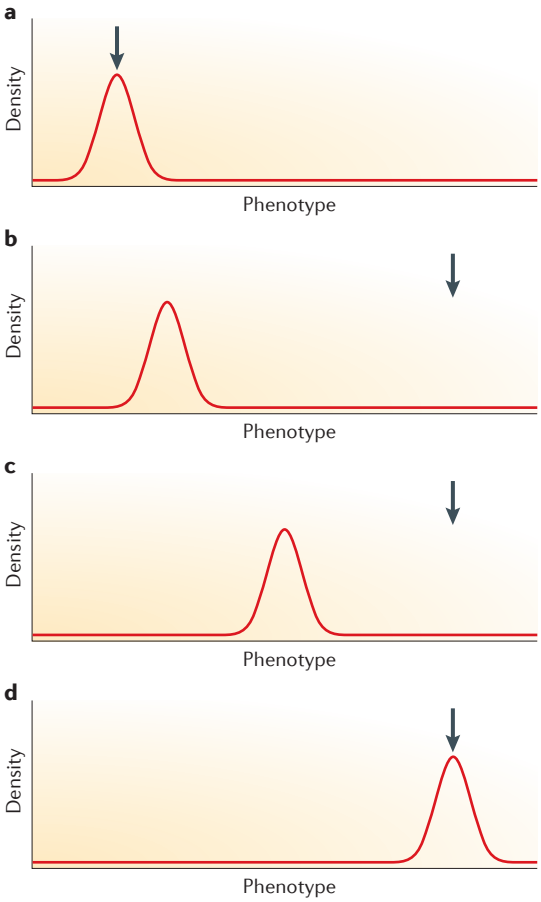


Figure 1 | **A conceptual experimental evolution experiment.** **a** | Starting with a population of organisms, cells or *in vitro* molecules, the initial distribution of phenotypes will track the average fitness conditional on phenotype (arrow) in the ancestral environment. **b** | In an experimental evolution experiment, the fitness optimum is manipulated through a shift in the environmental conditions under which the system is propagated (for example, by changing the temperature, adding a chemical to the media or forcing molecules to bind to a ligand). This shift redefines the phenotypic optimum relative to the population's average phenotype. **c**,**d** | The population will then attempt to track the new optimum via natural selection, using standing variation and/or newly arising mutations. The speed and mode of adaptation will depend on the system.

Standing genetic variation
Pre-existing genetic diversity in a population of interest.

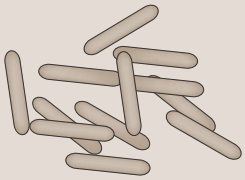
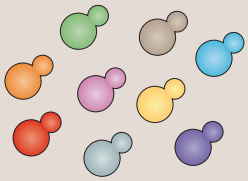
Selection coefficients
Measurements of the proportional change in the fitness of a genotype owing to a mutation (represented by the variable *s*). The fitness of that genotype is calculated as 1 – *s*.

Fixation
When an allele of an initially polymorphic locus or haplotype reaches 100% relative frequency in the population.

E&R: the systems

In vitro *systems.* *In vitro* selection and evolution experiments[32–34] start with highly diverse libraries of short DNA oligonucleotides (~10^{14}–10^{16} different molecules) that typically consist of random regions of ~30–200 bp flanked by primer-binding sequences. *In vitro* evolution consists of a transcription step, to generate a corresponding RNA pool; an enrichment step, during which the RNA population is exposed to a defined chemical condition (for example, binding to ATP-coated beads or ligation to another nucleic acid); and an amplification step, in which the enriched library is reverse-transcribed back to DNA and amplified by PCR[10]. The enrichment and amplification steps are typically repeated for 10–20 cycles. The starting bulk population of molecules often does not show any detectable activity in the assay, but after 3–5 rounds there is appreciable activity, and after 8–12 rounds the population demonstrates strong

Table 1 | **The four main E&R experimental paradigms discussed in this Review***

				
	In vitro	**Microbial isogenic**	**Microbial outbred**	**Obligate sexual higher eukaryotes**
Example models	Synthetic DNA or RNA molecules	Bacteria, haploid yeast or diploid yeast	Diploid yeast	*Drosophila melanogaster*
Sexual recombination	Mostly absent; occurs at a low frequency owing to template switching during PCR; can be intentionally facilitated through techniques such as synthetic shuffling	Mostly absent, but some experiments use plasmid exchange in bacteria; optional in yeast	Starting population obtained by crossing different strains; optional during evolution	Obligate
Population size	Up to 10^{16}	Up to 10^{10} per ml	Up to 10^{9} per ml	10^{2}–10^{3}
Initial genetic variation present	Extremely high, but limited by oligonucleotide synthesis technology	None	Moderate; unique haplotypes genome-wide obtained via recombination from 2–16 naturally occurring founders	High; experiments typically initiated from ~100 naturally occurring strains obtained from the wild
Initial variation in fitness	Extremely high, with the vast majority of molecules having a fitness of close to zero	None	High, and limited only by natural variation in fitness	High, and limited only by natural variation in fitness
Main source of genetic variation that is responsible for adaptation	Variation present in the starting population, with some modifier mutations	Newly arising mutations	Standing genetic variation, with newly arising mutations after several hundred generations	Standing genetic variation
Role of clonal interference	Strong	Strong	Strong when evolution is asexual, weak otherwise	Weak

E&R, evolve and resequence. *This table highlights salient information on the type of model, sexual recombination, population size and initial level of genetic diversity in the different model systems.

activity in the assay and converges on a few dominant motifs[35]. Similar experiments have been performed directly with single-stranded DNA[10,36,37], 'mosaic' nucleic acids (in which the backbone was scrambled between ribose and deoxyribose)[38], nucleic acid analogues[39,40] and proteins linked to their coding mRNAs or cDNAs[41,42].

Owing to the small 'genome' sizes being considered and the relatively low complexity of the population following *in vitro* selection, Sanger sequencing has been used for decades to identify the RNA motifs that 'won' the evolutionary competition. Furthermore, the total diversity of the population over rounds of selection could be followed using restriction enzymes[43]. NGS allows testing of the selected population in earlier rounds and provides a much more complete picture of the final array of winning genotypes[44]. In fact, direct experimental measurements of the fitness landscapes of RNA ligase[15], kinase[45], Diels–Alderase[46] and self-splicing[47] ribozymes are now obtained at a much higher resolution than ever before. Combined with microfluidic analytical platforms, NGS can be directly coupled to activity measurement, revealing the fitness landscape of a functional RNA[48,49].

Ribozymes
RNA molecules that are capable of catalysing chemical reactions. Natural ribozymes include ribosomal RNAs, spliceosomal RNAs, RNase P RNA, self-splicing introns and self-cleaving ribozymes.

Asexual microorganisms with an isogenic starting population. The dominant model consists of an initially isogenic population of 10^{6}–10^{8} bacteria or yeast that is evolved asexually. The evolution protocol relies on the renewal of the media the microorganisms use to grow. This renewal can be made either through serial transfers, in which a fraction of a saturated culture is regularly diluted into fresh media[50], or continuously using a chemostat[51]. Based on the dilution factor or flow rate, the evolutionary rate can be computed (for example, a daily 100-fold dilution implies $\log_2(100) = 6.64$ generations per day). Given estimated growth rates, hundreds to thousands of generations of evolution can occur in a few weeks or months in a microbial system. The keystone example of microbial experimental evolution is the *Escherichia coli* long-term evolution experiment (LTEE), which has been running for more than 25 years and has now reached more than 60,000 generations of evolution[52]. To put this in perspective, in humans each generation takes ~20 years to produce the next, so going back 60,000 generations would take us to ~1.2 million years ago, a time that predates the emergence of the species *Homo sapiens*.

Box 1 | **Population genetics of adaptation**

Population genetics can model the forces that contribute to rates of adaptation.

Emergence of beneficial mutations

Changes in the mean fitness of the population require variation in fitness in a population. In an initially clonal population, this diversity comes from newly arising mutations, although for adaptation to occur a subset of newly arising mutations must be beneficial. Population size then affects the rate of adaptation in two ways. First, the number of beneficial newly arising mutations is a linear function of population size. Second, population size determines the fluctuations in allele frequency from one generation to the next through a process known as random genetic drift. In a finite population, a beneficial mutation with a selective advantage smaller than the reciprocal population size is effectively neutral (and its probability of ultimate fixation is thus $1/N$, where N is the population size)[145]. Even if a beneficial allele has a selective advantage larger than $1/N$, it can still be lost owing to drift, with the probability of eventual fixation being roughly proportional to its selective coefficient[146]. Nevertheless, beneficial alleles that are lost tend to be lost early on, as after an allele reaches a copy number greater than the reciprocal of its fitness effect its dynamics are essentially deterministic. From these simple, well-known population genetics results[147], adaptation of an initially clonal population will be marked by a delay corresponding to the emergence of low-frequency beneficial mutations, their survival against drift and their deterministic increase in frequency until they detectably affect the mean fitness of the population.

Fixation of beneficial mutations

The fate of mutations that survive drift is then highly affected by the existence of genetic exchange. In the absence of recombination, mutations stay coupled to the genetic background on which they arise. Consequently, the beneficial mutations that survived drift compete with one another in a process called clonal interference. Mutations with the highest fitness, or the combination of mutations with the highest fitness, will eventually reach fixation, and other beneficial mutations will be lost. The number of beneficial mutations ultimately lost in the clonal population is very large[71]. By contrast, in the presence of high levels of recombination, mutations can switch backgrounds, and many mutations can simultaneously increase in frequency in a population. Intermediate levels of recombination can result in genetic 'hitchhiking', which involves neutral passenger mutations on the same background as (and closely linked to) a beneficial allele being swept with the beneficial allele and achieving a high frequency due solely to linkage. Distinguishing a beneficial allele from a neutral hitchhiker may require functional assays.

Adaptation from standing genetic variation

When the population starts with a high genetic diversity, as opposed to a single clonal individual, much of the initial variation in fitness in the population is due to standing genetic variation. Current models assume that there are many loci at intermediate frequencies that are essentially neutral in the starting population, but that when that population is placed in a novel environment those alleles start to change in allele frequency owing to natural selection. Initially, change will be strongest for alleles at intermediate frequency (as they contribute the most to the variance in fitness), and as a result phenotypes or fitness can exhibit rapid change[12]. Furthermore, in synthetic populations resulting from crosses among a limited number of different clones, the initial allele frequency of each allele is markedly higher than the drift threshold, and allele frequency change can be virtually deterministic (unless fitness is affected by *de novo* mutations or complex epistatic interactions between distant sites). For this reason, experimental evolution experiments in *Drosophila melanogaster*, in which the starting populations are typically derived from around 100 individuals, may be very different from those in yeast populations derived from four isogenic founders. In addition, experimental evolution populations of *D. melanogaster* will have effective population sizes several orders of magnitude smaller than those of yeast, thus requiring a higher initial frequency for each mutation to be above the drift threshold.

Long-term adaptation

Population genetics can be predictive for the early stage of adaptation, provided that we know the distribution of fitness effects, the mutation rate and the recombination rate. When adaptation involves the combination of several adaptive mutations, we need to know further global properties of the adaptive landscape, in particular those characterizing microscopic and macroscopic epistasis (BOX 2), which define the shape of the landscape.

