Chapter 7
Experimental Evolution and Next Generation Sequencing Illuminate the Evolutionary Trajectories of Microbes

Mario A. Fares

Introduction

In his book "The origin of species by means of natural selection" (Darwin 1859), Charles Darwin manifested a deep frustration justified by the realization that natural selection is too slow to be observed in real time. He admittedly based all his conclusions on observations or indirect measurements of the action of natural selection and reported many evidence supporting his conclusion: "That natural selection will always act with extreme slowness I fully admit."

Darwin, if lived today, would be enthralled by the fact that the process of natural selection and the mechanisms underlying them could be directly tested in a reasonable short time using microbes. Microbes offer a unique opportunity to observe and test the mechanism of natural selection and the general principles of evolution. This is mainly due to the short generation times, small genome sizes, and deep microbes genetic and physiological characterization. These features and the feasibility of evolving microbes in the laboratory with the current technology under controlled conditions and at high "speeds" make them ideal systems to put the main principles of evolution to test and unearth the dynamics underlying the evolution of biological complexity (Kawecki et al. 2012). In addition to the possibility of conducting laboratory-supervised evolution experiments, the next generation sequencing technology (NGS) has enabled sequencing hundreds of microbial genomes at once, linking particular genome dynamics to microbes' lifestyles.

M.A. Fares (✉)
Department of Abiotic Stress, Instituto de Biología Molecular y Celular de Plantas (CSIC-UPV), Ingeniero Fausto Elio, Valencia 46022, Spain

Department of Genetics, Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin, Ireland
e-mail: mfares@ibmcp.upv.es, faresm@tcd.ie

G. Sablok et al. (eds.), *Advances in the Understanding of Biological Sciences Using Next Generation Sequencing (NGS) Approaches*,
DOI 10.1007/978-3-319-17157-9_7

In this chapter, I will discuss the many different scenarios under which microbes have been evolved in the laboratory, how did NGS contribute to the understanding of the genomes dynamics behind specific adaptive processes, and the main conceptual breakthroughs derived from these studies.

What Makes Microbes Attractive to Test Evolutionary Processes?

Eighty-five years ago, August Krogh articulated a principle (Krogh principle) after which experimentalists should choose the model organism that can best foster a clear and direct experimental design and a rigorous and unambiguous result and interpretation (Krogh 1929). Krogh principle is particularly useful when testing evolutionary processes, as these are often dominated by very complex patterns that are intermingled and many times shaped by the environment.

The general principles of Evolutionary biology has been historically built based on indirect theoretical and comparative studies (Futuyma 1998), lacking rigorous experimentation proof. There are several reasons for the lack of experimental studies probing principles of evolution. Mainly, it remains difficult identifying the dynamics of natural selection leading to the fixation of advantageous mutations at specific episode of organisms' evolution. Some of the reasons for this difficulty are the impracticality of replicating the complex mix of environmental conditions under which populations grew at some stage during their evolution and the slow pace at which natural selection acts. In this sense, microorganisms offer a unique opportunity for studying evolution as they present large populations sizes, short generation times, small genome sizes, and enormous physiological plasticity. Noticeably, microorganisms are not equipped with complex homeostasis systems, and thus their phenotype is largely the result of their genetic composition interacting with the environment. In addition to this convenient feature, microbes present a puzzling diversification whether measured in terms of the number of species (Dykhuizen 1998; Gans et al. 2005), habitat range (Pikuta et al. 2007), or the breadth of energy sources and biochemical pathways they can exploit in order to survive (Pace 1997).

The hallmarks of experimentation of any kind are control and replication. In evolutionary biology, controlling environmental conditions, especially when conducting experiments out of the laboratory, is difficult if not impossible. However, the fact that enormous population size of microbes could grow in tiny spaces (for example, a drop of culture medium) makes it feasible growing hundreds of microbial populations in a standard laboratory space. Moreover, microbiologists have successfully harnessed bacterial evolution and domesticated them to grow under laboratory-controlled conditions. Hundreds of microbial populations can be then propagated and analyzed simultaneously. If maintained evolving separately, with no cross-contamination, such populations can be used to test the repeatability of evolutionary processes (Lenski et al. 2000), to understand the physiological plasticity of bacteria growing under different carbon sources, and reproduce ecological scenarios of more complex organisms. In summary, experimental evolution allows determining

the selective forces operating, and by virtue of replicating the experiment, researchers can distinguish between deterministic and stochastic effects.

Environmental control is one of the most important advantages of using microbial populations because we can grow homogenously distributed populations in an environment in which single factors can be modified. In this new single-factor modified environments, that reproduces ancestral environments, many hypotheses can be tested, including how novel physiologies emerge to adapt to a new environment, the population dynamics of generalists and specialists, and the role of contingency in the adaptation to novel conditions and the trade-offs that such adaptations involve (Bennett and Lenski 2007; Bronikowski et al. 2001; Lee et al. 2009).

