Molecular Systems Biology 7; Article number 509; doi:10.1038/msb.2011.42
Citation: ***Molecular Systems Biology* 7:509**

www.molecularsystemsbiology.com

molecular systems biology

REVIEW

Microbial laboratory evolution in the era of genome-scale science

Tom M Conrad, Nathan E Lewis and Bernhard Ø Palsson*

Department of Bioengineering, University of California San Diego, La Jolla, CA, USA
* Corresponding author. Department of Bioengineering, University of California San Diego, 9500 Gilman Drive, Mail Code 0412, La Jolla, CA 92093-0412, USA. Tel.: +1 858 534 5668; Fax: +1 858 822 3120;
E-mail: palsson@ucsd.edu

Received 21.2.11; accepted 12.5.11

Laboratory evolution studies provide fundamental biological insight through direct observation of the evolution process. They not only enable testing of evolutionary theory and principles, but also have applications to metabolic engineering and human health. Genome-scale tools are revolutionizing studies of laboratory evolution by providing complete determination of the genetic basis of adaptation and the changes in the organism's gene expression state. Here, we review studies centered on four central themes of laboratory evolution studies: (1) the genetic basis of adaptation; (2) the importance of mutations to genes that encode regulatory hubs; (3) the view of adaptive evolution as an optimization process; and (4) the dynamics with which laboratory populations evolve.
Molecular Systems Biology **7**: 509; published online 5 July 2011; doi:10.1038/msb.2011.42
Subject Categories: metabolic and regulatory networks
Keywords: epistasis; flux-balance analysis; metabolic engineering; mutation; regulatory hub

Introduction

Due to the lack of experimental approaches, the study of the process of evolution has been mostly limited to theoretical studies. Thus, this field is rich with abstract concepts that were for a time difficult or impossible to test experimentally. Recent advances in DNA sequencing, high-throughput technologies, and genetic manipulation systems have enabled empirical studies that directly characterize the molecular and genomic bases of evolution. Specifically, these tools allow the direct measurement and analysis of short-term evolutionary processes, and thus the testing of conceptual evolutionary models and theories (Wagner, 2008).

Microbes are well suited to adaptive laboratory evolution (ALE) studies for a number of reasons: fast generation times, repeatability, the ease of maintaining large population sizes, and the ability to store populations for later examination (Elena and Lenski, 2003). In the wild, evolution has diverse causes such as environmental change or isolation of small populations. ALE studies, also referred to as experimental evolution studies, seek to observe microbes under controlled scenarios in which evolution is expected to occur. While the controlled environments may not always precisely represent those in the wild, ALE studies have proven successful in grounding evolutionary theories to the actual molecular and mechanistic bases of evolution. In this review, we will highlight a few central lessons obtained through microbial ALE studies thus far. First, genomic sequencing can determine the complete set of mutations responsible for an evolved phenotype, and has led to the discovery that interactions between these mutations are very common. Second, adaptive mutations frequently target regulatory mechanisms. Third, principles of systems-level optimization underlie the genetic changes seen in adaptive evolution, and with a systems-level understanding, these optimization principles can be harnessed for the purposes of metabolic engineering. Fourth, mutant subpopulations of improved fitness invariably arise in growing populations, but their dynamics in the population are complicated due to factors such as natural selection, clonal interference, drift, and frequency-dependent selection.

Full genomic changes occurring during ALE can be directly observed

Darwin's theory of evolution preceded any knowledge of the molecular mechanisms of heredity, and the development of evolutionary theory since has frequently omitted these mechanisms. Yet, the molecular details are at the heart of the evolutionary process and, thus, of central interest. The development of next-generation sequencing technology has at last provided the means to study the molecular basis of evolution on a genome scale (Metzker, 2010; Brockhurst *et al*, 2011). Many recent studies have utilized next-generation technologies to find mutations on a genome-wide basis (Albert *et al*, 2005; Friedman *et al*, 2006; Herring *et al*, 2006; Velicer *et al*, 2006; Gresham *et al*, 2008; Barrick *et al*, 2009; Conrad *et al*, 2009; Araya *et al*, 2010; Atsumi *et al*, 2010; Charusanti *et al*, 2010; Kishimoto *et al*, 2010; Lee and Palsson, 2010; Lee *et al*, 2010).

Mutations found by whole-genome resequencing (WGS) of several non-mutator evolved clones from *E. coli* ALE experiments as compared with the sequence of the starting strain are summarized in Figure 1. In the studies summarized in Figure 1 (Herring *et al*, 2006; Barrick *et al*, 2009; Conrad *et al*, 2009; Charusanti *et al*, 2010; Kishimoto *et al*, 2010; Lee and Palsson, 2010), single-nucleotide mutations were the most common type of mutation observed (61% of observed mutations) and c>t/g>a mutations were disproportionally observed (48 of 111 SNPs). Deletions (29%), insertions (7%), and insertion sequence movements (3%) were also commonly observed.