Microorganisms with a synthetic outbred starting population. A less-mature branch of microbial experimental evolution uses synthetic populations of yeast generated through the intercrossing of two or more isogenic highly characterized founder strains[53,54]. Although these populations maintain all of the desired features of the microbial model, including large population size and replicate populations, they also introduce the important innovations of sexual reproduction and standing genetic variation in the starting population. The evolution of these populations potentially mimics that of obligate sexual higher eukaryotes more closely than initially isogenic asexual systems. These synthetic outbred populations are evolved asexually via serial passage in liquid or in solid culture for several hundred generations at effective population sizes of 10^6–10^8, with entire pooled populations being resequenced in a high-coverage manner. This 'Pool-seq' protocol (which can be applied to diverse species and systems)[20,55] results in an estimate of allele frequency in the population at every single-nucleotide polymorphism (SNP) in the genome. Despite yeast synthetic populations being derived from a small number of founder genomes, multiple rounds of intercrossing result in a near infinite number of genome-wide haplotype combinations (FIG. 2).

Obligate sexual higher eukaryotes with an outbred starting population. *Drosophila* species are long established as a system for the study of experimental evolution in the laboratory[56,57]. In the past decade, several groups have resequenced pooled DNA samples from experimentally evolved laboratory populations of *D. melanogaster* to identify regions of the genome responding to laboratory-imposed selection[7,14,22,58–60]. The details of E&R experiments in *D. melanogaster* are quite different from those previously discussed. The starting populations are founded from collections of flies sampled from the wild, which can harbour hundreds of natural haplotypes, so individual genetic variants can be rare. Population sizes are modest by the standards of microorganism or RNA molecule populations, ranging from 100 to 1,000 individuals. Evolution experiments in *D. melanogaster* are also modest in terms of the number of generations of selection carried out: a 1–3-year experimental evolution experiment in *D. melanogaster* covers only 25–75 generations, as the egg-to-egg life cycle of a fly is almost 2 weeks. Nevertheless, there are a few longer-term (200–300-generation) ongoing experiments that are being actively investigated[22,61,62].

Population dynamics of loci during selection

Despite population genetics theory having a wealth of models dealing with the dynamics of alleles responding to selection, there is a relative dearth of empirical data sets in which allele trajectories are directly observed. E&R experiments coupled with temporal sampling of populations allow the direct study of how allele frequencies change over time. The speed and the magnitude of allele frequency changes tell us a great deal about the selective advantages associated with different variants. Other patterns of change, such as alleles plateauing at

Box 2 | **Adaptive landscapes**

The evolutionary fate of a population depends on the particular mutations sampled during its evolution. The aim of the adaptive landscape metaphor is to find a visual way to illustrate such possibilities. The genetic space is virtually infinite. For a genome composed of *n* biallelic loci, the genetic space is a hypercube with 2^n states, which cannot be visualized for more than five loci. Hence, a landscape metaphor is used in which the vertical axis is fitness and the horizontal axes represent a continuous vision of the genetic space, in which proximity suggests genetic similarity (see the figure). Although metaphoric, analysis of the adaptive landscape can also be quantitative at a local and a global scale. In parts **a**–**d** of the figure, we illustrate experimental approaches for exploring the adaptive landscape relative to a single genotype, represented by the blue dot.

Local scale

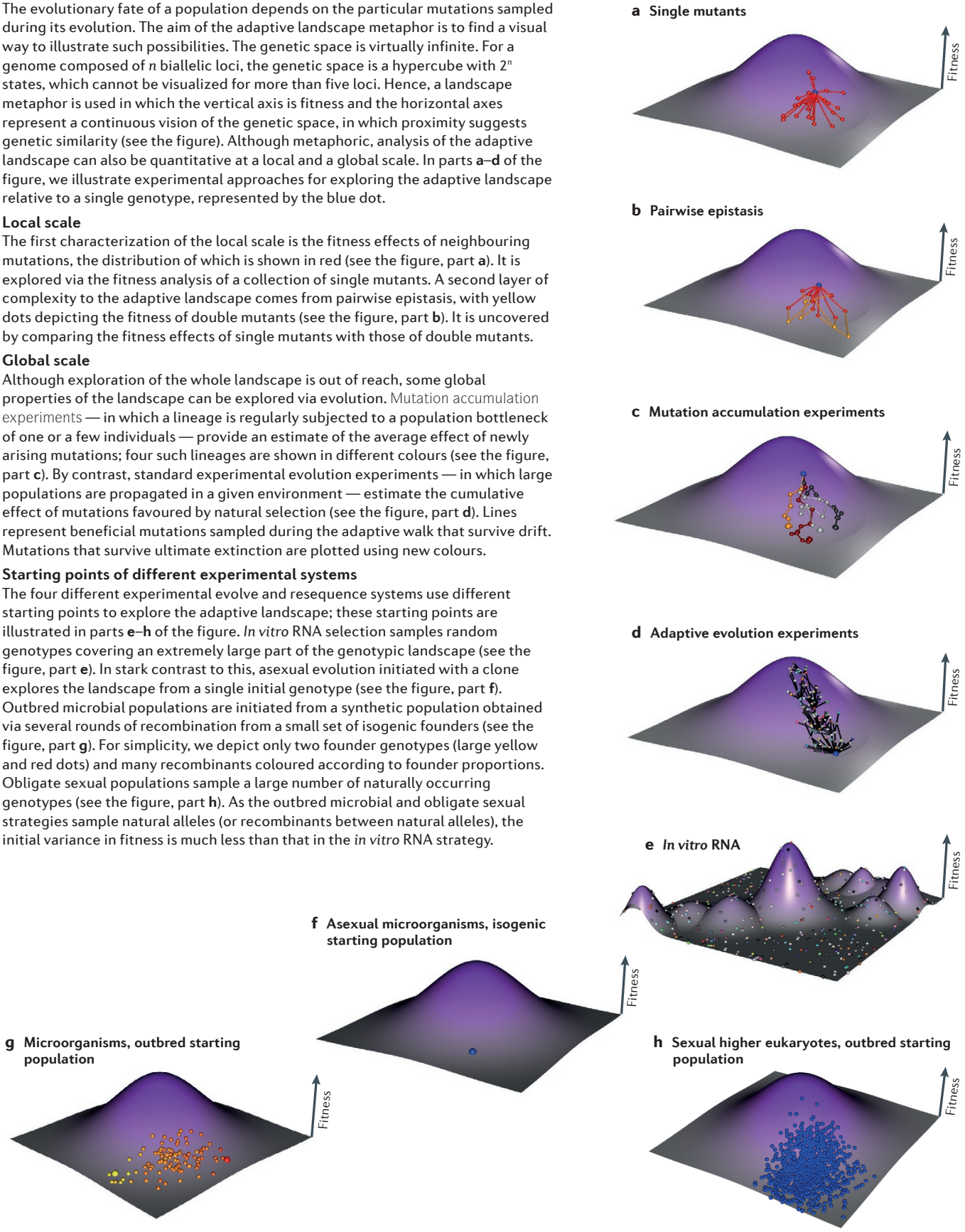
The first characterization of the local scale is the fitness effects of neighbouring mutations, the distribution of which is shown in red (see the figure, part **a**). It is explored via the fitness analysis of a collection of single mutants. A second layer of complexity to the adaptive landscape comes from pairwise epistasis, with yellow dots depicting the fitness of double mutants (see the figure, part **b**). It is uncovered by comparing the fitness effects of single mutants with those of double mutants.

Global scale

Although exploration of the whole landscape is out of reach, some global properties of the landscape can be explored via evolution. Mutation accumulation experiments — in which a lineage is regularly subjected to a population bottleneck of one or a few individuals — provide an estimate of the average effect of newly arising mutations; four such lineages are shown in different colours (see the figure, part **c**). By contrast, standard experimental evolution experiments — in which large populations are propagated in a given environment — estimate the cumulative effect of mutations favoured by natural selection (see the figure, part **d**). Lines represent beneficial mutations sampled during the adaptive walk that survive drift. Mutations that survive ultimate extinction are plotted using new colours.

Starting points of different experimental systems

The four different experimental evolve and resequence systems use different starting points to explore the adaptive landscape; these starting points are illustrated in parts **e**–**h** of the figure. *In vitro* RNA selection samples random genotypes covering an extremely large part of the genotypic landscape (see the figure, part **e**). In stark contrast to this, asexual evolution initiated with a clone explores the landscape from a single initial genotype (see the figure, part **f**). Outbred microbial populations are initiated from a synthetic population obtained via several rounds of recombination from a small set of isogenic founders (see the figure, part **g**). For simplicity, we depict only two founder genotypes (large yellow and red dots) and many recombinants coloured according to founder proportions. Obligate sexual populations sample a large number of naturally occurring genotypes (see the figure, part **h**). As the outbred microbial and obligate sexual strategies sample natural alleles (or recombinants between natural alleles), the initial variance in fitness is much less than that in the *in vitro* RNA strategy.



intermediate frequencies, suggest that simple population genetics models of adaptation do not fully capture the dynamics of adaptation.

In vitro selection systems are characterized by beneficial alleles having dramatic selective advantages. A variant starting at a frequency of 10^{-16} can fix in 10 rounds of selection[43], suggesting that the selective coefficient (or coefficients) associated with beneficial oligonucleotide haplotypes is greater than 50%. Initially there are extremely high levels of genetic diversity but, as oligonucleotides are randomly synthesized, the initial fitness of the vast majority of molecules is essentially zero (FIG. 2a). Starting from initially random sequences and selecting for a simple biochemical function results in a rapid loss of diversity and dramatic fitness improvement.

The dynamics of allele frequency change in initially isogenic microorganisms (such as bacteria or yeast) evolving asexually are surprisingly complex. In the absence of genetic exchange, these mutations compete with one another through a process called clonal interference[12,63] (BOX 1). The total number of potential beneficial mutations genome-wide is large enough that it is virtually impossible for a beneficial mutation to arise and reach fixation without having to compete with some other beneficial mutation that arose in a competing lineage in the population. Hence, as shown elegantly in yeast[64], in large asexual populations evolution is not characterized by single beneficial mutations reaching fixation but instead by 'lucky' combinations of mutations[34,65] that manage to outcompete other such combinations. Consequently, the allele frequency of a beneficial mutation may initially increase but then may decrease or even reverse as alternative combinations of mutations are being selected for[66–68] (FIG. 2b). Although the selective advantage of the beneficial mutations observed depends drastically on the selective regime[69], it is quite common to observe *s* values ranging from 1% to 20%[70,71]. These lower values, in comparison to those from *in vitro* evolution, may be due to pleiotropic constraints imposed on newly arising mutations. Those pleiotropic effects may increase with organismal complexity and favour alleles of more modest effect. Conversely, as fitness effects are measured relative to average fitness of the population as a whole, fitness effects may seem larger in *in vitro* evolution purely as a result of most oligonucleotides in the initial population having a fitness of essentially zero.

The dynamics of allele frequency change are quite different if selection is initiated from an outbred starting population, as in yeast (FIG. 2c). As initial genetic diversity is high and every allele starts from a high initial frequency (for example, ~25% in '4-way' populations derived from 4 isogenic founders), the response to selection is rapid. The selection coefficients inferred from the rate of allele frequency change are generally ~1%, unless the selective pressure is intense, in which case *de novo* mutations of strong effect may also come into play. Genome-wide patterns of allele frequency change over time allow inferences about how variation present at the start of the experiment is sorted out during adaptation. Specifically, regions of a few kilobases in size harbouring beneficial or detrimental alleles can be precisely identified[53,72]. Furthermore, for the first several hundred generations of evolution, virtually all adaptation is from standing genetic variation present in the starting population[53,54]. Through manipulations of the culturing regime, yeast populations can be forced to participate in sexual reproduction at regular intervals, providing insights into long-term effects of sex in evolving populations[73,74]. Yeast synthetic populations adapting to a novel laboratory environment with weekly episodes of sexual recombination show adaptation patterns comparable to those of asexually evolving synthetic yeast populations, at least for the first several hundred generations[75].