The large population sizes of microbes offers an analytical advantage, which is concerned with the higher likelihood of originating novel adaptations through mutations. The rationale is simple: in a small space of culture liquid billions of microbial cells can be kept and propagated, thereby avoiding the effect of genetic drift and directly testing the role of natural selection. During DNA replication, or even protein translation, there is a low but finite probability of an error in replication. The probability of occurrence of such a mutation is the product of the population size and the mutation rate. Therefore, the larger the population size the greater is the number of mutations originating in the population and the higher is the probability of a mutational novelty emerging. Because selection is strong when population sizes are large, the probability of fixation of beneficial mutations is very high. It follows then that the rate at which evolution occurs is high in microbial populations, making it possible reproducing adaptive evolution in real time. Indeed, in long-term evolutionary studies on microbial populations, every single nucleotide base pair should have experienced at least one mutation, and thus have undergone selection filtering (Lenski et al. 2003).

Finally, unlike multicellular organisms that require at least days or weeks to produce a new generation, microbes require minutes or hours. This allows beneficial mutations to become quickly fixed in the populations. For example, thermo-resistant mutations can become fixed in the microbial population within 15–20 days after initiating an evolution experiment (Bennett and Lenski 2007; Elena and Lenski 2003). Moreover, the enormous linkage disequilibrium of microbes ensures their clonal transmission for thousands of generations preserving the ancestral genetic background. This, in addition to the possibility of freezing evolved cells that can be thawed again, permits building a microbial fossil record and perform genome archeology at any time of the evolution experiment (Lenski et al. 2003; Ostrowski et al. 2008).

Experimental Evolution and Mutation Accumulation Dynamics

Experimental evolution combined with whole-genome re-sequencing is a promising strategy for investigating the dynamics of evolutionary change. One of the questions that have motivated efforts in reproducing an evolutionary scenario is how repeatable is evolution. The fragmentary nature of the fossil record cannot provide a full

picture that would allow answering this question, and even if it did we are not certain what kinds of environments or adaptations have not been explored by nature. Instead, reproducing fine-tuned scenarios in a test tube containing billions of bacterial cells can shed light on the complexity of evolutionary patterns.

Evolution experiments start with an initial population of microbes genetically identical and adapted to an ancestral environment (Fig. 7.1). Adaptation is determined by the Malthusian growth parameter of the population and is considered to be proportional to the relative fitness of the population. Fitness in experimentally evolved populations is measured as the capacity of such descendent populations to compete head-to-head with their ancestors. These two populations, the evolved one and its parental ancestral population, can be compared because they can be brought together in the same place at the same time. We can compare the performance of the descendant and ancestral populations by quantifying the number of offspring that each leaves in the next generations in an environment in which the carbon source is common for the two differentiated populations. Populations are propagated between

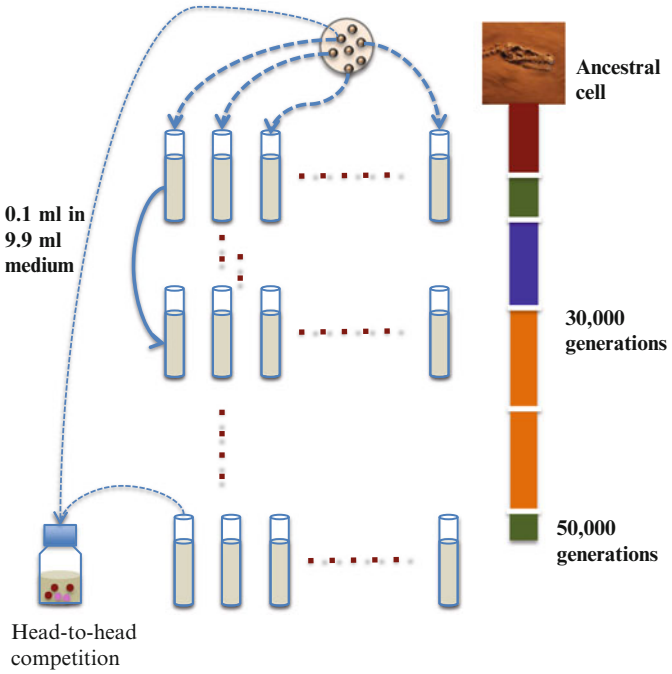

Fig. 7.1 Experimental evolution of microbes in the laboratory. From a single cell (ancestral cell) many replicate populations are generated, all being genetically identical and evolving for many generations independently. To assay the biological fitness of the evolved population at any time point, equal proportions of this population and of its ancestral one are mixed in the same medium. Both cells, the evolved and ancestral ones, should be distinguishable, for example through a metabolic marker that yields distinctly colored cells, to determine the relative frequency of each population at the start and end of the competition experiment. Improved fitness of the evolved population is reflected in a higher proportion of evolved cells than ancestors in the competition experiment

generations by diluting 0.1 ml of the grown culture in 9.9 ml of a new culture. To determine how repeatable is evolution, many different independently evolving lineages are generated from the same ancestor, and thus originally presenting the same genetic background and evolved in parallel (Fig. 7.1). The many different evolutionary paths followed by each of the independent evolving lines can be then compared and their differences quantified.