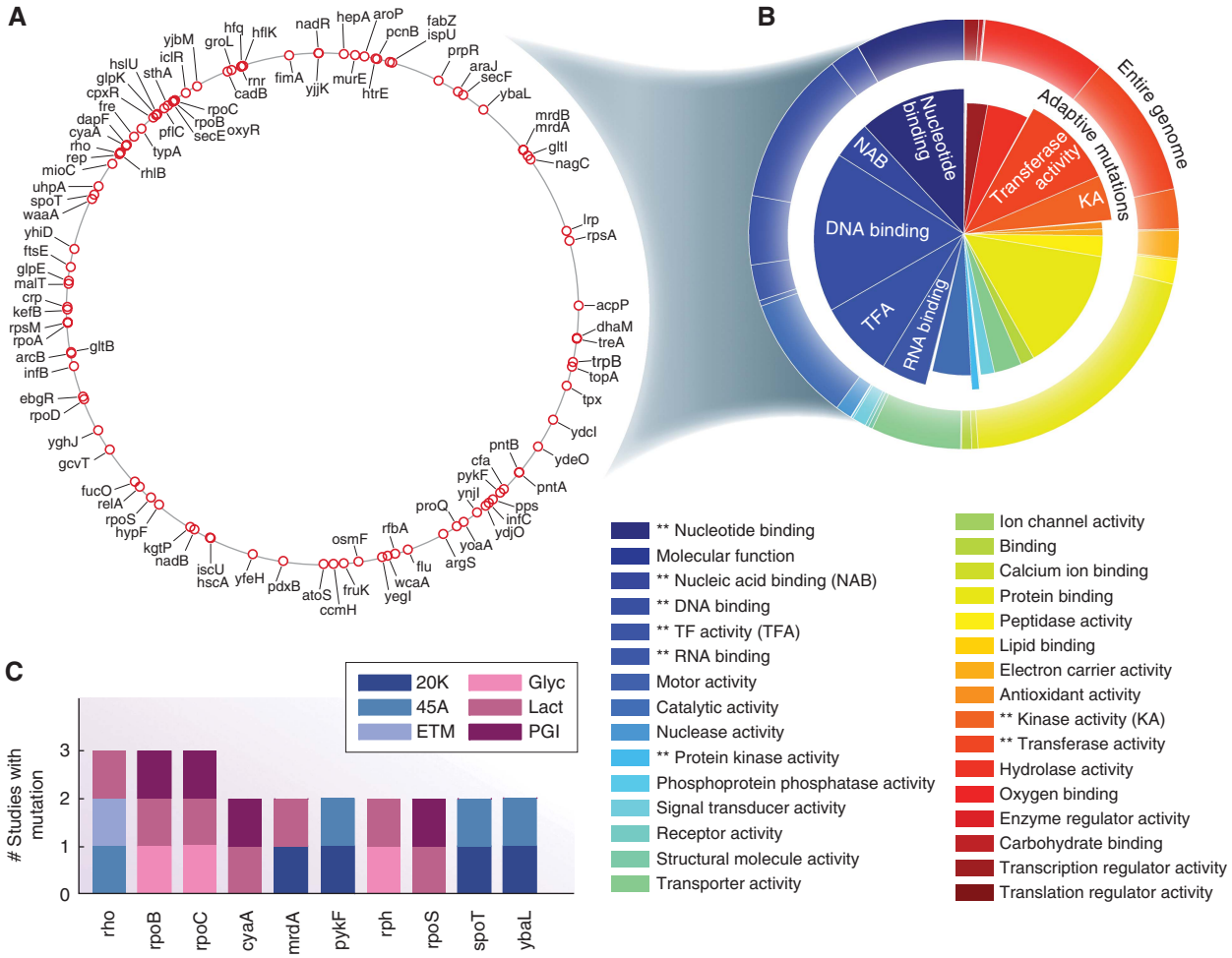

Figure 1 Intragenic mutations identified in *E. coli* ALE studies. (**A**) Single-nucleotide substitutions, insertions, and deletions found within the open reading frames by whole-genome sequencing in multiple *E. coli* ALE studies (Herring *et al*, 2006; Barrick *et al*, 2009; Conrad *et al*, 2009; Charusanti *et al*, 2010; Kishimoto *et al*, 2010; Lee and Palsson, 2010) are shown on a circular representation of the *E. coli* chromosome. (**B**) The set of genes displayed on the *E. coli* chromosome was subjected to enrichment analysis for Gene Ontology Slim (GOslim) categories (Camon *et al*, 2004). Wedges that protrude outward represent statistically enriched GOslim categories (also marked by ** in the legend). (**C**) Genes that were mutated in multiple studies are shown. 20K=growth on glucose minimal medium for 20 000 generations (Barrick and Lenski, 2009), 45A=adaptation to high temperature (Kishimoto *et al*, 2010), ETM=adaptation of ethanol tolerance (Goodarzi *et al*, 2009), Glyc=growth on glycerol minimal medium (Herring *et al*, 2006), Lact=growth on lactate minimal media (Conrad *et al*, 2009), and PGI=growth on glucose minimal media following the deletion of *pgi* (Charusanti *et al*, 2010).

WGS of ALE strains has uncovered the single-nucleotide substitutions, small insertions and deletions (indels), large genomic duplications (Herring *et al*, 2006; Conrad *et al*, 2009), large deletions (Barrick *et al*, 2009; Charusanti *et al*, 2010), and transposable element insertions (Gabriel *et al*, 2006; Gresham *et al*, 2008; Atsumi *et al*, 2010) acquired during adaptation. WGS of several strains evolved in replicate shows the extent to which evolutionary genetic paths are constrained. Parallel evolution also reveals the mutations conferring the largest benefit, since only the few mutations conferring large fitness gains should be repeatedly selected for in multiple replicates (Cooper *et al*, 2003; Elena and Lenski, 2003; Herring *et al*, 2006; Conrad *et al*, 2009; Charusanti *et al*, 2010). In another approach, WGS has been applied in a longitudinal manner, in which the genome was sequenced at various time points. This approach was applied to *E. coli* evolved in glucose minimal medium for 40 000 generations (Barrick *et al*, 2009), and to *E. coli* adapted to growth in successively higher temperatures for 7500 generations (Kishimoto *et al*, 2010). In both cases, the rate of mutation appearance increased in the later phase of the adaptation due to the development of a mutator phenotype. Surprisingly, the *E. coli* strain grown in glucose minimal medium acquired beneficial mutations at a nearly linear rate over 20 000 generations before obtaining a mutator phenotype (Barrick *et al*, 2009).

Through the use of allelic replacement, the fitness conferred by a mutation can be assessed using growth rate measurements or competition experiments (Applebee *et al*, 2008). Allelic replacement experiments have shown that most mutations found in non-mutator populations are beneficial when introduced to the ancestral strain (Cooper *et al*, 2003; Herring *et al*, 2006; Giraud *et al*, 2008; Conrad *et al*, 2009;

Goodarzi *et al*, 2009; Auriol *et al*, 2011; Tremblay *et al*, 2011). In addition to enhanced fitness, single allelic replacement has shown other genotype–phenotype relationships of adaptive mutations. For instance, it helped to link a mutation of a gene involved in Wsp signaling to the 'wrinkly spreader' phenotype in *Pseudomonas fluorescens* (Knight *et al*, 2006; Bantinaki *et al*, 2007), to show contribution of mutations to the *spoT* regulatory gene to expression patterns found in evolved *E. coli* (Cooper *et al*, 2003), and to show the effect of an adaptive mutation to a *Geobacter sulfurreducens* gene encoding a periplasmic cytochrome on Fe_2O_3 reduction (Tremblay *et al*, 2011). Furthermore, allelic replacement has been used to iteratively reconstruct the complete multiple-mutation genotype found by WGS of the adapted strain. Complete recovery of the adapted phenotype in the reconstructed strain suggests that all relevant mutations have been identified (Herring *et al*, 2006; Conrad *et al*, 2009; Goodarzi *et al*, 2009; Lee and Palsson, 2010).

Allelic replacement experiments have shown that some mutations exhibit a neutral or deleterious effect on fitness when introduced into the ancestral background, but appear to confer positive fitness in the presence of other mutations acquired during adaptation. These interactions between mutations are examples of epistasis (non-additive or non-multiplicative fitness effects of multiple mutations) and appear to be a very common feature observed in ALE and genetic interaction studies (Dykhuizen and Hartl, 1980; Herring *et al*, 2006; Spencer *et al*, 2007; Applebee *et al*, 2008; Beaumont *et al*, 2009; Conrad *et al*, 2009; Charusanti *et al*, 2010; Costanzo *et al*, 2010; He *et al*, 2010; Lee and Palsson, 2010). Epistatic mutations are sometimes compensatory, providing a remedy for complications introduced to the cell by earlier adaptive mutations (Maisnier-Patin and Andersson, 2004). Alternatively, sometimes the fitness attributed to the mutation is positive only in a fast growing cell (Chou *et al*, 2009). Epistatic interactions can severely constrain viable evolutionary trajectories, since there may be relatively few trajectories in which fitness increases monotonically with each single mutation (Weinreich *et al*, 2006; Poelwijk *et al*, 2007).

WGS and allelic replacement experiments determine the genetic basis of evolution. However, in the cases where many mutations are found (e.g., in populations with high mutation rates), it is difficult or unrealistic to iteratively construct allelic replacement collections. Two recent methods may help to overcome this limitation. First, multiplex automated genome engineering provides an inexpensive and high-throughput platform for introducing specific sequence diversity into the population (Wang *et al*, 2009). Under this approach, a set of synthetic oligonucleotides representing the mutations found by WGS randomly undergoes recombination with genomic DNA, creating a library of allelic replacement strains. Second, array-based discovery of adaptive mutations (ADAM) uses genetic foot printing to selectively identify mutations that provide a competitive advantage to the cell (Goodarzi *et al*, 2009). When reconstructing phenotypes through allelic replacement, ADAM allows the researcher to avoid the wasted effort of reconstructing non-beneficial mutations. These methods have the potential to accelerate and increase the scope of ALE studies.

Broad and systemic regulatory changes found in ALE strains

Microarrays have been used extensively to measure the global gene expression phenotype of evolved bacteria. Routine analyses exist to retrieve information from the large, complex data sets generated by these experiments (Eisen *et al*, 1998; Cui and Churchill, 2003). Clustering can group genes based on similar expression patterns across multiple samples. Genes showing statistically significant differential expression can further be analyzed for enrichment within functional categories or established regulons. Sometimes, the significant gene functions that come out of these analyses have obvious relevance to the strain in the adaptive growth condition, while other times the meaning of the expression changes is hard to interpret based on our current understanding of the function of the mutated gene product. Measurement of expression changes in multiple replicate evolved lines can identify the changes that are more likely to be relevant to adaptation by finding genes that are consistently differentially expressed. Likewise, fitness profiling using gene overexpression and gene knockout libraries can provide coarse-grain determination of the specific gene expression changes that are adaptive for a strain (Goodarzi *et al*, 2010).