Experiments in *D. melanogaster* are initiated from outbred starting populations harbouring hundreds of natural haplotypes at each genetic locus. Thus, they are similar to the outbred yeast system discussed above, except that many more natural haplotypes are used to found the starting population. Furthermore, because recombination is obligate in *D. melanogaster*, clonal interference is unlikely to impede evolution, and two

Figure 2 | **E&R experiments reveal the dynamics of adaptation on a genome-wide scale.** From sequencing data sets of individuals or pooled populations (Pool-seq), the evolution of haplotype diversity (upper panel of each figure part) and that of allele diversity at each site of the genome (bottom panel of each figure part) can be uncovered. Haplotypes are coloured according to the initial genome-wide haplotypes, and the middle panels of each figure part show example haplotypes from the start, middle and end of the evolve and resequence (E&R) experiment. Population heterozygosity is shown at each site of the simulated genome, with the colours red to green indicating decreasing levels of heterozygosity. **a** | For *in vitro* experiments, the great initial haplotypic diversity is lost within a few generations, as the best haplotype (or haplotypes) quickly increases in frequency, and the large majority of initial haplotypes have a fitness of nearly zero. Per-site heterozygosity is homogeneous and mostly decays through the adaptive process. **b** | In asexual microbial evolution from an isogenic starting population, the initial diversity is minimal, and it can build up only through newly arising mutations. In the clonal interference regime depicted here, several haplotypes compete with one another to reach fixation (upper panel). Diversity in the genome is maintained at only the beneficial sites and is therefore highly heterogeneous across the different sites of the genome (lower panel). New mutations are eventually lost or fixed, so heterozygosity is transitory. **c** | In yeast asexual evolution from a synthetic outbred starting population, the initial diversity is high but organized in blocks resulting from recombination between the founding parents. Similarly to in the *in vitro* studies, adaptation is characterized by a genome-wide loss of haplotype diversity. However, the rate of heterozygosity loss is lower than that in the *in vitro* case, as there is much lower variation in initial fitness and perhaps selection is not as strong. The regional loss of diversity will depend on variation in the fitness of the different haplotype blocks present, with regions not affecting fitness losing diversity later. For details on the parameters used to generate the panels in this figure, see Supplementary information S1 (box).

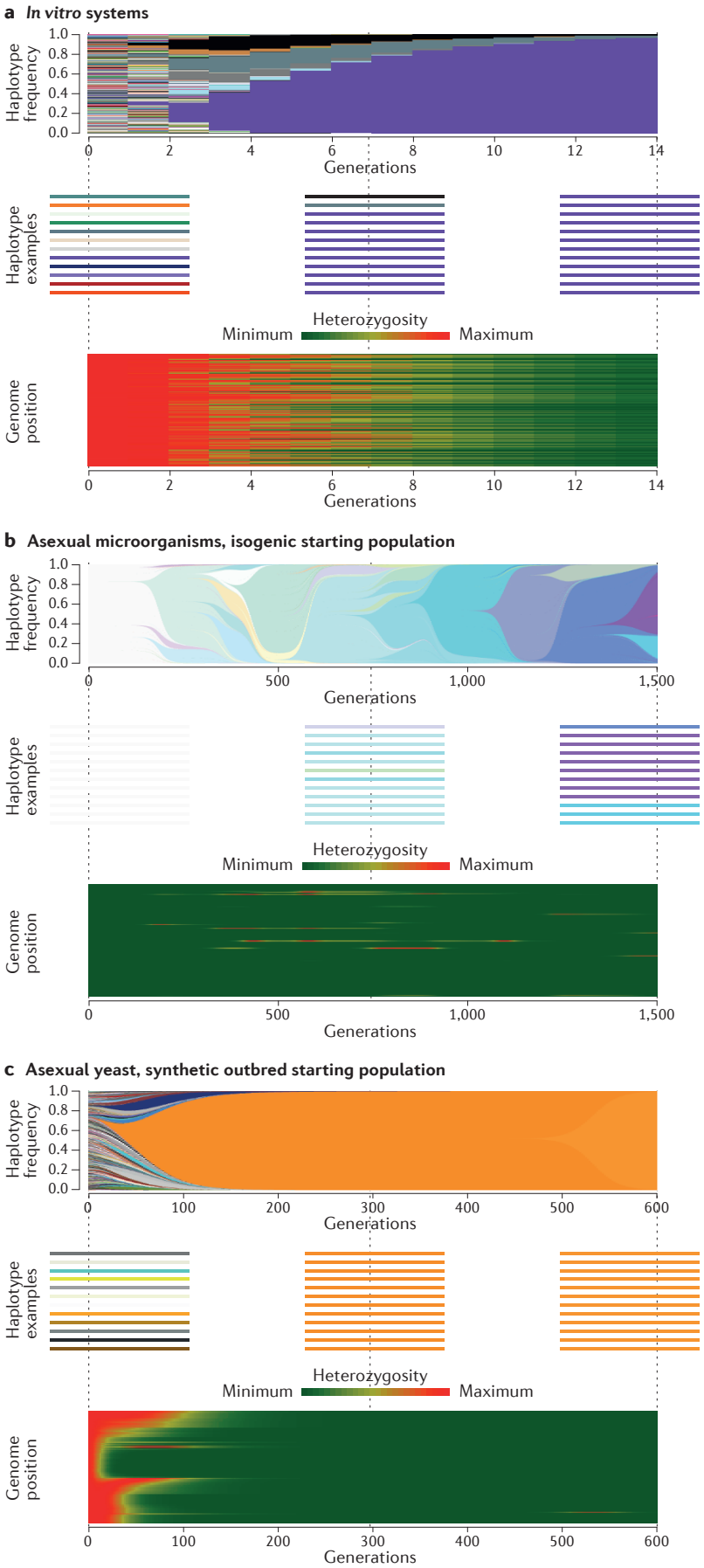
Haplotype
The ordered collection of alleles along a single chromosome.

Mutation accumulation experiments
Experiments in which an initially isogenic strain is propagated for many generations with severe population-size bottlenecking (often to a single cell or individual) without voluntary selection. The mutations that distinguish the accumulation strain from its ancestor can be used to estimate mutation rates.

Clonal interference
A phenomenon observed in asexually evolving systems. Owing to a lack of recombination, clones harbouring different combinations of mutations compete against one another to reach fixation.

Pleiotropic
A genetic change affecting more than one phenotype.

Average fitness
The average fitness of a population is defined as the weighted sum of the fitness values associated with each genotype, where the weights are the frequencies of those genotypes. In an *in vitro* evolution experiment, there could initially be several million genotypes, with the vast majority having fitness values close to zero.



regions of the genome can independently respond to natural selection. Finally, the populations of *D. melanogaster* used to found E&R experiments are typically derived from hundreds of wild haplotypes; thus, alleles with minor allele frequencies of 1–5% can contribute to selection response, unlike in the outbred yeast system. Consistent with these properties, sliding window plots of heterozygosity versus genome position tend to show several dozen regions of the genome experiencing dramatic reductions in heterozygosity punctuated by the vast majority of the genome maintaining substantial variation (FIG. 3), with changes in absolute allele frequency showing similar local spikes[22]. Notably, regions showing reductions in heterozygosity never show a drop to zero, consistent with the idea that adaptation is due to selection on standing genetic variation, that selective sweeps are happening but have had insufficient time to reach fixation or that advantageous alleles plateau in frequency before they reach fixation.

Overall, the strength of the selective advantage achieved because of beneficial alleles seems to vary across systems, as well as within systems, and it depends on the experimental system and on the nature of the selective pressure. When these factors are taken into account, population genetics provides a good qualitative description of the dynamics of adaptation of these highly variable systems. An important observation that is currently being actively debated is the extent to which allele frequency change can plateau in longer-term experimental evolution experiments initiated from outbred sexual populations[58,75,76]. If the alleles that are important in adaptation plateau before reaching fixation, then average selection coefficients in longer-term experiments will not accurately reflect an initially very rapid change in allele frequency at selected sites.

Types of mutations recruited

By virtue of sequencing the entire genome of evolved clones or populations, E&R experiments can potentially unravel the molecular nature of beneficial alleles and determine some notion on their functions, frequencies and interactions.

The complexity of the DNA pools used for *in vitro* experimental evolution is limited by the capabilities of DNA-synthesis platforms, which typically yield ~10^{16} different molecules. The theoretical diversity of a random DNA polymer is 4^N, where N is the length of the random region. For DNA pools longer than 26 random nucleotides, the number of different molecules possible is larger than 10^{16}, which means that the actual diversity in these experiments is lower, often vastly so, than the theoretical limit (for example, for a 60-nucleotide random pool, the theoretical limit is 4^{60}, which is ~10^{36} — 20 orders of magnitude above the actual complexity). Thus, many *in vitro* experimental evolution experiments are initiated with pools that greatly under-sample the sequence space, leaving much room for fine-tuning during the subsequent selection rounds. In one example, selection for GTP aptamers directly yielded an optimal family of aptamer sequences, suggesting that these sequences were present in the initial pool; by contrast, other aptamer