As I explained earlier, microbes are genetically represented by one chromosome. The gamma-proteobacterium *Escherichia coli* strain K12 MG1655 is the one most used in experimental evolution of microbes. Most bacteria, including *E. coli*, present highly dense genomes, with the genome size reflecting the number of genes (Giovannoni et al. 2005; Mira et al. 2001). The high gene density of these genomes and large linkage disequilibrium means that the mutational load is expected to increase as generations pass by without disrupting previous genetic backgrounds and that most changes will be affecting coding genes or regulatory regions. This means that we can directly associate particular nucleotide mutations to specific phenotypes and follow the history of interesting mutations since the last common ancestor of all the founded bacterial populations. Likewise, the yeast *Saccharomyces cerevisiae* has been used in its haploid or diploid genetic structure as a model to test specific evolutionary processes through experimental evolution. Here I provide examples of how NGS performs a powerful tool when combined with experimental evolution to unearth the rules governing fundamental evolutionary processes.

The Evolutionary Trajectories of Adaptive Mutations

NGS has been developed reaching a stage in which single minority mutations can be identified at low frequencies and their origin traced through reviving evolved cells at different time points of an evolution experiment. For example, the final stages of the fixation of an adaptive mutation can be identified by mixing equal proportions of bacterial cells labeled with two different tags (Hegreness et al. 2006). Combining cost-effective Illumina re-sequencing with experimental evolution makes it possible to sequence several hundreds of individuals from an evolved population, generating estimates of allele frequencies at millions of single-nucleotide polymorphisms (SNPs) genome-wide (Burke 2012; Burke et al. 2010; Burke and Long 2012; Futschik and Schlotterer 2010). This is important not only to identify rare variants but also to determine with unprecedented accuracy the evolutionary trajectories of adaptive mutations.

Evolution experiments seeking to identify adaptive evolution derive populations from a single ancestral genotypes, and thus genetically identical, in a constant environment or an environment with constant fluctuations. This is achieved by a continuous culture of populations in which the input of resources and the removal of individuals occur at a constant and controlled way. Alternatively, a fraction of the grown population is passaged to a new culture medium. When an adaptive mutation emerges in such an environment, this drives the evolutionary dynamic of the population, so that the

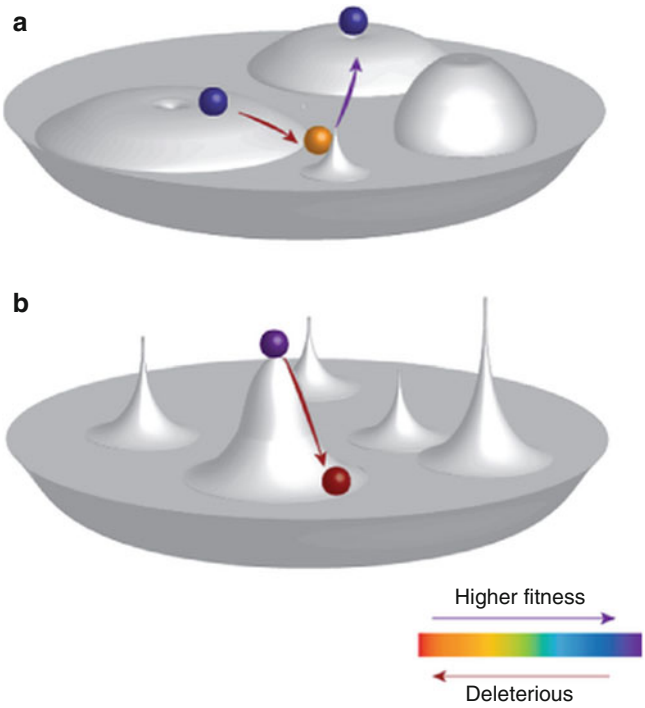

Fig. 7.2 Fitness landscape of an evolving population. Peaks represent regions of maximum relative biological fitness while valleys are regions of low fitness. In a smooth landscape (**a**) populations (*spheres*) can cross the valleys of low fitness without yielding lethal phenotypes (e.g., these populations go from high mean fitness to intermediate mean fitness). In rugged and complex landscape (**b**), crossing the valleys of low fitness is lethal and precludes populations from reaching new local fitness maxima through gradual evolution. This figure is taken from (Henderson et al. 2013), with author's permission