Sometimes, a large degree of parallelism is observed in the gene expression changes of replicate evolved populations (Cooper *et al*, 2003; Fong *et al*, 2005b; Stoebel *et al*, 2009). Sequencing of these populations shows that similar regulatory genes acquire mutations within replicate populations. A single mutation to a regulatory hub can result in broad-scale changes to the expression state of cells in the mutated population (Philippe *et al*, 2007). In general, such mutations of large effect were expected to be deleterious (Fisher, 1930). However, ALE experiments have shown that mutations to regulatory hubs can provide a large increase to fitness (Cooper *et al*, 2003; Crozat *et al*, 2005; Herring *et al*, 2006; Giraud *et al*, 2008; Conrad *et al*, 2010). Due to their large benefit, mutations to regulatory hubs are common and are frequently found in replicate populations (Figure 1B and C).

One of the most central transcriptional regulatory hubs is likely the RNA polymerase (RNAP) itself (Klein-Marcuschamer *et al*, 2009). ALE studies subjecting the *E. coli* K-12 MG1655 strain to continuous exponential growth in minimal media have found that the genes encoding the RNAP core enzyme are frequently mutated (Friedman *et al*, 2006; Herring *et al*, 2006; Conrad *et al*, 2009, 2010; Charusanti *et al*, 2010; Figure 1C). Mutations to the *rpoC* gene, encoding the RNAP β′ subunit, were found in ∼80% of MG1655 strains that were adapted to glycerol M9 minimal media (Conrad *et al*, 2010). These *rpoC* mutations resulted in a 60% increase to growth rate in glycerol minimal medium when introduced to the ancestor strain, while simultaneously changing the expression pattern of 20–27% of genes in the genome. These genes fell into certain functional classifications, suggesting that elements of the RNAP have transcription factor-like activity. *In vitro* assays indicated that separate adaptive *rpoC* mutations resulted in similar changes in transcription kinetics of the mutants, including decreased open-complex half-life at a sensitive promoter, decreased pausing half-life, and increased elongation rate. Thus, these mutations to the core enzyme may

affect broad and complex changes to the transcriptome by altering the transcription kinetics.

In bacteria, sigma factors bind core RNAP, providing preference for specific promoters and assisting in open-complex formation. There are several different types of sigma factors that compete with each other for binding to free RNAP core enzyme to form the holoenzyme. Each sigma factor recognizes different sets of promoters, and thus affects the global transcription profile of the bacteria. The *rpoS* gene of *E. coli*, which encodes the stationary phase sigma factor and is needed for the global stress response (Weber *et al*, 2005), is frequently affected by mutations (Notley-McRobb *et al*, 2002; Cooper *et al*, 2003; Ferenci, 2008; Spira *et al*, 2008; Conrad *et al*, 2009, 2010; Charusanti *et al*, 2010; Wang *et al*, 2010). Whereas binding of the RpoS sigma factor to the RNAP core enzyme results in a global stress response, the predominant sigma factor during exponential growth is Sigma 70, which is needed for transcription of many metabolic and growth-associated genes. The competition between RpoS and Sigma 70 for RNAP core enzyme sets up a tradeoff between stress survival and robust growth that has been termed the 'self-preservation and nutritional competence' (SPANC) balance (Ferenci, 2005). Therefore, mutations to *rpoS* and its regulators can serve to adjust the SPANC balance by reducing or eliminating the size of the RpoS pool that is able to bind RNAP core enzyme. Under conditions of strong growth pressures and weak environmental stress, such mutations are expected to be favorable.

These examples of regulatory hub mutations demonstrate that while the transcriptional regulatory network of *E. coli* is not wired to optimally respond to all environments, the network structure allows a single mutation to a hub to rewire the network into a more efficient state. While the process of discovering these mutations by WGS and introducing them into the ancestor through allelic replacement has become routine, careful characterization of the genotype–phenotype relationship of mutations to regulatory genes is difficult due to the large scale and complexity of phenotypic consequences. Understanding can be further impeded by unknown secondary functions of the mutated gene. These challenges to uncovering the genotype–phenotype relationship of mutations to regulatory hubs can to some extent be overcome using genome-scale fitness profiling (Goodarzi *et al*, 2010) and computational models (Lewis *et al*, 2010). The hard work of uncovering the genotype–phenotype relationship may be rewarded by insights into secondary and unknown functions of the mutated gene (Copley, 2003; James and Tawfik, 2003). ALE can thus serve as a discovery tool where the hypotheses about function are generated through the evolutionary process.

Cellular optimization in ALE

Models of evolution frequently invoke the concept of adaptive landscapes (Orr, 2005). These are often depicted graphically as three-dimensional landscapes in which the sequence space serves as a plane on which to map the fitness of each genotype. Adjacent points on this plane represent similar sequences. The landscape can be smooth (Figure 2A), with a single optimum, or rugged (Figure 2B), with multiple optima. In a rugged fitness landscape, accessible optima may depend on the starting point within the landscape. Thus, there may be a diversity of evolutionary outcomes. Results of ALE experiments suggest that bacteria encounter both smooth landscapes, in which the population is constrained to acquire specific mutations in an exact order (Weinreich *et al*, 2006; Poelwijk *et al*, 2007; Stoebel *et al*, 2009), and rough landscapes, in which WGS shows few patterns in the genes that acquire mutations among replicate populations (Gresham *et al*, 2008; Conrad *et al*, 2009; Wang *et al*, 2010). The continual fitness improvement of *E. coli* over tens of thousands of generations of adaptive evolution, albeit with diminishing returns, suggests that populations may only asymptotically approach local optima on the fitness landscape (Barrick *et al*, 2009).

The concept of traversing the fitness landscape depends on the assumption that cells will modify their genotype to optimize their fitness in the given environment. Therefore, insight into cellular optimization is gained through the use of mathematical models. When the organism's metabolic network and biomass composition are known, flux-balance analysis (FBA; Orth *et al*, 2010) and other constraint-based modeling methods can be used to predict optimal phenotypes and their corresponding flux distribution based on measured nutrient uptake rates. Essentially, FBA calculates the maximum rate that biomass can be produced given the uptake rates of nutrients. FBA results can be presented on phenotypic phase-plane (PhPP) plots of model-predicted optimal biomass production (growth rate) versus carbon source uptake rate (SUR) and oxygen uptake rate (OUR). The graph of optimal OUR as a function of SUR is linear under physiological constraints. When mapped on to the PhPP, this line defines the line of optimality (LO) describing the most efficient ratio of SUR and OUR for biomass synthesis. Under several conditions, the experimentally measured *E. coli* phenotype corresponds to the LO of the PhPP (Edwards *et al*, 2001). When *E. coli* growth is not consistent with the LO, populations migrate toward the LO through adaptive evolution (Ibarra *et al*, 2002; Figure 2C). Migration toward the model-predicted optimal phenotype has also been observed in *Lactobacillus plantarum* adapting to growth in glycerol (Teusink *et al*, 2009). These studies show that cells will optimize their metabolic efficiency consistent with the demands from the selective pressure.