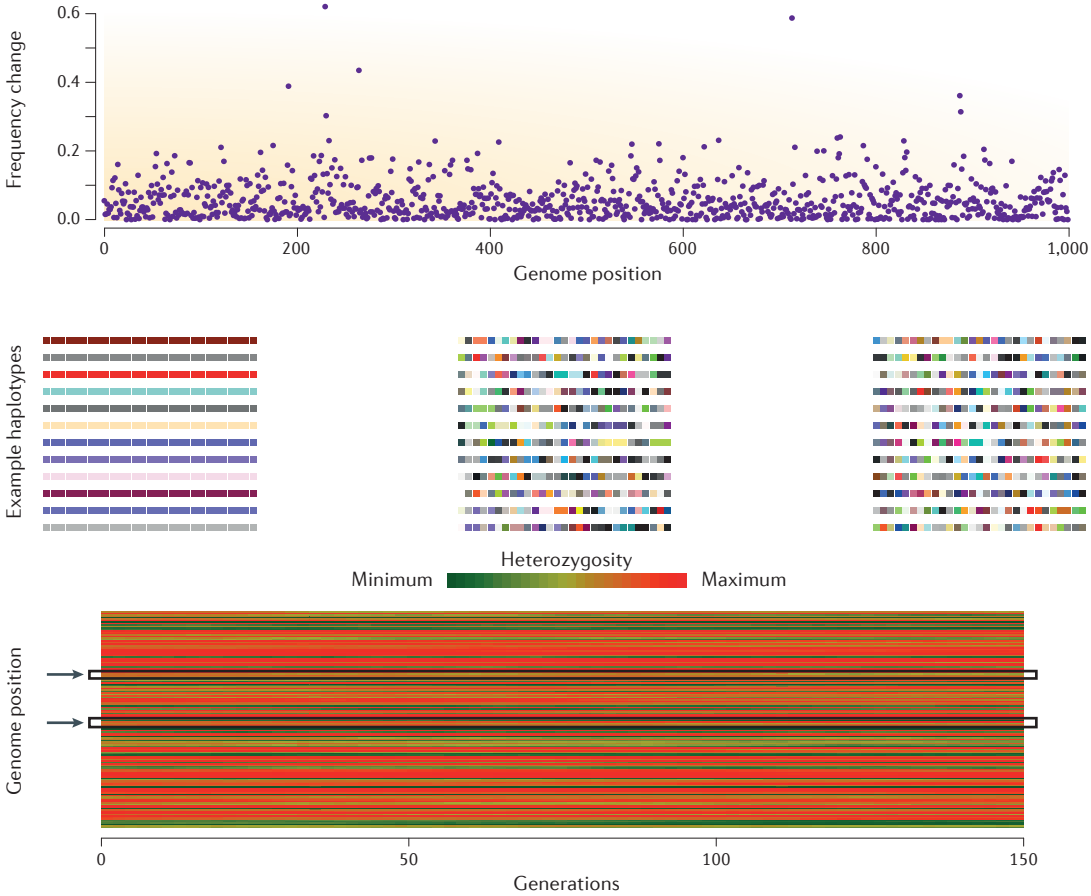


Figure 3 | **E&R experiments in sexually reproducing species.** Evolutionary patterns in obligate sexual evolve and resequence (E&R) experiments initiated from an outbred population are different from those in the asexually evolving examples shown in FIG. 2. As a result of recombination, haplotypes at the beginning of the experiments are shuffled, and therefore genome-wide haplotype evolution cannot be tracked by sequencing pools of individuals (Pool-seq). Instead, investigators tend to track sliding window haplotype change over the course of the entire experiment as a function of genome position[75] as presented in the top panel. Population sizes are also typically much smaller than those in the systems shown in FIG. 2; as a result, the sampling variance in haplotype and allele frequency change is an important consideration. Recombination occurring in the course of the adaptation further shuffles the initial haplotypes, as shown in the middle panels of example haplotypes. The pattern of heterozygosity presented in the lower panel contrasts with the ones shown in FIG. 2. Heterozygosity remains globally high over the genome, apart from in the few regions harbouring the variants that are important for adaptation, which show reduced diversity (two cases of this are indicated by arrows). For details on the parameters used to generate the lower panels in this figure, see Supplementary information S1 (box).

families improved their binding affinity by several orders of magnitude on mutagenesis and reselection, suggesting that the initial sequences of these aptamers were relatively far from optimal and required subsequent newly arising mutations for optimal activity[35,77]. To date, the evidence from NGS of *in vitro*-selected pools suggests that individual families of functional RNAs exist in the starting pools but greatly benefit from mutagenesis[15,46,77] or synthetic shuffling[45] to uncover the most active variants, which then do not tend to drift towards other peaks in the fitness landscape[78].

In initially isogenic asexual systems, despite clonal interference increasing the time to fixation, genetic diversity within the populations remains low compared with in the other systems that we discuss in this Review. As these systems are initially isogenic, total variation in the population is limited to newly arising mutations. In a highly replicated E&R experiment in bacteria carried out for ~2,000 generations, the average number of genetic events distinguishing evolved clones from the ancestral strain was 11, with several lines of evidence suggesting that the majority of these events were adaptive[20,79] (see the parallel evolution section below).

Strikingly, these experiments suggest that, for virtually every novel environment, myriad beneficial mutations are accessible. Bacterial and yeast asexual E&R experiments starting from an isogenic population have recovered many different types of mutations in evolved lines. Although point mutations dominate in number[79–83], small insertions or deletions, large duplications or deletions, and transposition events also contribute to the selection response[79,84]. Most selected alleles seem to be the result of clear loss-of-function mutations, such as premature stop codons, gene deletions and transposable

Selective sweeps
When selection drives a genetic polymorphism to fixation, closely linked regions of the genome will follow along to fixation with the adaptive allele. The size of the swept region depends on the starting allele frequency of the beneficial allele, the strength of selection and the local recombination rate.

element insertions into genes[79,83,85]. Presumably the selective regime favours the loss of certain cellular functions. However, there are also clear examples of non-synonymous nucleotide changes being positively selected[64,79]. An interesting observation in bacteria is that, among the first large-effect adaptive steps, selection seems to recruit mutations in global regulators. For example, RNA polymerase (*rpo*) is a major mutational target in many different adaptations, ranging from high temperature to glycerol minimal media[86]. Several distinct novel *rpoBC* alleles are found in many different experimental evolution contexts, which is surprising given that this operon is generally highly conserved among bacterial species (FIG. 4a). It is unclear why changes in this operon are so often favoured in experimental evolution settings despite the function of these genes being highly conserved in nature.

In asexual yeast evolving from an isogenic starting population, large-scale duplications are often observed in response to environmental stresses[87–89], and similar genomic signatures are observed in extant populations[90–92]. Amplification of gene regions is thought to be a general mechanism for compensation of deleterious mutations[93] or adaptation to a limiting substrate. For example, laboratory-evolved yeast populations and synthetic yeast libraries harbouring chromosomal amplifications showed that increasing the number of copies of whole chromosomes or chromosomal regions has beneficial effects for adapting to limiting substrate conditions[94]. Another recent study evolved 180 strains with different single-gene deletions and observed compensatory evolution via duplicated regions in 22% of the initial strains[81]. One potential explanation for the ubiquity of this response is that the rate of mutation to aneuploidy or amplified gene regions is several orders of magnitude higher than that for point mutations[95]. Thus, selection may be favouring available, as opposed to optimal, solutions.

The few reports of E&R experiments on outbred microbial populations with standing genetic variation paint a different picture, in which selection acts mainly on standing genetic variation[53,54,75]. Large-scale structural and copy-number variants are not observed, apart from dramatic depletion of mitochondrial DNA copy number triggered by the accumulation of reactive oxygen species during heat stress[53]. There are several potential explanations for this apparent discrepancy. These include differences in experimental design that could lead to different rates of formation of complex rearrangements, different selective regimes, dominance in haploid versus diploid yeast, and the extent to which meiosis tolerates such events. Alternatively, large-effect structural mutations with deleterious pleiotropic side effects may be tolerated in experimental evolution experiments that opportunistically rely on beneficial newly arising mutations, whereas standing genetic variation may be preferred if it is available, as any standing variants are less likely to have strong deleterious pleiotropic side effects (as they are segregating in natural populations). Another possibility is that structural variants are occurring but are poorly queried via Pool-seq experiments. Sliding window plots of sequence coverage versus genome location suggest that large insertions and deletions are not present, but the Pool-seq data await more sophisticated analyses[55].

To date, most E&R experiments on *D. melanogaster* have been unable to pinpoint the causative regions responding to selection to single genes; instead, they typically identify several dozen genomic regions of hundreds of kilobases in size that show reductions in heterozygosity and/or frequency changes of SNPs[7,14,22,58,59]. Nevertheless, when adaptation is due to causative loci of large effect relative to standing variation, single-nucleotide resolution is achievable[60]. The usual failure to identify the genes harbouring causative variants in E&R experiments on *D. melanogaster* should not come as a surprise, as simulation studies show that for modest-effect causative alleles precise localization is possible only with higher levels of experimental replication, larger population sizes and/or more generations of evolution than are typically achieved[96,97]. More-highly replicated (number of experimental replicates >15), larger population size (effective population size >1,000) E&R experiments in *D. melanogaster* are possible and should allow for finer localization of beneficial alleles, but their execution requires considerable fortitude.

The relative importance of newly arising mutations versus standing genetic variation as contributors to adaptation is currently debated. Patterns of within-species variation and between-species divergence in humans suggest that standing variation is more important in adaptation than newly arising mutations[98], whereas in flies the opposite seems to be the case[99]; more generally, the role of newly arising mutations seems to be larger in species with large population sizes[100]. In E&R experiments initiated from an isogenic starting population, only newly arising mutations matter, so these systems cannot address this question. E&R experiments in *D. melanogaster* have concluded that most adaptation is due to standing variation[7,14,22,58–60], but the population sizes in these experiments are extremely small, a situation that favours standing variants[2]. Experiments in outbred yeast systems could begin to address the relative importance of newly arising mutations versus standing variation, but if the starting populations are derived from a small number of founders (as they have been in prior studies), rare standing alleles are not effectively captured. Furthermore, if the evolution itself is carried out without recombination, then the experiment is effectively sorting lineages as opposed to modelling evolution in sexually reproducing organisms. Despite these shortcomings, outbred yeast populations can be created with more founders, and evolution can be carried out in large populations with sex[75]. Thus, E&R methodology provides a direct experimental opportunity to address the importance of standing variation versus newly arising mutations in contributing to selection response.

Parallel evolution

Parallel evolution is said to have occurred if two lineages independently evolving from the same starting population converge on the same solution at some level of organization. Unlike studies of evolution in the wild,

Aneuploidy
Having an abnormal chromosome number owing to gain or loss of entire chromosomes.