average fitness of the population increases gradually. When several adaptive mutations emerge, synergistic epistasis among them, that is interactions between mutations that increases the effects of single mutants on fitness in a non-linear fashion, leads to diminishing-returns epistasis: each mutation has lower beneficial effect for the individuals in the presence of another beneficial mutations than if it appeared alone in the ancestral genetic background (Chou et al. 2011; Khan et al. 2011; Kvitek and Sherlock 2011). Regardless of whether or not diminishing returns take place, beneficial mutations will lead populations to climb peaks in a fitness landscape (Fig. 7.2) (Orr 2009a, b). In the absence of interfering mutations, beneficial mutations will undergo refinement and selective sweep in the population (Atwood et al. 1951; Barrick and Lenski 2013). However, in asexual populations it is more frequent to observe cases in which the beneficial mutation needs to displace other beneficial mutations emerging during its fixation, thereby slowing down the fixation rate of

adaptive mutations. This effect, known as clonal interference (Fogle et al. 2008; Miralles et al. 1999), has been shown to be frequent in asexual populations of influenza (Strelkowa and Lassig 2012), the bacteriophage phiX174 (Pepin and Wichman 2008), bacteria (de Visser and Rozen 2006), and yeast (Kao and Sherlock 2008; Lang et al. 2013) but has only been characterized in yeast by deep sequencing yeast populations at frequent intervals (Lang et al. 2013).

Adaptive mutations need to be distinguished from those that are innovative, leading to new phenotypes adaptable to novel environments. Many research studies in this area have shown that such innovative mutations are often sudden and involve only one-to-few mutations. The identification of these mutations has been possible through the use of NGS, which has also enabled disentangling beneficial mutations from innovative ones. For example, in a recent study, Marchetti and colleagues showed that an experimentally evolved chimeric *Ralstonia solanacearum* strain, derived from a plant pathogen, could establish a symbiotic mutualistic association once evolved experimentally. This change in lifestyle occurred upon colonizing root nodules and was due to a single non-synonymous (amino acid replacing) mutation in the gene *hrpG* that encodes a protein regulating the expression of several virulent factors (Marchetti et al. 2010). In another study in which authors conducted a long-term evolution experiment with *E. coli* (LTEE), *E. coli* adapted to a glucose-limited medium, which also contained the bacterium-unusable citrate, evolved the ability to metabolize citrate after 30,000 generations in one of the 12 original replicate populations with which the experiment commenced (Blount et al. 2008). The emergence of this innovation required a single genome event in earlier generations (an enabling mutation), consistent on a chromosomal duplication that placed a transcription promoter upstream of a Citrate transporter-encoding gene (Blount et al. 2012).

The concept of genetic background and enabling mutations is very important to understand the term "evolvability"—the capacity of individuals or genotypes to evolve and adapt to a wide set of different conditions. Indeed, the combination of alleles existing in the population may well condition and constrain the evolutionary trajectories of new alleles, through either altering mutation rates or conditioning the nature and strength of epistatic interactions with new mutations (Meyer et al. 2012). The actual dynamics underlying the enabling effect of neutral mutation networks has been investigated in very simple systems, such as RNA folding (Wagner 2008), however, the role of enabling mutations versus compensatory mutations—those compensating the effects of destabilizing innovative mutations—remains the ground of intense investigation and debate.

As discussed earlier, populations with high mutation rates increase the per-capita chance of acquiring a beneficial mutation. In LTEE, the frequency of hyper-mutators is high, rising mutation rates 100-fold compared to that of the ancestral population (Mao et al. 1997). However, in recent studies it has been shown that hyper-mutators in experimental populations are generally followed by phenotypes with slow mutation rates, probably because such phenotypes prevent the loss of adaptive mutations in the populations and lower genetic load (Sniegowski et al. 2000; Wielgoss et al. 2013).

