Tracking Winners and Losers in E. coli Evolution Experiments

Within large bacterial populations, unexpected winners include strains that avoid dead-end mutations

Lindsey N. Wolf and Jeffrey E. Barrick

Without watching evolution in action, it is difficult to imagine how stepwise mutations from a common ancestor can add up to the differences between Escherichia coli and elephants. Field studies, such as those of Darwin’s finches in the Galapagos, show that evolution can be observed as it happens in natural populations. But general principles can prove difficult to extrapolate from such studies, in part because evolutionary outcomes have an element of chance and can take a long time to play out.

To overcome these limitations, we and others in the field of experimental evolution study microbial populations in the lab, taking advantage of their capacity to replicate rapidly. With bacteria, for example, one can watch hundreds or thousands of generations pass in a matter of weeks or months, making it possible to observe significant evolutionary change as it happens.

Furthermore, setting up many separate populations of microbes in precisely controlled environments is easy. Each culture tube is independent, meaning differences that evolve among initially identical populations can be used to tease out how much chance contributes to an outcome. Unlike Galapagos finches, intermediate-stage samples of microbes can be preserved during such experiments and later revived for further study.

Finally, tools for genetically manipulating bacteria such as E. coli make it possible to dissect evolutionary trajectories mutation by mutation. By combining this capability with the systems biology toolbox (genomics, transcriptomics, proteomics, and metabolomics) and decades of accumulated knowledge about microbial physiology, one can begin to understand precisely the molecular details of how mutations lead to specific differences in replication and survival that determine evolutionary success.

Evolution in an Erlenmeyer

In 1988, Richard Lenski, then at the University of California (UC), Irvine, started a deceptively simple experiment. From E. coli colonies on a petri dish, he placed 12 populations into flasks filled with glucose and other nutrients, then allowed the cells to multiply until they exhausted that sugar. The next day, he transferred 1% of each culture to a new set of 12 flasks filled

Summary

- During more than 50,000 generations, lineages of Escherichia coli cells cultured in a controlled environment explored different mutational paths to higher fitness.
- Mutations, particularly those in regulatory genes, can affect whether subsequent mutations in other genes do or do not confer fitness benefits.
- Despite not having the most advantageous initial mutations, an E. coli strain with greater potential to evolve prevailed amid other similar strains within this population of bacteria.
- Second-order selection for evolvability may enable populations of disease-causing microbes to avoid long-term evolutionary dead ends.
with the same glucose-containing solution and allowed the E. coli to grow again until they exhausted that nutrient. Lenski and his colleagues at UC Irvine and later Michigan State University have continued to feed and periodically freeze samples of these 12 populations of E. coli for more than 20 years, allowing more than 50,000 generations of bacterial history to accumulate.

What has happened in these flasks? Each one began with a uniform population of E. coli, but genomic copying errors and DNA damage led to mutations in some offspring. These changes might be lethal, be deleterious but non-fatal, have no effect on cell function, or improve a mutant’s ability to compete for glucose. Many millions of bacteria and many different branches of offspring simultaneously “try out” different mutational paths in one flask (Fig. 1). Subpopulations and their descendants wax and wane in frequency as each randomly mutates, until one gets far enough ahead in the race for glucose to drive the others extinct, and then diversification begins anew. Compared to ancestral E. coli cells, those that were cultured for 50,000 generations replicate twice as fast, producing roughly 100 descendant cells for every one cell of the ancestor when they compete in the same flask for one day of growth!

How many mutations accumulated over 50,000 generations? The rate of change is surprisingly slow: individuals picked from one population differ from their progenitors by around 20 mutations for every 10,000 generations of evolution (or roughly one mutation every two and a half months). About two-thirds of these mutations are base substitutions that change only a single letter in the genome, which for these bacteria consists of 4.6 million DNA bases. The remaining one-third of these mutations includes mobile DNA elements inserting copies into new places in the genome, other insertions and deletions ranging from a few to several thousand DNA bases, and other chromosomal rearrangements.

Most of the mutations that accrue appear to be beneficial. The same genes tend to be “hit” by mutations in more than one population, suggesting a strong selective pressure to alter this particular set of genes. Moreover, many of these mutations improve the ability of E. coli to compete for glucose. Single mutations have been moved into the ancestral strain and shown to increase competitive fitness or taken out of evolved genomes and found to decrease fitness. Although we still do not fully understand how every change affects cell physiology, these mutations can be categorized as mainly affecting either metabolic or regulatory processes.

For example, the gene pykF is a metabolic target whose activity is decreased by mutations in all 12 populations of the long-term experiment. PykF is one of two E. coli versions of the enzyme phosphofructokinase, which is responsible for converting phosphoenolpyruvate (PEP) to pyruvate during glycolysis. Although disrupting an enzyme involved in glucose breakdown might seem disadvantageous, slowing this step might lead to PEP buildup and drive the import of more glucose via the phosphotransferase system. Thus, an increase in nutrient uptake rate may account for the large fitness benefit from mutating pykF.

While metabolic genes code for proteins involved in specific pathways, regulatory genes typically activate or repress many genes at once. Thus, a single mutation in a regulatory gene might alter interactions among many genes, exerting larger effects on fitness than a mutation within a single metabolic gene. Indeed, regulatory mutations that affect chromosomal topology and responses to nutrient starvation are among the earliest and most beneficial mutations that occur in these long-term evolution experiments involving E. coli. Part of their benefit apparently is due to how they repress certain
costly genes that are not needed in this environment.