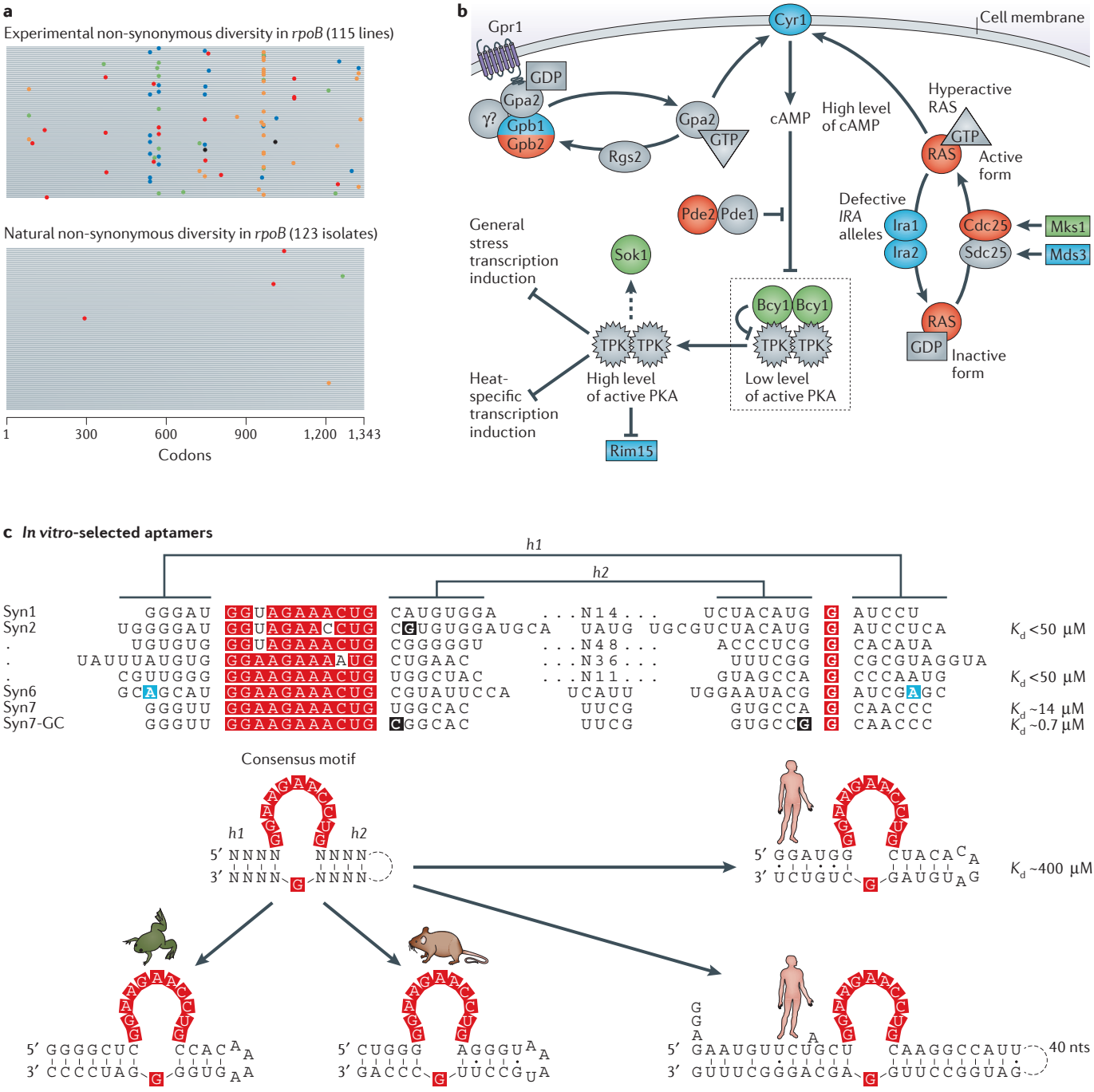


Figure 4 | **The molecular bases of adaptation.** **a** | RNA polymerase B (*rpoB*) sequence alignments for replicate populations (rows) evolved at high temperature[79] (top panel) and those in naturally occurring isolates[148] (bottom panel) show that *rpoB* is targeted repeatedly in laboratory-based adaptation to high temperature, but is largely invariant between naturally occurring *Escherichia coli* strains. Colours correspond to different base changes (A, green; C, blue; G, orange; T, red; and deletions, black). **b** | The RAS–cyclic AMP pathway is targeted by both *de novo* mutations and standing variants involved in yeast adaptation to multiple stress conditions. Colours indicate alleles detected in experiments initiated from a single clone (red), an outbred synthetic population (green) or both (blue). **c** | Adenosine aptamer sequences recovered from an *in vitro* selection experiment and subsequently optimized for strong binding[116]. The pattern of sequence evolution is different for the ligand-binding loop (red) and the helical segments (marked as *h1* and *h2*), which are defined solely by sequence covariation. A structure-based search for the identified motif and genomic systematic evolution of ligands by exponential enrichment (SELEX) for novel adenosine or ATP aptamers revealed the adenosine aptamer sequences in the bullfrog, mouse and human, suggesting molecular convergence between *in vitro*-evolved molecules and genomic sequences[120]. Bcy1, bypass of cyclic AMP requirement 1; Cdc25, cell division cycle 25; Cyr1, cyclic AMP requirement 1; Gpa2, G protein subunit α2; Gpb, G protein subunit-β; Gpr1, G protein-coupled receptor 1; Ira, inhibitory regulator of the RAS–cAMP pathway; K_d, dissociation constant; Mds3, Mck1 dosage suppressor 3; Mks1, multicopy kinase suppressor 1; Pde, phosphodiesterase; PKA, protein kinase A; Rgs2, regulator of heterotrimeric G protein signalling 2; Rim15, regulator of Ime2; Sdc25, suppressor of Cdc25; Sok1, suppressor of kinase 1; TPK, Takashi's protein kinase. Part **b** is adapted with permission from REF. 53, Cold Spring Harbor Laboratory Press.

E&R experiments are able to study multiple evolving lineages under identical selection conditions, providing a powerful opportunity to directly detect parallel evolution. Knowing the relative prevalence of this event will help to address a long-standing question regarding the replicability of evolution at the molecular level. Parallel evolution in independent lineages can also be used as a tool to identify functionally relevant changes[16,79,101,102] and distinguish them from non-selected passenger mutations that may be found in the evolved genomes.

When evolution is initiated from an isogenic starting population, the observation of similar changes recovered in independent lineages suggests either a high mutation rate or the filtering action of natural selection. For experimental evolution of asexually evolving isogenic yeast or bacteria, in most cases and at most loci, the high-mutation-rate hypothesis can be rejected, as replicate populations have different mutations affecting the same target. Indeed, in most targeted genes, numerous adaptive mutations seem to exist[79,103,104], even if the gene is essential (as seen in the *rpoB* example shown in FIG. 4a). Consequently, parallelism is rarely observed at the level of individual mutations[79,105], but instead is seen at the gene level and often at higher levels of functional integration, such as the operon or sets of functionally related genes (for example, genes involved in cell shape)[79,103,104]. The genomic precision of targeting seems to depend on the selective regime used, including the media used for selection[89] as well as the mode of culture (chemostats versus batch culture)[106]. For example, antibiotic treatments select almost exclusively for mutations in the active site of specific target genes[107], whereas for less-specific environmental stressors in highly replicated E&R experiments many genes and alternative mutations among genes are recovered[64,79]. With less-specific environmental stressors, most genes are hit in less than half of the evolved replicates. These observations suggest that there are perhaps thousands of mutations available that improve fitness in some novel environment, and that the particular set used in any given evolutionary realization depends on the particular mutations that occur early on in the experiment, the genetic backgrounds they occur on and the action of clonal interference.

Despite the stochastic nature of evolutionary change in asexual systems initiated from an isogenic starting population, there is enough parallelism such that many functional targets have been identified. The functions identified to date as targets of selection are very diverse, ranging from genes important in metabolism[18,89] to genes involved in cell shape[20,79] and the stress response[104,108,109], as well as genes of unknown function[79]. Interestingly, mutations in global regulators are often observed to be of large effect[110], to be beneficial and to occur early in the experimental evolution experiment, and they may even be recovered across different laboratory-induced selective regimes[111], with different alleles being discovered in different settings[112]. For example, in yeast, the RAS–cyclic AMP signalling pathway[64,66,85,113] is a target of selection both in experimental evolution initiated from an isogenic population as well as in experiments initiated from a synthetic outbred population, in which the pathway seems to harbour multiple deleterious SNPs[53] (FIG. 4b).

The observation of convergence can be equally as illuminating when selection is initiated from an outbred population potentially harbouring millions of polymorphic sites, as only a very small fraction is likely to be the target of selection and hence show convergent change across replicate populations. Indeed, modelling experiments of evolution initiated from an outbred population suggest that replication is perhaps the easiest way to increase the power to detect causative sites[96,97]. In contrast to in microbial systems, in which parallelism at the base-pair level is almost non-existent and gene-level parallelism is detectable but modest, in systems with standing genetic variation the degree of parallelism is high. In published replicated experiments in yeast and *D. melanogaster*, convergence is almost always observed. The probable explanation for this result is perhaps obvious: outbred sexual systems start with the same standing variation available for natural selection to act on in independent replicates, the total number of pre-existing beneficial alleles of modest-to-large effect is somewhat limited, and these same alleles are targeted in replicate populations. Sexual recombination results in an absence of clonal interference, which means that unlinked beneficial alleles can independently and simultaneously increase in frequency.

A caveat in the sexual systems is that even small amounts of accidental gene flow between replicate populations will virtually guarantee parallel allele frequency change across replicate populations, with allele frequency drift in the metapopulation masquerading as parallel adaptive change[22]. Population genetics theory predicts that exchanging one migrant per generation between two populations is sufficient enough to homogenize neutral allele frequencies[114]. Because replicate populations are generally not marked in any way, it is difficult to know whether low levels of gene flow are an issue, and future experiments would benefit from DNA barcoding of replicate populations. In outbred yeast, this could be accomplished using barcodes that are identical for every individual within a population but different between populations. A simple PCR-multiplexed NGS reaction could detect contamination events with some sensitivity (such a barcoding system is described in REF. 71). However, in a system that uses more-complex organisms, such as *D. melanogaster*, it is technically more challenging to easily barcode an outbred population.

Although parallel evolution dominates the landscape in the outbred systems, this is a novel and recent observation, and one that would not necessarily have been predicted a priori. If selection tends to operate on rare alleles present in the starting population, the adaptive response may not be so replicable, as the probability of initially rare alleles being stochastically lost in a given replicate evolving population is ~1–2s (BOX 1). By contrast, if an allele starts out at a frequency of >5%, the probability of fixation is almost certain provided

$s >> 1/N$ (where N is the population size). Taking these two theoretical considerations into account, the observation of highly parallel evolution in yeast synthetic outbred E&R experiments should not be surprising at all. As the synthetic populations are derived from four isogenic founders, all alleles start out at a frequency of roughly one-quarter, one-half or three-quarters, and because populations sizes are $>10^6$, the loss of beneficial alleles is highly unlikely. Convergence is more surprising in *D. melanogaster*, in which minor allele frequencies in the starting population can be less than 1%, and population sizes are smaller (or much smaller) than 1,000 individuals. If gene flow is absent in E&R experiments in *D. melanogaster*, the observation of strong parallel evolution suggests that the targeted beneficial alleles start at minor allele frequencies of >5% and that selection is fairly strong. However, we note that studies in outbred systems (particularly those in flies) often have few replicates and are thus typically too underpowered to confidently distinguish beneficial from neutral alleles based on parallelism across replicates. Thus, apparent evidence of parallelism should be interpreted cautiously, and E&R studies in outbred yeast and flies should ultimately aim for highly replicated formats analogous to those in bacteria[80].

Beyond parallel evolution at the primary sequence or higher functional level, parallelism is also apparent when considering molecular structures. *In vitro* evolution experiments suggest that chemically functional RNAs (or DNAs and their analogues) fold into specific structures, often dominated by secondary structures, and it is usually these structures, rather than specific sequences, that are selected for. Hence, independent populations may converge towards a secondary structure that may be revealed by sequence covariation (for example, G–C versus C–G base pairs) but not primary sequence conservation, whereas key tertiary interactions and active and/or binding sites tend to be conserved on a primary-sequence level (FIG. 4c). Thus, genotypes that evolve are dominated by strong sequence conservation in key loops, surrounded by strongly co-varying (but generally sequence-independent) helical structures. The information content of co-varying segments is therefore higher than the primary sequence would suggest[77]; as a result, the degree to which the evolution of functional RNAs shows convergence may not be fully appreciated. Nevertheless, strong functional convergence is observed across experiments. For example, *in vitro* selection starting from a short, 24-nucleotide pool that extensively sampled the theoretical sequence space yielded the same GTP aptamers in two independent experiments[115]. Functional convergence can even be extended across biological systems: for instance, the same adenosine aptamer motif has been independently selected *in vitro* from random sequences at least four times[116–119] and occurs in genomes spanning various organisms, from bacteria to humans[120]. Similarly, the hammerhead ribozyme, a specific type of self-cleaving ribozyme with a conserved catalytic core and secondary structure, has been identified *in vitro* several times[121] and is widespread in nature[122–124].

Tertiary interactions
Molecular interactions stabilizing the overall (tertiary) structure of a functional RNA.