Convergent Evolution in Bacterial Experimental Populations

One of the most important questions yet unanswered is how repeatable is evolution. In particular, what is the role of contingency in the fixation of adaptive mutations? In a recent study (Tenaillon et al. 2012), authors evolved 115 *E. coli* populations for 2,000 generations of the bacterium to adapt to 42.2 °C, a complex environmental factor to which many pathways of the organism respond. To determine the diversity of adaptation of *E. coli* to high temperatures, they started the experiment from a single ancestral cell adapted to 37 °C. After 2,000 generations, the genome of one clone from each of the 115 experimentally evolving populations at 42.2 °C was sequenced. In addition to genome sequencing, the relative fitness of the evolved clones was assessed, observing a significant increase of fitness of the evolved strain at 42.2 °C in comparison with their ancestor. Interestingly, in 18 of the 115 lines, authors found a shared mutation in codon 966 of the RNA polymerase β-subunit (*rpoB*), and 17 lines contained an amino acid replacing mutation in codon 15 of the *rho* gene. In general, 20.2 % of genes mutated convergently in their experiment and 24.5 % of operons were convergently affected by mutations. This significant convergence was strongly driven by the epistatic interactions between new alleles. These experiments demonstrate that while the range of adaptive pathways may be bewildering, epistasis and genetic background can constrain the set of possible solutions to adapt to an environment, making evolution somewhat predictable.

Experimental Evolution Under Inefficient Natural Selection

To study the spectrum of mutations, researchers have evolved microbes, such as *E. coli* and *S. cerevisiae*, under controlled laboratory experiments and re-sequenced their genomes at different time points of the evolution experiment. Because the main objective of these experiments is to identify the breadth of mutations occurring in the genome, and calculate the rates of mutations, such populations have been evolved under very inefficient natural selection: replicates of evolving lines were single-colony transferred to new plates and this was repeated for hundreds or even thousands of generations (Fig. 7.3). These experiments have been useful to determine the spectrum and rate of mutations in *E. coli* (Lee et al. 2012) and *S. cerevisiae* (Lynch et al. 2008).

Purifying selection generally precludes the fixation of innovative mutations because they are generally destabilizing owing to the trade-off between current and novel adaptations (DePristo et al. 2005; Wilke et al. 2005; Zeldovich et al. 2007). There are a number of scenarios in which innovative mutations can be fixed under inefficient natural selection, including gene duplication (Ohno 1999), and systems with over-active mechanisms of mutational robustness, such as over-expressed molecular chaperones (Moran 1996).

How does gene duplication enable the fixation of innovative mutations? After the duplication of a gene, the two daughter copies are virtually identical, hence

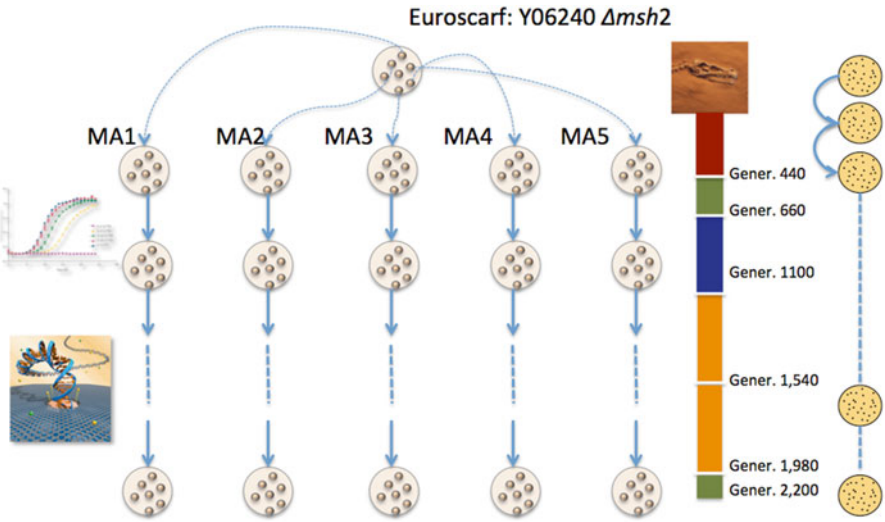

Fig. 7.3 Experimental evolution of populations of yeast under inefficient natural selection. Many replicate populations derive from a single yeast cell. To impose population bottlenecks and genetic drift a single colony is transferred to the new environment (plate). In the figure example, five independent lines of evolution started and evolved for many generations. At specific points of the evolution experiment, whole-genome sequencing and growth curves are conducted and mutations mapped in the reference genome

functionally redundant, with some exceptions that include non-duplicated regulatory elements, moving of one gene copy to a differently transcribed genome region or allele ancestral polymorphism (Lynch and Katju 2004). Such exceptions may well determine the spectrum of subsequent mutations of each gene copy, and consequently the functional fates of duplicates. The asymmetry between gene copies is avoided in many biological systems such as yeast through whole-genome duplication (WGD) but not through small-scale duplications (SSD). Accordingly, a number of studies have shown that the mechanism of duplication can determine the persistence of genes in duplicate, with WGDs being more prevalent among central genes in the network (although with some exceptions depending on the organism (Alvarez-Ponce and Fares 2012)), they are refractory to subsequent SSD events and dosage sensitive (Carretero-Paulet and Fares 2012; Conant and Wolfe 2006; Fares et al. 2013; Hakes et al. 2007; Makino and McLysaght 2010). These studies have shown that SSDs are more likely to present redundancy, hence mutational robustness and evolvability (Draghi et al. 2010), than WGDs. In particular, Fares and colleagues conducted a simple mutation accumulation experiment in which five lines of *S. cerevisiae* haploid strains derived from a single ancestor deficient in a mismatch repair gene (*msh2*) were evolved independently under strong genetic drift. They passaged these lines periodically by single colony transfers from one generation to the next for 2,200 generations. The whole genome of one colony was sequenced from each line and the distribution of non-synonymous SNPs in duplicates and singletons identified. As predicted by theory, SSDs showed significantly

larger number of non-synonymous SNPs than singletons and WGDs, supporting larger redundancy for SSDs than WGDs (Fares et al. 2013).

Experimental evolution has also been used to determine the role of a molecular chaperone in ameliorating the effects of deleterious non-lethal mutations. In an experiment in which several independent *E. coli* lines were subjected to single-colony passages, authors assessed the fitness of evolved population by competing them head-to-head to their ancestral population. After 3,200 generations of experimental bottlenecked evolution, cells presented half as much fitness as their ancestors owing to the increase in the deleterious mutational load owing to strong genetic drift effects. Over-expression of GroEL, a molecular chaperone essential in *E. coli* and which folds other proteins in the cell (Fayet et al. 1989; Lin and Rye 2006), allowed the recovery of about 88 % of the fitness of evolved cells (Fares et al. 2002). Interestingly, *groESL*, the operon encoding the chaperonin GroEL and its cofactor GroES, is abundantly synthesized in endosymbiotic mutualistic bacteria (Ahn et al. 1994; Aksoy 1995) that undergo strong genetic drift during their clonal transmission from mother host to the offspring (Buchner 1965). Experimental evolution of *E. coli* under inefficient natural selection reproduced therefore the transmission of endosymbiotic bacteria and identified GroEL as a mechanism of robustness against deleterious non-lethal mutations.

Concluding Remarks

Experimental evolution is a powerful tool to reproduce particular evolutionary processes with high repeatability and under tightly controlled environmental conditions. When combined with whole-genome sequencing, experimental evolution can inform on the dynamics underlying adaptations, speed of evolution, role of environment, and evolvability. Current studies have unveiled unprecedented and unexpected outcomes and have revealed complex dynamics to adaptation. While the general principles of evolution by natural selection clearly follow Darwinian laws, the evolutionary trajectories, contingency, constraints, and evolvability of organisms remain largely obscure. Future research in population genomics combined with NGS will be the key for understanding how do adaptations come about, how they interact, and where they lead.

References

Ahn TI, Lim ST, Leeu HK, Lee JE, Jeon KW (1994) A novel strong promoter of the groEx operon of symbiotic bacteria in *Amoeba proteus*. Gene 148:43–49

Aksoy S (1995) Molecular analysis of the endosymbionts of tsetse flies: 16S rDNA locus and over-expression of a chaperonin. Insect Mol Biol 4:23–29

Alvarez-Ponce D, Fares MA (2012) Evolutionary rate and duplicability in the *Arabidopsis thaliana* protein–protein interaction network. Genome Biol Evol 4:1263–1274. doi:10.1093/gbe/evs101

Atwood KC, Schneider LK, Ryan FJ (1951) Periodic selection in *Escherichia coli*. Proc Natl Acad Sci U S A 37:146–155
Barrick JE, Lenski RE (2013) Genome dynamics during experimental evolution. Nat Rev Genet. doi:10.1038/nrg3564
Bennett AF, Lenski RE (2007) An experimental test of evolutionary trade-offs during temperature adaptation. Proc Natl Acad Sci U S A 104(Suppl 1):8649–8654. doi:10.1073/pnas.0702117104
Blount ZD, Borland CZ, Lenski RE (2008) Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc Natl Acad Sci U S A 105:7899–7906. doi:10.1073/pnas.0803151105
Blount ZD, Barrick JE, Davidson CJ, Lenski RE (2012) Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature 489:513–518. doi:10.1038/nature11514
Bronikowski AM, Bennett AF, Lenski RE (2001) Evolutionary adaptation to temperature. VIII Effects of temperature on growth rate in natural isolates of Escherichia coli and Salmonella enterica from different thermal environments. Evolution 55:33–40
Buchner P (1965) Endosymbiosis of animals with plant microorganisms. Wiley, New York
Burke MK (2012) How does adaptation sweep through the genome? Insights from long-term selection experiments. Proc Biol Sci 279:5029–5038. doi:10.1098/rspb.2012.0799
Burke MK, Long AD (2012) What paths do advantageous alleles take during short-term evolutionary change? Mol Ecol 21:4913–4916
Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD (2010) Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature 467:587–590. doi:10.1038/nature09352
Carretero-Paulet L, Fares MA (2012) Evolutionary dynamics and functional specialization of plant paralogs formed by whole and small-scale genome duplications. Mol Biol Evol 29:3541–3551. doi:10.1093/molbev/mss162
Chou HH, Chiu HC, Delaney NF, Segre D, Marx CJ (2011) Diminishing returns epistasis among beneficial mutations decelerates adaptation. Science 332:1190–1192. doi:10.1126/science.1203799
Conant GC, Wolfe KH (2006) Functional partitioning of yeast co-expression networks after genome duplication. PLoS Biol 4:e109. doi:10.1371/journal.pbio.0040109
Darwin C (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Myrray (Ed). London
de Visser JA, Rozen DE (2006) Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. Genetics 172:2093–2100. doi:10.1534/genetics.105.052373
DePristo MA, Weinreich DM, Hartl DL (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. Nat Rev Genet 6:678–687. doi:10.1038/nrg1672
Draghi JA, Parsons TL, Wagner GP, Plotkin JB (2010) Mutational robustness can facilitate adaptation. Nature 463:353–355. doi:10.1038/nature08694
Dykhuizen DE (1998) Santa Rosalia revisited: why are there so many species of bacteria? Antonie Van Leeuwenhoek 73:25–33
Elena SF, Lenski RE (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat Rev Genet 4:457–469. doi:10.1038/nrg1088
Fares MA, Ruiz-Gonzalez MX, Moya A, Elena SF, Barrio E (2002) Endosymbiotic bacteria: groEL buffers against deleterious mutations. Nature 417:398. doi:10.1038/417398a
Fares MA, Keane OM, Toft C, Carretero-Paulet L, Jones GW (2013) The roles of whole-genome and small-scale duplications in the functional specialization of *Saccharomyces cerevisiae* genes. PLoS Genet 9:e1003176. doi:10.1371/journal.pgen.1003176
Fayet O, Ziegelhoffer T, Georgopoulos C (1989) The groES and groEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. J Bacteriol 171:1379–1385
Fogle CA, Nagle JL, Desai MM (2008) Clonal interference, multiple mutations and adaptation in large asexual populations. Genetics 180:2163–2173. doi:10.1534/genetics.108.090019
Futschik A, Schlotterer C (2010) The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. Genetics 186:207–218. doi:10.1534/genetics.110.114397
Futuyma DJ (1998) Evolutionary biology. Sinauer, Sunderland, MA

Gans J, Wolinsky M, Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 309:1387–1390. doi:10.1126/science.1112665

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappe MS, Short JM, Carrington JC, Mathur EJ (2005) Genome streamlining in a cosmopolitan oceanic bacterium. Science 309:1242–1245. doi:10.1126/science.1114057

Hakes L, Robertson DL, Oliver SG, Lovell SC (2007) Protein interactions from complexes: a structural perspective. Comp Funct Genomics 49356. doi: 10.1155/2007/49356

Hegreness M, Shoresh N, Hartl D, Kishony R (2006) An equivalence principle for the incorporation of favorable mutations in asexual populations. Science 311:1615–1617. doi:10.1126/science.1122469

Henderson B, Fares MA, Lund PA (2013) Chaperonin 60: a paradoxical, evolutionarily conserved protein family with multiple moonlighting functions. Biol Rev Camb Philos Soc 88:955–987. doi:10.1111/brv.12037

Kao KC, Sherlock G (2008) Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. Nat Genet 40:1499–1504. doi:10.1038/ng.280

Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC (2012) Experimental evolution. Trends Ecol Evol 27:547–560. doi:10.1016/j.tree.2012.06.001

Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF (2011) Negative epistasis between beneficial mutations in an evolving bacterial population. Science 332:1193–1196. doi:10.1126/science.1203801

Krogh A (1929) Progress of physiology. Am J Physiol 90:9

Kvitek DJ, Sherlock G (2011) Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. PLoS Genet 7:e1002056. doi:10.1371/journal.pgen.1002056

Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, Botstein D, Desai MM (2013) Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. Nature 500:571–574. doi:10.1038/nature12344

Lee MC, Chou HH, Marx CJ (2009) Asymmetric, bimodal trade-offs during adaptation of Methylobacterium to distinct growth substrates. Evolution 63:2816–2830. doi:10.1111/j.1558-5646.2009.00757.x

Lee H, Popodi E, Tang H, Foster PL (2012) Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. Proc Natl Acad Sci U S A 109:E2774–E2783. doi:10.1073/pnas.1210309109

Lenski RE, Rose MR, Simpson SC, Stadler SC (2000) Long-term experimental evolution in *Escherichia coli*. Am Nat 138:27