Achieving an optimal metabolic flux state requires sufficient expression of enzymes, yet proteins are costly for the cell to express (Kurland and Dong, 1996; Scott *et al*, 2010) due to limited quantities of free RNAP enzymes and ribosomes in the cell (Stoebel *et al*, 2008). Therefore, there is expected to be an optimal protein expression pattern. The principle of optimization of protein expression is exemplified by a cost-benefit analysis of Lac operon expression in the presence of lactose (Dekel and Alon, 2005). The mathematical model developed in this study shows that the optimal Lac expression level varies with the lactose concentration in the media. ALE experiments revealed that strains that initially expressed suboptimal levels of LacZ for a given lactose concentration rapidly evolved to produce the predicted optimal amount of LacZ, demonstrating optimization of gene expression for a single operon. Efforts have also been made to show economy of protein expression on the genome scale. Constraint-based modeling analyses of

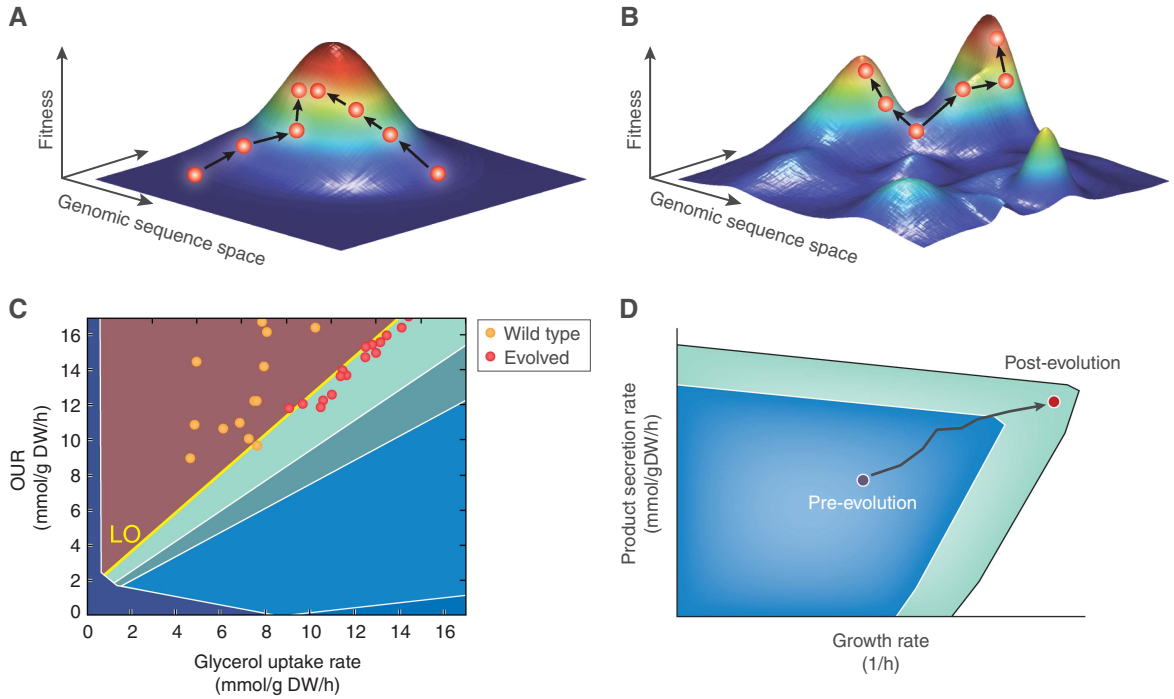

Figure 2 Optimality principles in adaptive evolution. (**A**) A smooth fitness landscape consisting of a single peak. Circles represent points on the landscape and arrows indicate the pathway of genetic change through the landscape. On a smooth landscape, there is a tendency for evolutionary convergence toward the single optimum, regardless of the starting point on the landscape. (**B**) A rough fitness landscape consists of multiple peaks. With multiple optima, there tends to be evolutionary divergence, sometimes even when starting from the same location on the landscape. (**C**) A phenotypic phase plane is a representation of how two fluxes in a metabolic network relate to each other and affect *in silico*-predicted optimal growth. Distinct planes are represented by several colors. Here, the line of optimality (LO, yellow) defines the ratio of glycerol uptake rate to oxygen uptake rate that leads to optimal biomass production. On glycerol, wild-type *E. coli* initially has a phenotype that maps to a suboptimal region of the portrait. After a growing for several hundred generations on glycerol, the *E. coli* phenotype migrates to the line of optimality. (**D**) Optimality principles can used to design strains of bacteria in which growth at the maximal rate requires the secretion of a product of interest. When subjected to ALE, the designed strains increase both their growth rate and product secretion rate. The two colored regions indicate accessible flux states before and after evolution.

ALE outcomes have shown that evolved strains exhibit a general pattern of increased expression of genes and proteins associated with the optimal flux distribution, and decreased expression of genes and proteins associated with unused pathways (Becker *et al*, 2007; Lewis *et al*, 2010). Thus, the optimization of expression levels is a significant goal of adaptation.

The concept of evolution toward optimal phenotypes has a practical application in metabolic engineering. Strain engineering aims to coerce a cell to produce large amounts of product. However, this often inhibits growth. As a consequence, production strains often lose their ability to excrete product as the cellular fitness improves. Algorithms have been developed to predict gene deletions that couple fast growth with metabolite excretion (Burgard *et al*, 2003; Pharkya *et al*, 2004; Patil *et al*, 2005; Lun *et al*, 2009; Kim and Reed, 2010). In these growth-coupled designs, the metabolic pathways of an organism are altered such that the predicted optimal growth rate requires the secretion of the desired product (Figure 2D). Thus, adaptation for a faster growth rate can result in higher product yields, and the production strains become more stable. Many growth-coupled designs have been predicted *in silico* (Burgard *et al*, 2003; Bro *et al*, 2006; Feist *et al*, 2010) and improved product yield following the adaptation has been observed in designs for lactate and ethanol producers in the laboratory (Fong *et al*, 2005a; Trinh and Srienc, 2009).

Currently, the most extensive and detailed optimization models used to assess ALE experiments are constraint-based models (Feist *et al*, 2009). However, these focus primarily on metabolism, which accounts for less than one third of the genes in *E. coli*. Therefore, mechanistic conclusions drawn from metabolic models alone might be incomplete, since they do not directly include important biochemical mechanisms such as regulation, cellular maintenance and protection, and other non-metabolic functions. Efforts are being made to incorporate the functions of transcriptional regulation (Gianchandani *et al*, 2006; Chandrasekaran and Price, 2010) and the transcription/translation machinery (Thiele *et al*, 2010), thereby increasing the genome coverage by these models. In addition, alternative modeling frameworks are also being developed to model multiple cellular processes beyond metabolism (Covert *et al*, 2008; Molenaar *et al*, 2009). However, mutations in regulatory proteins still will pose a challenge regardless of the modeling framework. Mutations may alter the binding kinetics and motifs of these proteins, thereby changing the topology of the regulatory network. Despite this potential challenge, the assessment of high-throughput data in the model context may allow for the elucidation of the topological changes through an analysis of genes and proteins that failed to behave as predicted by the optimization models. Therefore, the use of optimization models in the assessment of ALE experiments will

complement the analysis and provide hypotheses to possibly guide the difficult assessment of individual mutations.

Mutational dynamics and diversity within evolving populations

When placed into a challenging environment, natural selection acts on bacteria to select for populations harboring mutations that provide improved growth or survival. ALE studies have addressed the dynamics with which these mutations are acquired and propagated through a population. In many cases, there are expansions of diversity as additional beneficial mutations appear in the population, followed by contractions in which a sub-population of superior fitness eliminates the inferior alleles. In other cases to be discussed, mechanisms for stable population diversity have been observed.