Epistasis

Epistasis measures the extent to which allelic effects depend on the genetic background in which they appear. It is a property of the adaptive landscape that conditions the dynamics of adaptation. Epistasis can be defined at a microscopic level, as a functional interaction between alleles such that the fitness of a double mutant differs from that expected based on the combined fitness effects of the single mutants, or at a macroscopic level, where the fitness of a mutant depends on some higher-level property of the overall genetic background it occurs on. An emerging question is whether robust statistical properties may be apparent at a higher macroscopic level[24] when all microscopic epistatic interactions are accounted for. If rules exist at the macroscopic epistasis level, long-term predictions on the adaptive process may be possible despite a poor knowledge of the details of epistasis at the microscopic level[24]. What have E&R experiments told us about microscopic and macroscopic epistasis?

Microscopic epistasis. As discussed above, *in vitro* evolution experiments select RNA molecules with sequence-conserved loops flanked by strongly co-varying but generally sequence-independent helical structures (FIG. 4c). The epistatic interactions among the partners of the helical structures can be identified using NGS reads: covariation is extracted from the entire population-wide haplotype of the adapting 30–100-mers-long sequence. It is worth noting that these epistatic interactions are easily detected in this system because the distances between interacting sites in the linear nucleotide chain are typically shorter than the sequencing read lengths, so interaction sites can be collectively analysed in single sequencing reads. It is important to appreciate that all evolutionary change occurring in the stem part of a ribozyme or an aptamer is epistatic in nature; thus, epistasis is pervasive in this system[77,78,116].

E&R experiments also reveal microscopic epistasis among beneficial mutations in asexual isogenic systems. For instance, when a particular gene, pathway or function is targeted by some mutation, especially a loss-of-function mutation, little advantage is gained from additional mutations in the same gene, pathway or function. The old adage that 'there is little use in beating a dead horse' comes to mind. Moreover, populations evolving more than a few hundred generations will eventually fix multiple beneficial mutations contributing to adaptation. Provided that enough E&R replicates have been carried out, order-of-fixation epistasis can be detected as non-random orders of fixation events in replicate-evolved populations. Tenaillon *et al.*[79] observed many significant epistatic interactions between independent functional targets of selection. The strongest such example was that early substitution events in *rpoB* tended to preclude later events in *rho* (and vice versa). Other studies coupling E&R methods with experiments have revealed many cases of mutations showing sign epistasis, which is when a mutation is beneficial in the ancestral background but deleterious when coupled with another mutation[125–128].

Despite the evidence for pervasive epistasis in *in vitro* systems and in microbial systems initiated from an isogenic population, there is little evidence for epistasis in systems in which experimental evolution is initiated from an outbred population. This might mean that epistasis is less important in sexual systems, as any favourable allele increasing in frequency with time will find itself in numerous different genetic backgrounds, whereas in asexual systems a newly arising mutation is always associated with the genome-wide haplotype on which it arises. Thus, a beneficial allele in an asexual system needs to be advantageous only in the background on which it arises, whereas the advantage of an allele in a sexual population is its average effect over all backgrounds it is likely to encounter. It follows that, in sexual systems, natural selection will favour only those alleles that combine beneficially across all backgrounds, whereas in asexual systems natural selection will favour alleles that are beneficial only in the background in which they occur. Contrarily, epistasis may be commonplace but rarely observed in experiments initiated from outbred populations owing to it being difficult to detect, given the way experiments are carried out in these systems. The easiest way to detect fitness epistasis in an outbred population is to identify unlinked loci in linkage disequilibrium[129]. E&R experiments from an outbred starting population prepare genomic DNA from an entire population, sequence it as a pool and then estimate the frequency of every SNP in the genome as a function of treatment and/or time. Given that these data report population-wide allele frequencies but not which alleles co-occur in individuals, linkage disequilibrium between unlinked loci is not estimated, so it is very difficult to show that particular combinations of alleles are more favoured than others. Nevertheless, at least in the yeast system, when causal sites have been identified, gene-replacement experiments could be used to shed light on the problem.

Macroscopic epistasis. In all the different systems, there seems to be a change in the dynamics of adaptation over time. For *in vitro* E&R experiments, after the early stages of fast adaptation, further selective cycles result in diminishing returns with respect to biochemical activity in the assay[130]. Similarly, in isogenic asexual populations, rates of adaptation slow over the course of the experiment[52]. The explanation for diminishing returns over the course of an experimental evolution experiment has often been based on Fisher's geometric model of adaptation[131]. Under this model, as an evolving population approaches a new optimum in phenotypic space, the beneficial alleles favoured by natural selection are both less common and of smaller selective effect[131,132]. According to the model, large-effect mutations, which may also have deleterious pleiotropic side effects, are favoured in the initial phases of evolution, but as the optimum is approached fine-tuning mutations are instead favoured[133]. In yeast, allele-replacement techniques are straightforward, and they have revealed a form of macroscopic epistasis in which the effect of a beneficial mutation is a function of the fitness of the clone into which it was inserted, as opposed to the specific combination of other mutations present in that clone[82]. The higher the initial fitness of the recipient background, the lower the incremental effect of the introduced beneficial mutation on fitness[134]. This overall diminishing rate of adaptation as fitness increases seems to be conserved from viral systems through to yeast[81,82,135–139] and is consistent with the analysis of epistasis among combinations of beneficial mutations[70,140–142]. An interesting remaining question is: to what extent does this observation apply to evolution in outbred sexual populations?

Linkage disequilibrium
The condition in which the frequency of a particular haplotype for two loci is significantly different from that expected if the loci were assorting independently.

Conclusions and future perspectives

Despite the vastly different complexity of different molecules and organisms, ranging from 30 bp oligonucleotides transcribed into RNA to complex multicellular higher eukaryotes, the principles governing their evolution are constant. The large differences we see in how quickly these systems respond to environmental challenges and the molecular bases of adaptation are largely outcomes of the nature of the starting populations used and the intensity of natural selection experienced in each system. Thus, each population is exploring its adaptive landscape in a slightly different manner, and the exploratory process governs evolutionary outcomes. In asexual systems, clonal interference results in genome-wide haplotypes competing with one another for eventual fixation, whereas in sexual systems different genomic regions can respond to evolution somewhat independently. In systems starting from standing variation, adaptation can be much more deterministic, as the same range of alleles is available to all replicate populations. By contrast, in systems starting from a single isogenic population, all evolution must proceed from newly arising mutations; thus, it is somewhat predicated on the stochastic nature and timing of these events. Even when evolution is initiated from populations with standing variation, it will proceed differently if that starting variation is in randomly synthesized oligonucleotides (in which the vast majority of alleles have an initial fitness of nearly zero) compared with if the starting variation is naturally occurring, in which case the variance in fitness among starting alleles is probably much more subtle. But the role of very strong and/or molecularly targeted selection cannot be discounted. In bacteria, the nature of evolutionary response can be very different when the selection pressure is relatively specific and strong (for example, antibiotic resistance) versus more weak but pervasive (for example, high temperature). In *in vitro* systems, the selective agent directly queries a specify biochemical activity, so it is difficult to imagine a system that consistently experiences a more-specific selective pressure.

There is still a rather large disconnect between the communities who believe the principles of evolutionary change can be completely elucidated via experimental evolution in the laboratory and those who believe that the rules in nature are somehow different. The often-observed phenomenon of mutations in global regulators (such as *rpoB*) being important in laboratory microbial evolution, despite these proteins being highly conserved

over evolutionary time, supports the idea that evolution in the laboratory does not fully recapitulate what occurs in the wild. Nevertheless, the controlled and replicated nature of laboratory experiments is an extremely appealing feature of the E&R approach. In a few cases, replicated evolution has certainly occurred in nature, and these systems can yield incredible insights[143], but it is difficult to routinely identify such 'natural experiments'. Another avenue might be to introduce isogenic bacteria to replicate sets of more-natural environments[144] than chemostats or Erlenmeyer flasks, or to introduce outbred laboratory populations back into natural replicated environments.

The overall comparability between experimental evolution systems, after the underlying population genetics and nature of the adaptive landscape are taken into account, belies some nagging differences. Both microscopic epistasis and macroscopic epistasis seem to be pervasive in all of the lower-complexity systems, but epistasis seems to be much-less common in the outbred, sexual, higher-complexity systems. An important future task is to determine whether this difference is a feature of the systems or is imposed by different experimental designs. Individual-based sequencing seems to be cost-prohibitive in these systems, but epistasis could be studied by gene-replacement experiments. A second, somewhat striking, finding is that larger-scale structural variants often play a part in adaptation in asexual evolution initiated from an outbred population but that these same events are not generally observed in sexual systems initiated from an outbred population. Future experiments should address whether these observations are replicable in systems in which recombination is occurring during evolution (currently, it is more difficult to evolve these systems for as many generations). Finally, it is fascinating that the field of *in vitro* E&R has 'evolved' largely independently of that of the *in vivo* systems. Given the increasing desire to understand E&R experiments within a systems biology framework, the distinction between *in vivo* and *in vitro* systems seems to be increasingly arbitrary. The time is ripe for practitioners in the different systems to learn from one another.