Lenski RE, Winkworth CL, Riley MA (2003) Rates of DNA sequence evolution in experimental populations of *Escherichia coli* during 20,000 generations. J Mol Evol 56:498–508. doi:10.1007/s00239-002-2423-0

Lin Z, Rye HS (2006) GroEL-mediated protein folding: making the impossible, possible. Crit Rev Biochem Mol Biol 41:211–239. doi:10.1080/10409230600760382

Lynch M, Katju V (2004) The altered evolutionary trajectories of gene duplicates. Trends Genet 20:544–549. doi:10.1016/j.tig.2004.09.001

Lynch M, Sung W, Morris K, Coffey N, Landry CR, Dopman EB, Dickinson WJ, Okamoto K, Kulkarni S, Hartl DL, Thomas WK (2008) A genome-wide view of the spectrum of spontaneous mutations in yeast. Proc Natl Acad Sci U S A 105:9272–9277. doi:10.1073/pnas.0803466105

Makino T, McLysaght A (2010) Ohnologs in the human genome are dosage balanced and frequently associated with disease. Proc Natl Acad Sci U S A 107:9270–9274. doi:10.1073/pnas.0914697107

Mao EF, Lane L, Lee J, Miller JH (1997) Proliferation of mutators in A cell population. J Bacteriol 179:417–422

Marchetti M, Capela D, Glew M, Cruveiller S, Chane-Woon-Ming B, Gris C, Timmers T, Poinsot V, Gilbert LB, Heeb P, Medigue C, Batut J, Masson-Boivin C (2010) Experimental evolution

of a plant pathogen into a legume symbiont. PLoS Biol 8:e1000280. doi:10.1371/journal.pbio.1000280
Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE (2012) Repeatability and contingency in the evolution of a key innovation in phage lambda. Science 335:428–432. doi:10.1126/science.1214449
Mira A, Ochman H, Moran NA (2001) Deletional bias and the evolution of bacterial genomes. Trends Genet 17:589–596
Miralles R, Gerrish PJ, Moya A, Elena SF (1999) Clonal interference and the evolution of RNA viruses. Science 285:1745–1747
Moran NA (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc Natl Acad Sci U S A 93:2873–2878
Ohno S (1999) Gene duplication and the uniqueness of vertebrate genomes circa 1970–1999. Semin Cell Dev Biol 10:517–522. doi:10.1006/scdb.1999.0332
Orr HA (2009a) Fitness and its role in evolutionary genetics. Nat Rev Genet 10:531–539. doi:10.1038/nrg2603
Orr HA (2009b) Testing natural selection. Sci Am 300:44–50
Ostrowski EA, Woods RJ, Lenski RE (2008) The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. Proc Biol Sci 275:277–284. doi:10.1098/rspb.2007.1244
Pace NR (1997) A molecular view of microbial diversity and the biosphere. Science 276:734–740
Pepin KM, Wichman HA (2008) Experimental evolution and genome sequencing reveal variation in levels of clonal interference in large populations of bacteriophage phiX174. BMC Evol Biol 8:85. doi:10.1186/1471-2148-8-85
Pikuta EV, Hoover RB, Tang J (2007) Microbial extremophiles at the limits of life. Crit Rev Microbiol 33:183–209. doi:10.1080/10408410701451948
Sniegowski PD, Gerrish PJ, Johnson T, Shaver A (2000) The evolution of mutation rates: separating causes from consequences. Bioessays 22:1057–1066. doi:10.1002/1521-1878(200012)22:12<1057::AID-BIES3>3.0.CO;2-W
Strelkowa N, Lassig M (2012) Clonal interference in the evolution of influenza. Genetics 192:671–682. doi:10.1534/genetics.112.143396
Tenaillon O, Rodriguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, Gaut BS (2012) The molecular diversity of adaptive convergence. Science 335:457–461. doi:10.1126/science.1212986
Wagner A (2008) Neutralism and selectionism: a network-based reconciliation. Nat Rev Genet 9:965–974. doi:10.1038/nrg2473
Wielgoss S, Barrick JE, Tenaillon O, Wiser MJ, Dittmar WJ, Cruveiller S, Chane-Woon-Ming B, Medigue C, Lenski RE, Schneider D (2013) Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. Proc Natl Acad Sci U S A 110:222–227. doi:10.1073/pnas.1219574110
Wilke CO, Bloom JD, Drummond DA, Raval A (2005) Predicting the tolerance of proteins to random amino acid substitution. Biophys J 89:3714–3720. doi:10.1529/biophysj.105.062125
Zeldovich KB, Berezovsky IN, Shakhnovich EI (2007) Protein and DNA sequence determinants of thermophilic adaptation. PLoS Comput Biol 3:e5. doi:10.1371/journal.pcbi.0030005