Mutations provide the raw material for population diversity and adaptation. Thus, the mutation rate has profound effects on population diversity and evolution dynamics. A low mutation rate will result in slow discovery of adaptive mutations, while a fast mutation rate will increase the rate of occurrence of both deleterious and beneficial alleles, and result in more diverse populations. Mutator phenotypes frequently appear during adaptation and are characterized by an increased mutation rate, including a higher fraction of synonymous mutations (Sniegowski *et al*, 1997; Gresham *et al*, 2008; Barrick *et al*, 2009; Kishimoto *et al*, 2010). Mutator phenotypes may be transiently useful for accelerated discovery of a beneficial mutation, allowing the mutator allele to 'hitchhike' during an adaptive sweep. However, they may be harmful over the long term due to accumulation of deleterious mutations (increased genetic load) (Giraud *et al*, 2001; Elena and Lenski, 2003). Furthermore, the mutation rate varies with environment and exposure to antibiotics (Notley-McRobb *et al*, 2003; Sakai *et al*, 2006; Kohanski *et al*, 2010). Frequent DNA damage or oxidative stress due to environmental stresses could provide a mechanism by which the mutation rate increases during a stressful condition (Notley-McRobb and Ferenci, 1999b; Notley-McRobb *et al*, 2003).

The mutation rate of a bacterial strain can be measured by a fluctuation test (Luria and Delbruck, 1943), in which the mutation rate is approximated by frequency of phage resistance or antibiotic resistance in an initially sensitive population, or by a mutation accumulation experiment (Kibota and Lynch, 1996; Lind and Andersson, 2008), where serial passage of single cell bottlenecks is used to measure the rate of appearance of mutations in a lineage over time. However, rates measured by either of these methods greatly over-approximate the rate at which mutations appear at a detectable frequency in laboratory evolution cultures (Lenski and Travisano, 1994). This discrepancy points to the importance of mutational dynamics within the population. First, many mutations are deleterious and will often be removed by selective pressures from the culture before reaching detectable frequency. Second, in large bacterial populations, multiple sub-populations can exist that possess differing beneficial alleles. When one sub-population gains and maintains a substantial fitness advantage over the rest of the population, it can increase its fraction in the population and even sweep the population (Muller, 1932), eliminating other sub-populations completely or to undetectable levels in a phenomenon known as clonal interference. Clonal interference may also impose a 'speed limit' on adaptation, in which an increase in the mutation only results in an increase in adaptation rate to a certain point, after which selection, not the supply of beneficial mutations, becomes rate limiting (Arjan *et al*, 1999). Finally, many laboratory evolution experiments include 'bottlenecks' in which a random fraction of the population is transferred to fresh medium to maintain exponential rates of growth. In serial passage experiments, the fraction of the population discarded tends to be very high. Thus, even beneficial mutations are routinely lost to drift. In chemostat experiments, cells are discarded continuously as quickly as new cells replicate. Therefore, the experiments are free of bottlenecks, with higher numbers of mutants remaining in the population.

To demonstrate the effect of drift and clonal interference in a serial passage experiment, consider *E. coli* grown in glycerol M9 minimal medium where cultures are maintained in continuous exponential growth by transferring a small amount of the culture to fresh media once each day before late exponential phase is reached (Herring *et al*, 2006). If there is an average of 2.5×10^{-3} mutations across the genome per cell division (Drake, 1991) and 10^{11} cells in a 250-ml culture of *E. coli* at the time of passage, there will be $\sim 2.5 \times 10^{8}$ mutations appearing during a single passage cycle. Because only $\sim 10^{8}$ cells or fewer are passaged, only a small fraction of the mutants avoids extinction by drift each passage. However, mutants with a growth advantage will replicate more often before the next passage, giving a higher probability of surviving subsequent passages. For example, during adaptation in glycerol M9 minimal medium, an *rpoC* mutant was identified that increased the growth rate by 60%. Because the *rpoC* mutant sub-population grows much faster than the rest of the population, a single *rpoC* mutant cell passaged among 2×10^{8} cells of wild-type growth rate will increase to represent on average 46 of the 2×10^{8} cells of the subsequent passage. Within 144 h, >98% of the population is expected to have the mutant *rpoC* allele (Figure 3A), and eventually the small, shrinking sub-population with the wild-type allele will be lost to drift. As an example of clonal interference, the population sweep of *rpoC* mutants will be slowed down if it enters the population in the presence of another mutant sub-population that grows at a rate intermediate between the wild-type and *rpoC* mutant growth rates (Figure 3B). As in the previous scenario, the *rpoC* mutant eventually sweeps the culture, this time eliminating a lesser beneficial allele from the population.

As the above examples suggest, homogeneous populations are often found in serial passage experiments, where strong selective pressures favor sweeps and small bottlenecks purge diversity (Herring *et al*, 2006; Conrad *et al*, 2009; Lee and Palsson, 2010). In contrast, sweeps are more rarely observed in chemostat ALE experiments (Kao and Sherlock, 2008). High levels of diversity within a single niche rapidly appear in chemostat conditions (Notley-McRobb and Ferenci, 1999a, b; Maharjan *et al*, 2006; Wang *et al*, 2010). Diversity within chemostat cultures is unintuitive, since, as the competitive exclusion principle (Hardin, 1960) states, if two mutants are

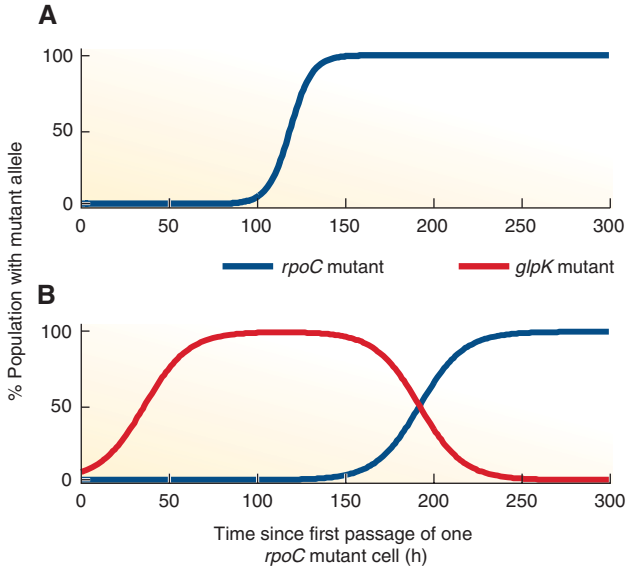

Figure 3 Simulation of the evolution dynamics of an *rpoC* mutant. Exponential growth in the number of cells N at time t is given by the function $N(t)=N_0e^{kt}$ where N_0 represents N at $t=0$ and k represents the growth rate. (**A**) The fraction of the population represented by an *rpoC* mutant ($k_{rpoC}=0.43$) in an otherwise wild-type population ($k_{wt}=0.27$) and which initially exists at a ratio of one *rpoC* mutant cell per 2×10^8 wild-type cells is shown. (**B**) The scenario in this graph is the same as before, except initially the one *rpoC* mutant (blue) cell exists in a population of 10^8 cells that is 95% wild-type and 5% *glpK* mutant (red, $k_{mut}=0.35$). This situation results in a situation of clonal interference between two beneficial alleles.

competing for the same resource, one mutant should win out if it has even a small edge of the other mutant. However, diversity may exist given the large number of generations needed to edge out a slightly less fit competitor, especially when even a nearly down-and-out competitor can make a comeback in the population through acquiring additional mutations (Kao and Sherlock, 2008). High population diversity may also be due to elevated mutation rates in slow-growing chemostat cultures (Notley-McRobb *et al*, 2003) and a lack of bottlenecks.

The stable coexistence of two clonal sub-populations on a single resource has been witnessed in ALE experiments. A number of phenomena can explain these populations. Evolution of cross-feeding commensalisms (Pfeiffer and Bonhoeffer, 2004) allow for a first sub-population to consume a primary nutrient (i.e., glucose) and excrete a secondary compound (i.e., acetate) that is metabolized by a second sub-population (Helling *et al*, 1987; Rosenzweig *et al*, 1994; Rozen and Lenski, 2000; Kinnersley *et al*, 2009). A related example of niche diversity is a 'fast switch' (FS) phenotype that was found to evolve in minimal media containing both glucose and acetate (Friesen *et al*, 2004; Spencer *et al*, 2007, 2008). The FS phenotype switches to metabolizing acetate faster than the ancestor when glucose is exhausted (diauxic switch). However, this comes at the cost of slower growth in glucose. In this scenario, the FS strain is able to stably coexist with a non-FS phenotype through frequency-dependent selection. Finally, ALE resulted in the development of stability of a syntrophic mutualism in which *Methanococcus maripaludis* fed on hydrogen produced in the fermentation of lactate by *Desulfovibrio vulgaris*. The consumption of hydrogen by *M. maripaludis* benefited *D. vulgaris* by providing the thermodynamic driving force for the fermentation of lactate that is necessary for growth (Hillesland and Stahl, 2010).

ALE studies have demonstrated that evolving populations often exhibit dynamics that are more complicated than the clonal replacement model suggests would occur. Furthermore, the discovery of stable diverse bacterial populations within a single niche due to frequency-dependent selection or mutualistic interactions has challenged the competitive exclusion principle (Hardin, 1960). While further characterization of diversity within chemostat and serial passage populations is of interest, most ALE studies using WGS have focused on sequencing clonal samples of the population. Recent work has described a detailed tracking of genetic diversity in a serial passage population throughout adaptation (Barrick and Lenski, 2009), and advances in sequencing technology may soon allow high resolution measurements of population diversity on a genome scale.

Conclusions

Experience from microbial ALE experiments under a variety of selective pressures is accumulating rapidly. Here, we have reviewed what has been learned from ALE experiments about the genetic basis of adaptation, regulatory changes, optimality in evolving populations, and population dynamics. Importantly (1) The genetic basis of adaptation can now be determined by next-generation sequencing technologies and can be reconstructed using allelic replacement. Epistasis is frequently observed between multiple reconstructed mutants and may indicate either compensatory effects for the earlier mutation or a dependence of the fitness benefit of the allele on the cellular growth rate. (2) Broad-acting mutations to regulatory hubs, including the RNAP itself, are often highly beneficial. (3) The target phenotype of an evolving population is likely guided by principles of optimality. Models based on methods such as FBA and cost-benefit analysis can predict aspects of these target phenotypes. (4) Diversity in evolving populations is related to the mutation rate, the fitness effects of beneficial mutations, and the size of bottleneck involved. The minimization of diversity through clonal replacement is more common in serial passage experiments than in chemostat experiments.

Mutations selected for in ALE experiments that cause broad expression changes can be thought of as systemic perturbation variables. For instance, a mutation that is found to increase tolerance to a toxic end product represents a particular change in the DNA that has a very particular systemic effect on network functions that is selected for. Biological function can be studied through combinations of environmental and genetic perturbations. The field typically considers perturbation variables to be environmental (chemical) or genetic (gene knockout). Dual perturbation experiments are used to discover gene function (Nichols *et al*, 2011). A collection of systemic perturbation variables resulting from ALE, that Figure 1 basically represents, gives the field of systems biology its own perturbation variables to work with. These perturbation variables, singly or in combination, confer some

optimality property on network functions that become known through the selection process implemented in ALE. Thus, they may be used to study gene function or as raw material for 'accelerated evolution' (Wang *et al*, 2009).

The frequent mutation of transcriptional regulators is consistent with recent evidence showing that regulatory networks evolve faster than other networks, such as genetic networks, protein interaction networks, and metabolic networks (Shou *et al*, 2011). Therefore, tools to improve understanding of changes to the regulatory network are of high priority. Omics measurements have shown great value for providing complete systems-level determination of changes in the state of the regulatory network. Yet, these data sets are impossible to analyze fully without a genome-scale context. As discussed earlier, use of genome-scale metabolic reconstructions as a scaffold has yielded insight into changes in the expression state of evolved populations (Lewis *et al*, 2010). Currently, progress is being made toward a comprehensive transcription regulatory network (TRN) for *E. coli* (Cho *et al*, 2008) reconstructed from genome-scale data sets. The comprehensive TRN will enable detailed analysis of transcription changes. Another form of genome-scale context is provided by fitness profiling (Goodarzi *et al*, 2010). This method determines the fitness change associated with a large increase or decrease in expression of a single gene, therefore, providing clues to changes in gene expression that are important for the mutant phenotype.

In addition to academic interest, ALE studies have a number of important practical applications. In the field of metabolic engineering, ALE is useful both for improved product yields in growth-coupled designs and for the study of tolerance to the strenuous conditions often implicit in industrial-scale growth (heat, pH, and toxic levels of alcohols) (Hughes *et al*, 2007; Atsumi *et al*, 2010; Horinouchi *et al*, 2010; Kishimoto *et al*, 2010). ALE studies can simulate the development of antibiotic resistance and drug sensitivity within the laboratory, thus indicating mechanisms by which dangerous drug-resistant bacterial strains can arise (Friedman *et al*, 2006). Other studies of adaptive evolution are moving beyond the laboratory and studying the evolution of bacteria within a human host (Zdziarski *et al*, 2010). Therefore, the coming years promise practical knowledge and applications to come from ALE experiments, in addition to important biological insights. Some of these insights will come through systems biology and the identification of systemic perturbation variables.

Acknowledgements

This work was supported by National Institutes of Health Grants R01 GM062791 and R01 GM057089.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Albert TJ, Dailidiene D, Dailide G, Norton JE, Kalia A, Richmond TA, Molla M, Singh J, Green RD, Berg DE (2005) Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*. *Nat Methods* **2:** 951–953

Applebee MK, Herrgard MJ, Palsson BO (2008) Impact of individual mutations on increased fitness in adaptively evolved strains of *Escherichia coli*. *J Bacteriol* **190:** 5087–5094

Araya CL, Payen C, Dunham MJ, Fields S (2010) Whole-genome sequencing of a laboratory-evolved yeast strain. *BMC Genomics* **11:** 88

Arjan JA, Visser M, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE (1999) Diminishing returns from mutation supply rate in asexual populations. *Science (New York, NY)* **283:** 404–406

Atsumi S, Wu TY, Machado IM, Huang WC, Chen PY, Pellegrini M, Liao JC (2010) Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Mol Syst Biol* **6:** 449

Auriol C, Bestel-Corre G, Claude JB, Soucaille P, Meynial-Salles I (2011) Stress-induced evolution of *Escherichia coli* points to original concepts in respiratory cofactor selectivity. *Proc Natl Acad Sci USA* **108:** 1278–1283

Bantinaki E, Kassen R, Knight CG, Robinson Z, Spiers AJ, Rainey PB (2007) Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. *Genetics* **176:** 441–453

Barrick JE, Lenski RE (2009) Genome-wide mutational diversity in an evolving population of *Escherichia coli*. *Cold Spring Harb Symp Quant Biol* **74:** 119–129

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461:** 1243–1247

Beaumont HJ, Gallie J, Kost C, Ferguson GC, Rainey PB (2009) Experimental evolution of bet hedging. *Nature* **462:** 90–93

Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protoc* **2:** 727–738

Bro C, Regenberg B, Forster J, Nielsen J (2006) In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab Eng* **8:** 102–111

Brockhurst MA, Colegrave N, Rozen DE (2011) Next-generation sequencing as a tool to study microbial evolution. *Mol Ecol* **20:** 972–980

Burgard AP, Pharkya P, Maranas CD (2003) Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* **84:** 647–657

Camon E, Magrane M, Barrell D, Lee V, Dimmer E, Maslen J, Binns D, Harte N, Lopez R, Apweiler R (2004) The Gene Ontology Annotation (GOA) database: sharing knowledge in uniprot with gene ontology. *Nucleic Acids Res* **32:** D262–D266

Chandrasekaran S, Price ND (2010) Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* **107:** 17845–17850

Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y, Palsson BO (2010) Genetic basis of growth adaptation of *Escherichia coli* after deletion of pgi, a major metabolic gene. *PLoS Genet* **6:** e1001186

Cho BK, Barrett CL, Knight EM, Park YS, Palsson BO (2008) Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc Natl Acad Sci USA* **105:** 19462–19467

Chou HH, Berthet J, Marx CJ (2009) Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. *PLoS Genet* **5:** e1000652

Conrad TM, Frazier M, Joyce AR, Cho BK, Knight EM, Lewis NE, Landick R, Palsson BO (2010) RNA polymerase mutants found through adaptive evolution reprogram *Escherichia coli* for optimal growth in minimal media. *Proc Natl Acad Sci USA* **107:** 20500–20505

Conrad TM, Joyce AR, Applebee MK, Barrett CL, Xie B, Gao Y, Palsson BO (2009) Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution

in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biol* **10:** R118

Cooper TF, Rozen DE, Lenski RE (2003) Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proc Natl Acad Sci USA* **100:** 1072–1077

Copley SD (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr Opin Chem Biol* **7:** 265–272

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, Prinz J, St Onge RP, VanderSluis B, Makhnevych T, Vizeacoumar FJ, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M *et al* (2010) The genetic landscape of a cell. *Science (New York, NY)* **327:** 425–431

Covert MW, Xiao N, Chen TJ, Karr JR (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* **24:** 2044–2050

Crozat E, Philippe N, Lenski RE, Geiselmann J, Schneider D (2005) Long-term experimental evolution in *Escherichia coli*. XII. DNA topology as a key target of selection. *Genetics* **169:** 523–532

Cui X, Churchill GA (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol* **4:** 210

Dekel E, Alon U (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436:** 588–592

Drake JW (1991) A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA* **88:** 7160–7164

Dykhuizen D, Hartl DL (1980) Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. *Genetics* **96:** 801–817

Edwards JS, Ibarra RU, Palsson BO (2001) In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol* **19:** 125–130

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* **95:** 14863–14868

Elena SF, Lenski RE (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev* **4:** 457–469

Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO (2009) Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol* **7:** 129–143

Feist AM, Zielinski DC, Orth JD, Schellenberger J, Herrgard MJ, Palsson BO (2010) Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metab Eng* **12:** 173–186

Ferenci T (2005) Maintaining a healthy SPANC balance through regulatory and mutational adaptation. *Mol Microbiol* **57:** 1–8

Ferenci T (2008) The spread of a beneficial mutation in experimental bacterial populations: the influence of the environment and genotype on the fixation of rpoS mutations. *Heredity* **100:** 446–452

Fisher RA (1930) *The Genetical Theory of Natural Selection*. Oxford: The Clarendon Press

Fong SS, Burgard AP, Herring CD, Knight EM, Blattner FR, Maranas CD, Palsson BO (2005a) In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol Bioeng* **91:** 643–648

Fong SS, Joyce AR, Palsson BO (2005b) Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res* **15:** 1365–1372

Friedman L, Alder JD, Silverman JA (2006) Genetic changes that correlate with reduced susceptibility to daptomycin in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **50:** 2137–2145

Friesen ML, Saxer G, Travisano M, Doebeli M (2004) Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in *Escherichia coli*. *Evolution* **58:** 245–260

Gabriel A, Dapprich J, Kunkel M, Gresham D, Pratt SC, Dunham MJ (2006) Global mapping of transposon location. *PLoS Genet* **2:** e212

Gianchandani EP, Papin JA, Price ND, Joyce AR, Palsson BO (2006) Matrix formalism to describe functional states of transcriptional regulatory systems. *PLoS Comput Biol* **2:** e101

Giraud A, Arous S, De Paepe M, Gaboriau-Routhiau V, Bambou JC, Rakotobe S, Lindner AB, Taddei F, Cerf-Bensussan N (2008) Dissecting the genetic components of adaptation of *Escherichia coli* to the mouse gut. *PLoS Genet* **4:** e2

Giraud A, Matic I, Tenaillon O, Clara A, Radman M, Fons M, Taddei F (2001) Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science (New York, NY)* **291:** 2606–2608

Goodarzi H, Bennett BD, Amini S, Reaves ML, Hottes AK, Rabinowitz JD, Tavazoie S (2010) Regulatory and metabolic rewiring during laboratory evolution of ethanol tolerance in *E. coli*. *Mol Syst Biol* **6:** 378

Goodarzi H, Hottes AK, Tavazoie S (2009) Global discovery of adaptive mutations. *Nat Methods* **6:** 581–583

Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, DeSevo CG, Botstein D, Dunham MJ (2008) The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet* **4:** e1000303

Hardin G (1960) The competitive exclusion principle. *Science (New York, NY)* **131:** 1292–1297

He X, Qian W, Wang Z, Li Y, Zhang J (2010) Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks. *Nat Genet* **42:** 272–276

Helling RB, Vargas CN, Adams J (1987) Evolution of *Escherichia coli* during growth in a constant environment. *Genetics* **116:** 349–358

Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK, Joyce AR, Albert TJ, Blattner FR, van den Boom D, Cantor CR, Palsson BO (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* **38:** 1406–1412

Hillesland KL, Stahl DA (2010) Rapid evolution of stability and productivity at the origin of a microbial mutualism. *Proc Natl Acad Sci USA* **107:** 2124–2129

Horinouchi T, Tamaoka K, Furusawa C, Ono N, Suzuki S, Hirasawa T, Yomo T, Shimizu H (2010) Transcriptome analysis of parallel-evolved *Escherichia coli* strains under ethanol stress. *BMC Genomics* **11:** 579

Hughes BS, Cullum AJ, Bennett AF (2007) Evolutionary adaptation to environmental pH in experimental lineages of *Escherichia coli*. *Evolution* **61:** 1725–1734

Ibarra RU, Edwards JS, Palsson BO (2002) Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420:** 186–189

James LC, Tawfik DS (2003) Conformational diversity and protein evolution–a 60-year-old hypothesis revisited. *Trends Biochem Sci* **28:** 361–368

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

Kibota TT, Lynch M (1996) Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. *Nature* **381:** 694–696

Kim J, Reed JL (2010) OptORF: optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst Biol* **4:** 53

Kinnersley MA, Holben WE, Rosenzweig F (2009) *E Unibus Plurum:* genomic analysis of an experimentally evolved polymorphism in *Escherichia coli*. *PLoS Genet* **5:** e1000713

Kishimoto T, Iijima L, Tatsumi M, Ono N, Oyake A, Hashimoto T, Matsuo M, Okubo M, Suzuki S, Mori K, Kashiwagi A, Furusawa C, Ying B-W, Yomo T (2010) Transition from positive to neutral in mutation fixation along with continuing rising fitness in thermal adaptive evolution. *PLoS Genet* **6:** e1001164

Klein-Marcuschamer D, Santos CN, Yu H, Stephanopoulos G (2009) Mutagenesis of the bacterial RNA polymerase alpha subunit for improvement of complex phenotypes. *Appl Environ Microbiol* **75:** 2705–2711

Knight CG, Zitzmann N, Prabhakar S, Antrobus R, Dwek R, Hebestreit H, Rainey PB (2006) Unraveling adaptive evolution: how a single

point mutation affects the protein coregulation network. *Nat Genet* **38:** 1015–1022
Kohanski MA, DePristo MA, Collins JJ (2010) Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Mol Cell* **37:** 311–320
Kurland CG, Dong H (1996) Bacterial growth inhibition by overproduction of protein. *Mol Microbiol* **21:** 1–4
Lee DH, Palsson BO (2010) Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a Nonnative carbon source, L-1,2-propanediol. *Appl Environ Microbiol* **76:** 4158–4168
Lee HH, Molla MN, Cantor CR, Collins JJ (2010) Bacterial charity work leads to population-wide resistance. *Nature* **467:** 82–85
Lenski RE, Travisano M (1994) Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc Natl Acad Sci USA* **91:** 6808–6814
Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, Adkins JN, Schramm G, Purvine SO, Lopez-Ferrer D, Weitz KK, Eils R, Konig R, Smith RD, Palsson BO (2010) Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol Syst Biol* **6:** 390
Lind PA, Andersson DI (2008) Whole-genome mutational biases in bacteria. *Proc Natl Acad Sci USA* **105:** 17878–17883
Lun DS, Rockwell G, Guido NJ, Baym M, Kelner JA, Berger B, Galagan JE, Church GM (2009) Large-scale identification of genetic design strategies using local search. *Mol Syst Biol* **5:** 296
Luria SE, Delbruck M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28:** 491–511
Maharjan R, Seeto S, Notley-McRobb L, Ferenci T (2006) Clonal adaptive radiation in a constant environment. *Science (New York, NY)* **313:** 514–517
Maisnier-Patin S, Andersson DI (2004) Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution. *Res Microbiol* **155:** 360–369
Metzker ML (2010) Sequencing technologies—the next generation. *Nat Rev* **11:** 31–46
Molenaar D, van Berlo R, de Ridder D, Teusink B (2009) Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol Syst Biol* **5:** 323
Muller HJ (1932) Some genetic aspects of sex. *Am Nat* **8:** 118–138
Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, Lee S, Kazmierczak KM, Lee KJ, Wong A, Shales M, Lovett S, Winkler ME, Krogan NJ, Typas A, Gross CA (2011) Phenotypic landscape of a bacterial cell. *Cell* **144:** 143–156
Notley-McRobb L, Ferenci T (1999a) Adaptive mgl-regulatory mutations and genetic diversity evolving in glucose-limited *Escherichia coli* populations. *Environ Microbiol* **1:** 33–43
Notley-McRobb L, Ferenci T (1999b) The generation of multiple co-existing mal-regulatory mutations through polygenic evolution in glucose-limited populations of *Escherichia coli*. *Environ Microbiol* **1:** 45–52
Notley-McRobb L, King T, Ferenci T (2002) rpoS mutations and loss of general stress resistance in *Escherichia coli* populations as a consequence of conflict between competing stress responses. *J Bacteriol* **184:** 806–811
Notley-McRobb L, Seeto S, Ferenci T (2003) The influence of cellular physiology on the initiation of mutational pathways in *Escherichia coli* populations. *Proc Biol Sci* **270:** 843–848
Orr HA (2005) The genetic theory of adaptation: a brief history. *Nat Rev* **6:** 119–127
Orth JD, Thiele I, Palsson BO (2010) What is flux balance analysis? *Nat Biotechnol* **28:** 245–248
Patil KR, Rocha I, Forster J, Nielsen J (2005) Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinformatics* **6:** 308
Pfeiffer T, Bonhoeffer S (2004) Evolution of cross-feeding in microbial populations. *Am Nat* **163:** E126–E135
Pharkya P, Burgard AP, Maranas CD (2004) OptStrain: a computational framework for redesign of microbial production systems. *Genome Res* **14:** 2367–2376
Philippe N, Crozat E, Lenski RE, Schneider D (2007) Evolution of global regulatory networks during a long-term experiment with *Escherichia coli*. *Bioessays* **29:** 846–860
Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445:** 383–386
Rosenzweig RF, Sharp RR, Treves DS, Adams J (1994) Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* **137:** 903–917
Rozen DE, Lenski RE (2000) Long-term experimental evolution in *Escherichia coli*. VIII. Dynamics of a balanced polymorphism. *Am Nat* **155:** 24–35
Sakai A, Nakanishi M, Yoshiyama K, Maki H (2006) Impact of reactive oxygen species on spontaneous mutagenesis in *Escherichia coli*. *Genes Cells* **11:** 767–778
Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science (New York, NY)* **330:** 1099–1102
Shou C, Bhardwaj N, Lam HY, Yan KK, Kim PM, Snyder M, Gerstein MB (2011) Measuring the evolutionary rewiring of biological networks. *PLoS Comput Biol* **7:** e1001050
Sniegowski PD, Gerrish PJ, Lenski RE (1997) Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* **387:** 703–705
Spencer CC, Bertrand M, Travisano M, Doebeli M (2007) Adaptive diversification in genes that regulate resource use in *Escherichia coli*. *PLoS Genet* **3:** e15
Spencer CC, Tyerman J, Bertrand M, Doebeli M (2008) Adaptation increases the likelihood of diversification in an experimental bacterial lineage. *Proc Natl Acad Sci USA* **105:** 1585–1589
Spira B, Hu X, Ferenci T (2008) Strain variation in ppGpp concentration and RpoS levels in laboratory strains of *Escherichia coli* K-12. *Microbiology (Reading, England)* **154:** 2887–2895
Stoebel DM, Dean AM, Dykhuizen DE (2008) The cost of expression of *Escherichia coli* lac operon proteins is in the process, not in the products. *Genetics* **178:** 1653–1660
Stoebel DM, Hokamp K, Last MS, Dorman CJ (2009) Compensatory evolution of gene regulation in response to stress by *Escherichia coli* lacking RpoS. *PLoS Genet* **5:** e1000671
Teusink B, Wiersma A, Jacobs L, Notebaart RA, Smid EJ (2009) Understanding the adaptive growth strategy of *Lactobacillus plantarum* by *in silico* optimisation. *PLoS Comput Biol* **5:** e1000410
Thiele I, Fleming RM, Bordbar A, Schellenberger J, Palsson BO (2010) Functional characterization of alternate optimal solutions of *Escherichia coli's* transcriptional and translational machinery. *Biophys J* **98:** 2072–2081
Tremblay PL, Summers ZM, Glaven RH, Nevin KP, Zengler K, Barrett CL, Qiu Y, Palsson BO, Lovley DR (2011) A c-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in *Geobacter sulfurreducens* revealed by adaptive evolution. *Environ Microbiol* **13:** 13–23
Trinh CT, Srienc F (2009) Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol. *Appl Environ Microbiol* **75:** 6696–6705
Velicer GJ, Raddatz G, Keller H, Deiss S, Lanz C, Dinkelacker I, Schuster SC (2006) Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. *Proc Natl Acad Sci USA* **103:** 8107–8112
Wagner A (2008) Neutralism and selectionism: a network-based reconciliation. *Nat Rev* **9:** 965–974
Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460:** 894–898
Wang L, Spira B, Zhou Z, Feng L, Maharjan RP, Li X, Li F, McKenzie C, Reeves PR, Ferenci T (2010) Divergence involving global regulatory gene mutations in an *Escherichia coli* population evolving under phosphate limitation. *Genome Biol Evol* **2:** 478–487
Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R (2005) Genome-wide analysis of the general stress response network in

Escherichia coli: sigmaS-dependent genes, promoters, and sigma factor selectivity. *J Bacteriol* **187:** 1591–1603

Weinreich DM, Delaney NF, Depristo MA, Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science (New York, NY)* **312:** 111–114

Zdziarski J, Brzuszkiewicz E, Wullt B, Liesegang H, Biran D, Voigt B, Gronberg-Hernandez J, Ragnarsdottir B, Hecker M, Ron EZ, Daniel R, Gottschalk G, Hacker J, Svanborg C, Dobrindt U (2010) Host imprints on bacterial genomes-rapid, divergent evolution in individual patients. *PLoS Pathog* **6:** pii: e1001078