1. Chevin, L. M. & Hospital, F. Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics* **180**, 1645–1660 (2008).
2. Hermisson, J. & Pennings, P. S. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352 (2005).
3. Kaplan, N. L., Hudson, R. R. & Langley, C. H. The 'hitchhiking effect' revisited. *Genetics* **123**, 887–899 (1989).
4. Sellis, D., Callahan, B. J., Petrov, D. A. & Messer, P. W. Heterozygote advantage as a natural consequence of adaptation in diploids. *Proc. Natl Acad. Sci. USA* **108**, 20666–20671 (2011).
5. Adams, J. & Rosenzweig, F. Experimental microbial evolution: history and conceptual underpinnings. *Genomics* **104**, 393–398 (2014).
6. Elena, S. F. & Lenski, R. E. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* **4**, 457–469 (2003).
7. Zhou, D. *et al.* Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **108**, 2349–2354 (2011).
8. Cassidy, J. J. *et al.* miR-9a minimizes the phenotypic impact of genomic diversity by buffering a transcription factor. *Cell* **155**, 1556–1567 (2013).
9. Soskine, M. & Tawfik, D. S. Mutational effects and the evolution of new protein functions. *Nat. Rev. Genet.* **11**, 572–582 (2010).
10. Wilson, D. S. & Szostak, J. W. *In vitro* selection of functional nucleic acids. *Annu. Rev. Biochem.* **68**, 611–647 (1999).
11. Joyce, G. F. Forty years of *in vitro* evolution. *Angew. Chem. Int. Ed. Engl.* **46**, 6420–6436 (2007).
12. Rose, M. R. & Lauder, G. V. *Adaptation* (Academic Press, 1996).
13. Kawecki, T. J. *et al.* Experimental evolution. *Trends Ecol. Evol.* **27**, 547–560 (2012).
14. Turner, T. L., Stewart, A. D., Fields, A. T., Rice, W. R. & Tarone, A. M. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet.* **7**, e1001336 (2011).
This study describes an E&R experiment in *D. melanogaster* and introduces the term 'evolve and resequence'.
15. Pitt, J. N. & Ferré-D'Amaré, A. R. Rapid construction of empirical RNA fitness landscapes. *Science* **330**, 376–379 (2010).
The first study using NGS to measure the shape of the fitness landscape of an *in vitro*-selected ribozyme and explain the mutations within the context of the ribozyme structure.
16. 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).
17. Shendure, J. *et al.* Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* **309**, 1728–1732 (2005).
18. Herring, C. D. *et al.* Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat. Genet.* **38**, 1406–1412 (2006).
19. 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).
20. Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**, 1243–1247 (2009).
A study that details the sequencing of *E. coli* lineages throughout 30,000 generations of evolution and is the first paper to link molecular evolution to fitness improvement.
21. Araya, C. L., Payen, C., Dunham, M. J. & Fields, S. Whole-genome sequencing of a laboratory-evolved yeast strain. *BMC Genomics* **11**, 88 (2010).
22. Burke, M. K. *et al.* Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* **467**, 587–590 (2010).
The first E&R experiment in an obligate sexual higher eukaryote. Despite several hundred generations of evolution, fixation events are not observed.
23. Illingworth, C. J. R., Parts, L., Schiffels, S., Liti, G. & Mustonen, V. Quantifying selection acting on a complex trait using allele frequency time series data. *Mol. Biol. Evol.* **29**, 1187–1197 (2012).
24. Good, B. H. & Desai, M. M. The impact of macroscopic epistasis on long-term evolutionary dynamics. *Genetics* **199**, 177–190 (2015).
25. Martin, A. & Orgogozo, V. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67**, 1235–1250 (2013).
26. Hoekstra, H. E. & Coyne, J. A. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995–1016 (2007).
27. Hietpas, R. T., Jensen, J. D. & Bolon, D. N. A. Experimental illumination of a fitness landscape. *Proc. Natl Acad. Sci. USA* **108**, 7896–7901 (2011).
28. Firnberg, E. & Ostermeier, M. PFunkel: efficient, expansive, user-defined mutagenesis. *PLoS ONE* **7**, e52031 (2012).
29. Stiffler, M. A., Hekstra, D. R. & Ranganathan, R. Evolvability as a function of purifying selection in TEM-1 β-lactamase. **160**, 882–892 (2015).
30. Slattery, M. *et al.* Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. **147**, 1270–1282 (2011).
31. Gu, G., Wang, T., Yang, Y., Xu, X. & Wang, J. An improved SELEX-Seq strategy for characterizing DNA-binding specificity of transcription factor: NF-κB as an example. *PLoS ONE* **8**, e76109 (2013).
32. Ellington, A. D. & Szostak, J. W. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818–822 (1990).
33. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
34. Robertson, D. L. & Joyce, G. F. Selection *in vitro* of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* **344**, 467–468 (1990).
35. Davis, J. H. & Szostak, J. W. Isolation of high-affinity GTP aptamers from partially structured RNA libraries. *Proc. Natl Acad. Sci. USA* **99**, 11616–11621 (2002).
36. Ellington, A. D. & Szostak, J. W. Selection *in vitro* of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature* **355**, 850–852 (1992).
37. Breaker, R. R. & Joyce, G. F. A. DNA enzyme that cleaves RNA. *Chem. Biol.* **1**, 223–229 (1994).
38. Trevino, S. G., Zhang, N., Elenko, M. P., Lupták, A. & Szostak, J. W. Evolution of functional nucleic acids in the presence of nonheritable backbone heterogeneity. *Proc. Natl Acad. Sci. USA* **108**, 13492–13497 (2011).
39. Yu, H., Zhang, S. & Chaput, J. C. Darwinian evolution of an alternative genetic system provides support for TNA as an RNA progenitor. *Nat. Chem.* **4**, 183–187 (2012).
40. Taylor, A. I. *et al.* Catalysts from synthetic genetic polymers. *Nature* **518**, 427–430 (2015).
41. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the *in vitro* selection of peptides and proteins. *Proc. Natl Acad. Sci. USA* **94**, 12297–12302 (1997).
42. Seelig, B. & Szostak, J. W. Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* **448**, 828–831 (2007).
43. Bartel, D. P. & Szostak, J. W. Isolation of new ribozymes from a large pool of random sequences. *Science* **261**, 1411–1418 (1993).
44. Kobori, S., Nomura, Y., Miu, A. & Yokobayashi, Y. High-throughput assay and engineering of self-cleaving ribozymes by sequencing. *Nucleic Acids Res.* **43**, e85 (2015).
45. Curtis, E. A. & Bartel, D. P. Synthetic shuffling and *in vitro* selection reveal the rugged adaptive fitness landscape of a kinase ribozyme. *RNA* **19**, 1116–1128 (2013).
This study describes NGS measurement of the fitness landscape of a kinase ribozyme, with explicit comparison of recombination versus point mutations in an RNA population.
46. Ameta, S., Winz, M.-L., Previti, C. & Jäschke, A. Next-generation sequencing reveals how RNA catalysts evolve from random space. *Nucleic Acids Res.* **42**, 1303–1310 (2014).

47. Hayden, E. J., Bratulic, S., Koenig, I., Ferrada, E. & Wagner, A. The effects of stabilizing and directional selection on phenotypic and genotypic variation in a population of RNA enzymes. *J. Mol. Evol.* **78**, 101–108 (2014).
48. Cho, M. *et al.* Quantitative selection and parallel characterization of aptamers. *Proc. Natl Acad. Sci. USA* **110**, 18460–18465 (2013).
49. Buenrostro, J. D. *et al.* Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nat. Biotech.* **32**, 562–568 (2014).
50. Atwood, K. C., Schneider, L. K. & Ryan, F. J. Periodic selection in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **37**, 146–155 (1951).
51. Novick, A. & Szilard, L. Experiments with the Chemostat on spontaneous mutations of bacteria. *Proc. Natl Acad. Sci. USA* **36**, 708–719 (1950).
52. Wiser, M. J., Ribeck, N. & Lenski, R. E. Long-term dynamics of adaptation in asexual populations. *Science* **342**, 1364–1367 (2013).
53. Parts, L. *et al.* Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Res.* **21**, 1131–1138 (2011).
This paper introduces the yeast outbred population system to the E&R field.
54. Cubillos, F. A. *et al.* High-resolution mapping of complex traits with a four-parent advanced intercross yeast population. *Genetics* **195**, 1141–1155 (2013).
55. Schlötterer, C., Tobler, R., Kofler, R. & Nolte, V. Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. *Nat. Rev. Genet.* **15**, 749–763 (2014).
56. L'Héritier, P. L., Neefs, Y. & Teissier, G. Apterisme de insects et selection naturelle. *Compt. Rend. Acad. Sci.* **204**, 907–909 (1937) (in French).
57. Powell, J. R. *Progress and Prospects in Evolutionary Biology: The Drosophila Model* (Oxford University Press, 1997).
58. Orozco-terWengel, P. *et al.* Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Mol. Ecol.* **21**, 4931–4941 (2012).
59. Teotônio, H., Chelo, I. M., Bradič, M., Rose, M. R. & Long, A. D. Experimental evolution reveals natural selection on standing genetic variation. *Nat. Genet.* **41**, 251–257 (2009).
60. Martins, N. E. *et al.* Host adaptation to viruses relies on few genes with different cross-resistance properties. *Proc. Natl Acad. Sci. USA* **111**, 5938–5943 (2014).
61. Teotônio, H. & Rose, M. R. Variation in the reversibility of evolution. *Nature* **408**, 463–466 (2000).
62. Passananti, H. B., Matos, M. & Rose, M. R. *Methuselah Flies: A Case Study in the Evolution of Aging* (World Scientific Publishing Company, 2004).
63. Gerrish, P. J. & Lenski, R. E. The fate of competing beneficial mutations in an asexual population. *Genetica* **102–103**, 127–144 (1998).
64. Lang, G. I. *et al.* Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500**, 571–574 (2013).
The sequencing of tens of yeast populations through time reveals the dynamics of clonal interference, in which groups of mutations compete with one another until one combination, possibly including neutral 'hitchhiker' mutations, reaches fixation.
65. Desai, M. M. & Fisher, D. S. Beneficial mutation selection balance and the effect of linkage on positive selection. *Genetics* **176**, 1759–1798 (2007).
66. Kao, K. C. & Sherlock, G. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat. Genet.* **40**, 1499–1504 (2008).
67. Lang, G. I., Botstein, D. & Desai, M. M. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* **188**, 647–661 (2011).
68. Rodríguez-Verdugo, A., Carrillo-Cisneros, D., González-González, A., Gaut, B. S. & Bennett, A. F. Different tradeoffs result from alternate genetic adaptations to a common environment. *Proc. Natl Acad. Sci. USA* **111**, 12121–12126 (2014).
69. Toprak, E. *et al.* Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat. Genet.* **44**, 101–105 (2012).
70. Khan, A. I. A., Dinh, D. M. D., Schneider, D. D., Lenski, R. E. R. & Cooper, T. F. T. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* **332**, 1193–1196 (2011).
71. Levy, S. F. *et al.* Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* **519**, 181–186 (2015).
Adding molecular markers to 500,000 yeast cells and following their frequency through time allowed the authors to uncover the diversity of beneficial mutations that are simultaneously present in a yeast population.
72. Ehrenreich, I. M. *et al.* Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature* **464**, 1039–1042 (2011).
73. Zeyl, C. & Bell, G. The advantage of sex in evolving yeast populations. *Nature* **388**, 465–468 (1997).
74. Goddard, M. R., Godfray, H. C. J. & Burt, A. Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434**, 636–640 (2005).
75. Burke, M. K., Liti, G. & Long, A. D. Standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. *Mol. Biol. Evol.* **31**, 3228–3239 (2014).
This study describes a yeast population, derived from four isogenic founders, that was experimentally evolved with forced sexual recombination roughly once every 30 cell divisions.
76. Burke, M. K. & Long, A. D. What paths do advantageous alleles take during short-term evolutionary change? *Mol. Ecol.* **21**, 4913–4916 (2012).
77. Carothers, J. M., Oestreich, S. C., Davis, J. H. & Szostak, J. W. Informational complexity and functional activity of RNA structures. *J. Am. Chem. Soc.* **126**, 5130–5137 (2004).
78. Petrie, K. L. & Joyce, G. F. Limits of neutral drift: lessons from the *in vitro* evolution of two ribozymes. *J. Mol. Evol.* **79**, 75–90 (2014).
79. Tenaillon, O. *et al.* The molecular diversity of adaptive convergence. *Science* **335**, 457–461 (2012).
The parallel adaptation of 115 lineages of *E. coli* reveals that convergence among replicates occurs mostly at a functional level and that many competing mutations may affect each of the functional targets of adaptation.
80. Blank, D., Wolf, L., Ackermann, M. & Silander, O. K. The predictability of molecular evolution during functional innovation. *Proc. Natl Acad. Sci. USA* **111**, 3044–3049 (2014).
81. Szamecz, B. *et al.* The genomic landscape of compensatory evolution. *PLoS Biol.* **12**, e1001935 (2014).
82. Kryazhimskiy, S., Rice, D. P., Jerison, E. R. & Desai, M. M. Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* **344**, 1519–1522 (2014).
The existence of macroscopic epistasis leading to a global diminishing return rate of adaptation as a function of fitness is shown using yeast.
83. Payen, C. *et al.* Empirical determinants of adaptive mutations in yeast experimental evolution. *bioRxiv* http://dx.doi.org/10.1101/014068 (2015).
84. Raeside, C. *et al.* Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. *mBio* **5**, e01377-14 (2014).
85. Kvitek, D. J. & Sherlock, G. Whole genome, whole population sequencing reveals that loss of signaling networks is the major adaptive strategy in a constant environment. *PLoS Genet.* **9**, e1003972 (2013).
86. Conrad, T. M. *et al.* RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc. Natl Acad. Sci. USA* **107**, 20500–20505 (2010).
87. Di Rienzi, S. C. *et al.* The dynamics of diverse segmental amplifications in populations of *Saccharomyces cerevisiae* adapting to strong selection. *G3 (Bethesda)* **4**, 399–409 (2014).
88. Chang, S.-L., Lai, H.-Y., Tung, S.-Y. & Leu, J.-Y. Dynamic large-scale chromosomal rearrangements fuel rapid adaptation in yeast populations. *PLoS Genet.* **9**, e1003232 (2013).
89. Gresham, D. *et al.* The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet.* **4**, e1000303 (2008).
90. Bergström, A. *et al.* A high-definition view of functional genetic variation from natural yeast genomes. *Mol. Biol. Evol.* **31**, 872–888 (2014).
91. Strope, P. K. *et al.* The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res.* **25**, 762–774 (2015).
92. Warringer, J. *et al.* Trait variation in yeast is defined by population history. *PLoS Genet.* **7**, e1002111 (2011).
93. Koszul, R., Caburet, S., Dujon, B. & Fischer, G. Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments. *EMBO J.* **23**, 234–243 (2004).
94. Sunshine, A. B. *et al.* The fitness consequences of aneuploidy are driven by condition-dependent gene effects. *PLoS Biol.* **13**, e1002155 (2015).
This study introduces a new methodology to dissect the underlying molecular basis of aneuploidy advantages.
95. Dujon, B. Yeast evolutionary genomics. *Nat. Rev. Genet.* **11**, 512–524 (2010).
96. Baldwin-Brown, J. G., Long, A. D. & Thornton, K. R. The power to detect quantitative trait loci using resequenced, experimentally evolved populations of diploid, sexual organisms. *Mol. Biol. Evol.* **31**, 1040–1055 (2014).
This is a simulation study that examined the power of E&R experiments in sexual outbred populations. A conclusion of the paper is that current experiments are too underpowered to routinely identify regions responding to selection.
97. Kofler, R. & Schlötterer, C. A guide for the design of evolve and resequencing studies. *Mol. Biol. Evol.* **31**, 474–483 (2014).
98. Hernandez, R. D. *et al.* Classic selective sweeps were rare in recent human evolution. *Science* **331**, 920–924 (2011).
99. Sattath, S., Elyashiv, E., Kolodny, O., Rinott, Y. & Sella, G. Pervasive adaptive protein evolution apparent in diversity patterns around amino acid substitutions in *Drosophila* simulans. *PLoS Genet.* **7**, e1001302 (2011).
100. Corbett-Detig, R. B., Hartl, D. L. & Sackton, T. B. Natural selection constrains neutral diversity across a wide range of species. *PLoS Biol.* **13**, e1002112 (2015).
101. Cooper, T. F., Rozen, D. E. & Lenski, R. E. Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **100**, 1072–1077 (2003).
102. Pelosi, L. *et al.* Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. *Genetics* **173**, 1851–1869 (2006).
103. Lee, M.-C. & Marx, C. J. Synchronous waves of failed soft sweeps in the laboratory: remarkably rampant clonal interference of alleles at a single locus. *Genetics* **193**, 943–952 (2013).
104. Maharjan, R. P., Liu, B., Feng, L., Ferenci, T. & Wang, L. Simple phenotypic sweeps hide complex genetic changes in populations. *Genome Biol. Evol.* **7**, 531–544 (2015).
105. Ostrowski, E. A., Woods, R. J. & Lenski, R. E. The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. *Proc. Biol. Sci.* **275**, 277–284 (2008).
106. Gresham, D. & Dunham, M. J. The enduring utility of continuous culturing in experimental evolution. **104**, 399–405 (2014).
107. Lindsey, H. A., Gallie, J., Taylor, S. & Kerr, B. Evolutionary rescue from extinction is contingent on a lower rate of environmental change. *Nature* **494**, 463–467 (2013).
108. Riehle, M. M., Bennett, A. F. & Long, A. D. Genetic architecture of thermal adaptation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **98**, 525–530 (2001).
109. Charusanti, P. *et al.* Genetic basis of growth adaptation of *Escherichia coli* after deletion of pgi, a major metabolic gene. *PLoS Genet.* **6**, e1001186 (2010).
110. Hindré, T., Knibbe, C., Beslon, G. & Schneider, D. New insights into bacterial adaptation through *in vivo* and *in silico* experimental evolution. *Nat. Rev. Microbiol.* **10**, 352–365 (2012).
111. Conrad, T. M., Lewis, N. E. & Palsson, B. Ø. Microbial laboratory evolution in the era of genome-scale science. *Mol. Syst. Biol.* **7**, 509 (2011).
112. Achaz, G., Rodríguez-Verdugo, A., Gaut, B. S. & Tenaillon, O. The reproducibility of adaptation in the light of experimental evolution with whole genome sequencing. *Adv. Exp. Med. Biol.* **781**, 211–231 (2014).
113. Wenger, J. W. *et al.* Hunger artists: yeast adapted to carbon limitation show trade-offs under carbon sufficiency. *PLoS Genet.* **7**, e1002202 (2011).
114. Hartl, D. L. & Clark, A. G. *Principles of Population Genetics* (Sinauer Associates Incorporated, 2007).
A population genetics textbook that should be read as often as it is cited.