**Not All Paths to Higher Fitness Are Equivalent**

As a graduate student with Lenski at Michigan State University, Robert Woods traced the fates of several beneficial mutations in one of the 12 long-term populations (Fig. 2). Not only did he reconstruct the timing and order of these mutations, he also learned that several other subpopulations with different muta-

tions vied with these long-term “winners” for dominance. One such “loser” subpopulation appeared by 500 generations and persisted until at least 1,500 generations. Knowing the final outcome might lead one to expect that the winners had a competitive advantage over these losers at every stage. However, when 500-generation isolates of each kind were revived and allowed to compete head-to-head, the long-term winners failed at this stage by a large margin. Had these losers and winners continued growing together in the original flask, the results would have been reversed: the winners would have been driven extinct within only 350 generations.

What explains this discrepancy? The winners must have acquired new mutations with large fitness benefits that allowed them to leapfrog competing subpopulations. Did they win because they were unusually lucky in just this one instance of evolution in the original flask? Or, were they somehow better at evolving, such that this outcome would be common if we could watch them evolve over and over again? Woods was able to test these possibilities by going back to the early frozen samples to replay what happened. He started 20 experiments from winner types and 20 experiments from loser types, and then transferred these populations daily under the same conditions as the first long-term experiment.

After 883 generations, he found a result consistent with the winner type being reliably “better” at evolving rather than simply “lucky.” On average, descendants of the 500-generation winners overcame their fitness deficits and surpassed the descendants of the loser 500-generation isolates during the replay experiments (Fig. 2). Both types gave rise to descendants with improved fitness, but apparently the mutational “road” that the winners had already started down led to better opportunities. In contrast, the loser types had mutations that impeded pathways to higher fitness.
Some Mutations Lead to Regulatory Blind Alleys

We sequenced the genomes of the winner and loser types to determine what mutations accumulated after 500 generations. Tim Cooper and his colleagues at the University of Houston reconstructed strains carrying combinations of the mutations that we found by sequencing. Subsequent analysis showed that key regulatory mutations occurred in the genes topA and spoT. Tweaking TopA, a DNA topoisomerase, alters the level of DNA twisting in the E. coli chromosome and, consequently, the expression of many genes. The winner mutant strains had a single amino acid substitution in topA, while the losers had a different substitution only one amino acid away. The overall effect was beneficial to fitness in the glucose environment in both cases.

SpoT, another global regulator, detects and responds across many genes to starvation conditions. A highly beneficial mutation in the spoT gene occurred after the winner topA mutation in the original long-term experiment. This mutation was also beneficial when added to the same genome as the winner topA mutation. However, it was not beneficial in combination with the loser topA mutation. Consistent with these results, we found that new mutations in spoT evolved in several of the winner replay experiments, but we did not find spoT mutations in any of the loser replay experiments. In contrast, we found that mutations in the metabolic gene pykF almost always evolved during replay experiments of both winner and loser types of bacteria, indicating that they were still among the most beneficial mutation steps regardless of which earlier mutational path was being followed.

We do not know why these two mutations in topA interact so differently with other mutations, including those in spoT. Probably, mutations in these global regulators adjust the ex-
pression and activities of many other genes. The E. coli strain used to start these experiments was not optimally adapted to its glucose-limited environment, meaning that changes in one of these regulators could probably adjust multiple genes more toward optimal levels.

Theoretically, a beneficial mutation adjusting one regulator to its optimal level can be a false step that leads to a dead end. The best overall solution may rely on mutating one regulator in a direction that is beneficial, but only so far that it leaves the door open for a subsequent beneficial mutation in a second regulator to reach an even better fitness (Fig. 3). We are now measuring gene expression in these strains to test whether this hypothesis explains our observations of winner and loser evolution. However, unlike this idealized example, we do not know what the optimal levels of expression are for genes in this environment or how changing the activity of any given target gene influences fitness, so it may be challenging to understand the results in terms of this simple model.

**Outlook**

The E. coli subpopulation that acquired mutations that initially rendered it more fit did not prove the most fit much later after further evolution. Instead, a competing lineage that maintained more potential for further adaptation eventually prevailed. This second-order selection for greater evolvability can arise in other contexts provided several basic conditions are met. First, competing strains must have time to explore multiple mutation paths. The larger a population and the higher the mutation rate among its members, the more likely significantly diverged lineages will coexist before one gains an advantage and drives others extinct. Second, the fitness effects of beneficial mutations must strongly depend on other mutations within a genome. Such fitness interactions may often involve regulatory genes.

While these experiments follow bacteria in a controlled environment, they help to illustrate what to expect outside the test tube. Other investigators, charting genomic changes within bacterial pathogens that chronically infect patients, report finding patterns that resemble what we observe in culture. This finding from a clinical setting leads us to wonder whether second-order selection for evolvability is important in populations of bacterial pathogens. Typically, we see only the microbial survivors because they did not blunder into long-term blind alleys. By understanding more about the systems biology underlying microbial evolution, we may better recognize what types of mutations lead to dead ends, and this understanding in turn might enable us to trick pathogens into evolutionary cul-de-sacs.

**SUGGESTED READING**