115. Jiménez, J. I., Xulvi-Brunet, R., Campbell, G. W., Turk-MacLeod, R. & Chen, I. A. Comprehensive experimental fitness landscape and evolutionary network for small RNA. *Proc. Natl Acad. Sci. USA* **110**, 14984–14989 (2013).
This study provides a NGS analysis of two parallel aptamer selections, starting from high-coverage pools and presenting isolation versus connectivity between peaks in the fitness landscape.
116. Sassanfar, M. & Szostak, J. W. An RNA motif that binds ATP. *Nature* **364**, 550–553 (1993).
117. Burke, D. H. & Gold, L. RNA aptamers to the adenosine moiety of *S*-adenosyl methionine: structural inferences from variations on a theme and the reproducibility of SELEX. *Nucleic Acids Res.* **25**, 2020–2024 (1997).
118. Gebhardt, K., Shokraei, A., Babaie, E. & Lindqvist, B. H. RNA aptamers to *S*-adenosylhomocysteine: kinetic properties, divalent cation dependency, and comparison with anti-*S*-adenosylhomocysteine antibody. *Biochemistry* **39**, 7255–7265 (2000).
119. Burgstaller, P. & Famulok, M. Isolation of RNA aptamers for biological cofactors by *in vitro* selection. *Angew. Chem. Int. Ed. Engl.* **33**, 1084–1087 (1994).
120. Vu, M. M. K. *et al.* Convergent evolution of adenosine aptamers spanning bacterial, human, and random sequences revealed by structure-based bioinformatics and genomic SELEX. *Chem. Biol.* **19**, 1247–1254 (2012).
121. Salehi-Ashtiani, K. & Szostak, J. W. *In vitro* evolution suggests multiple origins for the hammerhead ribozyme. *Nature* **414**, 82–84 (2001).
122. Seehafer, C., Kalweit, A., Steger, G., Gräf, S. & Hammann, C. From alpaca to zebrafish: hammerhead ribozymes wherever you look. *RNA* **17**, 21–26 (2011).
123. Jimenez, R. M., Delwart, E. & Lupták, A. Structure-based search reveals hammerhead ribozymes in the human microbiome. *J. Biol. Chem.* **286**, 7737–7743 (2011).
124. Perreault, J. *et al.* Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS Comput. Biol.* **7**, e1002031 (2011).
125. Maharjan, R. P. & Ferenci, T. Epistatic interactions determine the mutational pathways and coexistence of lineages in clonal *Escherichia coli* populations. *Evolution* **67**, 2762–2768 (2013).
126. Chou, H.-H., Delaney, N. F., Draghi, J. A. & Marx, C. J. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and sign epistasis between adaptive mutations. *PLoS Genet.* **10**, e1004149 (2014).
127. Chiotti, K. E. *et al.* The valley-of-death: reciprocal sign epistasis constrains adaptive trajectories in a constant, nutrient limiting environment. *Genomics* **104**, 431–437 (2014).
128. Kvitek, D. J. & Sherlock, G. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet.* **7**, e1002056 (2011).
129. Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L. & Ayroles, J. F. Genetic incompatibilities are widespread within species. *Nature* **504**, 135–137 (2013).
130. Hayden, E. J., Ferrada, E. & Wagner, A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474**, 92–95 (2011).
131. Fisher, R. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).
132. Tenaillon, O. The utility of Fisher's geometric model in evolutionary genetics. *Annu. Rev. Ecol. Evol. Syst.* **45**, 179–201 (2014).
133. Orr, H. A. Theories of adaptation: what they do and don't say. *Genetica* **123**, 3–13 (2005).
134. Couce, A. & Tenaillon, O. A. The rule of declining adaptability in microbial evolution experiments. *Front. Genet.* **6**, 1–6 (2015).
135. Sanjuán, R., Cuevas, J. M., Moya, A. & Elena, S. F. Epistasis and the adaptability of an RNA virus. *Genetics* **170**, 1001–1008 (2005).
136. Rokyta, D. R., Abdo, Z. & Wichman, H. A. The genetics of adaptation for eight microvirid bacteriophages. *J. Mol. Evol.* **69**, 229–239 (2009).
137. Perfeito, L., Sousa, A., Bataillon, T. & Gordo, I. Rates of fitness decline and rebound suggest pervasive epistasis. *Evolution* **68**, 150–162 (2014).
138. Moore, F. B., Rozen, D. E. & Lenski, R. E. Pervasive compensatory adaptation in *Escherichia coli*. *Proc. Biol. Sci.* **267**, 515–522 (2000).
139. MacLean, R. C., Perron, G. G. & Gardner, A. Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in *Pseudomonas aeruginosa*. *Genetics* **186**, 1345–1354 (2010).
140. Bull, J. J., Badgett, M. R. & Wichman, H. A. Big-benefit mutations in a bacteriophage inhibited with heat. *Mol. Biol. Evol.* **17**, 942–950 (2000).
141. Weinreich, D. M., Delaney, N. F., DePristo, M. A. & Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111–114 (2006).
142. 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).
143. Colosimo, P. F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).
144. Barroso-Batista, J. *et al.* The first steps of adaptation of *Escherichia coli* to the gut are dominated by soft sweeps. *PLoS Genet.* **10**, e1004182 (2014).
145. Kimura, M. Evolutionary rate at the molecular level. *Nature* **217**, 624–626 (1968).
146. Haldane, J. B. A mathematical theory of natural & artificial selection part X. Some theorems on artificial selection. *Genetics* **19**, 412–429 (1934).
147. Crow, J. F. & Kimura, M. *An Introduction to Population Genetics Theory* (Harper and Row, 1970).
148. Lescat, M. *et al.* The conserved nhaAR operon is drastically divergent between B2 and non-B2 *Escherichia coli* and is involved in extra-intestinal virulence. *PLoS ONE* **9**, e108738 (2014).

Acknowledgements
This work was supported by: the Borchard Scholar-in-Residence Program; ATIP-Avenir (Centre national de la recherche scientifique (CNRS/INSERM) and Institut National de la Santé et de la Recherche Médicale), FP7-PEOPLE-2012-CIG (322035), L'Agence nationale de la recherche (ANR-13-BSV6-0006-01 and ANR-11-LABX-0028-01), ARC (SFI20111203947) and La Ligue contre le cancer; US National Science Foundation Molecular and Cellular Biosciences (MCB1330606); and the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) European Research Council grant 310944.

Competing interests statement
The authors declare no competing interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (box)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF