A Test of the Repeatability of Measurements of Relative Fitness in the Long-Term Evolution Experiment with *Escherichia coli*

Jeffrey E. Barrick, Daniel E. Deatherage, and Richard E. Lenski

Abstract Experimental studies of evolution using microbes have a long tradition, and these studies have increased greatly in number and scope in recent decades. Most such experiments have been short in duration, typically running for weeks or months. A venerable exception, the long-term evolution experiment (LTEE) with *Escherichia coli* has continued for 30 years and 70,000 bacterial generations. The LTEE has become one of the cornerstones of the field of experimental evolution, in general, and the BEACON Center for the Study of Evolution in Action, in particular. Science laboratories and experiments usually have finite lifespans, but we hope that the LTEE can continue far into the future. There are practical issues associated with maintaining such a long-term experiment. One issue, which we address here, is whether key measurements made at one time and place are reproducible, within reasonable limits, at other times and places. This issue comes to the forefront when one considers moving an experiment like the LTEE from one lab to another. To that end, the Barrick lab at the University of Texas at Austin, measured the fitness values of samples from the 12 LTEE populations at 2,000, 10,000, and 50,000 generations and compared the new data to data previously obtained at Michigan State University. On balance, the datasets agree very well. More generally, this finding shows the value of simplicity in experimental design, such as using a chemically defined growth medium and appropriately storing samples from microbiological experiments. Even so, one must be vigilant in checking assumptions and procedures given the potential for uncontrolled factors (e.g., water quality) to affect outcomes. This vigilance is perhaps especially important for a trait like fitness, which integrates all aspects of organismal performance and may therefore be sensitive to any number of subtle environmental influences.

Key words Bacteria, experimental evolution, fitness, long-term studies, reproducible research

---

Jeffrey E. Barrick
Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX USA; and BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI USA; email: jbarrick@cm.utexas.edu

Daniel E. Deatherage
Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX USA; and BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI USA

Richard E. Lenski
Department of Microbiology and Molecular Genetics, and BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI USA; email: lenski@msu.edu

1 Introduction

Microorganisms have been used to study evolution in action for well over a century, dating back to work by William Dallinger, who corresponded with Charles Darwin about his research in the 1880s [18, 24], and continuing with pioneering modern experiments by Novick and Szilard [33] and Atwood *et al.* [1]. The field of microbial experimental evolution has expanded greatly in recent years with many laboratories using diverse viruses, bacteria, and fungi to address a wide range of questions [14, 25, 47]. The ability to sequence and compare the complete genomes of ancestral and experimentally derived samples has led to an even faster expansion of this field in the last decade [2, 7, 10, 47].

The speed with which many microbes can reproduce is one of the main reasons they have become experimental models for studying evolution. It is often possible to observe evolutionary changes within days, weeks, or months, depending on the organisms and environments used in an experiment. Even so, there are advantages to running such experiments for much longer periods. Some evolutionary phenomena, like speciation, may take thousands or even millions of generations to play out. Also, resolving subtly different models of evolutionary dynamics and how adaptive fitness landscape are structured may require very long time-series of data [49, 50].

In 1988, Richard Lenski began the long-term evolution experiment—the LTEE—with 12 replicate populations of *Escherichia coli*, all started from the same ancestor, except for a genetic marker embedded in the experimental design. The LTEE was intended to address three overarching questions [15, 29]. First, for how long can the bacteria continue to improve before they reach some limit to their fitness? Second, are adaptive changes repeatable across the replicate populations, or do the populations adapt in different ways to the same environment? Third, how tightly coupled are the dynamics of phenotypic adaptation and genetic change?

Lenski called this experiment "long-term" from the outset, including in the title of the first paper on the LTEE [15, 29]. He did so because he previously performed similar experiments that lasted for several hundred generations, and he realized from the fitness trajectories and resulting gains that the evolving populations in those experiments had probably experienced only one or two selective sweeps [6, 23]. So few fixations seemed unsatisfactory for addressing the questions that motivated the LTEE.

The LTEE started at the University of California, Irvine, in February of 1988. After Lenski accepted a faculty position at Michigan State University, the 12 populations were frozen in April of 1992, after reaching 10,000 generations. The LTEE was then restarted from the frozen samples at Michigan State University in October of 1992, where it has continued ever since. Competition experiments were performed in the experiment's old and new homes to confirm that the improvements in fitness seen in the original laboratory would be reproducible in its new environment. That experiment used samples from only a single time point, namely 10,000 generations, and it was not formally analyzed or published. However, the correspondence in fitness values was judged as being satisfactory.

Over its long history, samples from the LTEE populations have been shared with dozens of laboratories around the world. In general, observations made in the original lab have been highly

reproducible in other labs. For example, extensive genome sequencing by other labs shows that the populations have been successfully maintained without cross-contamination [17, 45]. Major phenotypic changes, including the evolution of hypermutability in some populations [3, 17, 43, 45] and the surprising appearance of citrate utilization (Cit^+ phenotype) in one population [4, 5, 31, 38, 39] have also been confirmed in other labs. Even more subtle phenomena, such as crossfeeding between two lineages that evolved and coexisted for tens of thousands of generations in one population have been confirmed in studies by other labs [21, 37, 41, 42].

Nonetheless, there have been occasional unexpected results in the LTEE. Notably, it has not been possible, even in the lab at Michigan State University, to repeat or explain the extinction of the Cit^- lineage in the LTEE population that evolved the Cit^+ phenotype [46]. When frozen populations from before this extinction occurred in the LTEE were revived and evolution was "replayed" many times from them, this extinction did not re-occur. The best guess in this case is that some unintended, transient perturbation in the conditions of the experiment caused the Cit^- ecotype to go extinct. It may have been especially susceptible to these types of fluctuations because the Cit^- ecotype had a small population size relative to the dominant Cit^+ ecotype.

Measurements of how *E. coli* fitness evolves on long timescales are arguably some of the most important and unique data from the LTEE. These fitness trajectories and the related mutational dynamics have been used to examine different models of evolution [3, 11, 12, 16, 20, 36, 49, 50]. Yet, it is rare in microbial evolution experiments to repeat fitness measurements in different labs or at intervals separated by decades in the same lab to examine the repeatability of these values and their evolutionary trajectories. Testing repeatability is especially important for a trait like fitness, which integrates across all aspects of organismal performance and might therefore be sensitive to any number of environmental influences, such as water quality or trace impurities in chemical components [13, 34], in an otherwise defined culture medium and environment. In light of the fact that all research groups have finite lives and the LTEE will move, sooner or later to another laboratory and likely another location, it seems important to examine the repeatability of these measurements. To that end, we measured the fitness values of samples from the 12 LTEE populations at 2,000, 10,000, and 50,000 generations at the University of Texas at Austin and compared these new data to data previously obtained at Michigan State University [50, 51]. As we show below, the two datasets agree quite well, given the inherent measurement noise associated with these assays.

2 Methods

2.1 Long-term evolution experiment

The LTEE consists of 12 populations of *E. coli*, all founded from essentially the same ancestor and propagated under identical conditions [29]. Six populations began with strain REL606 [19], and the other six started with REL607. These two strains differ by a mutation that allows the former, but not the latter, to grow on arabinose [29, 44]. This marker serves to distinguish competitors during assays of relative fitness (**see section 2.2**). The marker also helps guard against undetected cross-contamination of the populations. The six REL606-derived Ara^- and six REL607-derived Ara^+ populations are strictly alternated during the serial-transfer procedure. As a consequence, cross-contamination during successive transfers would introduce cells with the

wrong arabinose-marker state, which could be detected during quality-control checks performed when samples are frozen.

Every day, 1% of each population is removed and transferred to fresh medium. The 100-fold dilution and subsequent re-growth allow $\log_2 100$ generations (i.e., doublings), which is rounded to six and two-thirds generations per daily transfer. Every 75 days (500 generations), after the transfers are performed and the cells are plated on various test media, glycerol (a cryoprotectant) is added to each culture from the previous day. These whole-population samples are then split and stored in duplicate vials at –80°C, providing a "frozen fossil record" of cells that remain viable and can be revived for later analyses. After incubation, the test plates are used to inspect the bacteria for growth and colony appearance, with particular attention given to whether there is any evidence of cells with the wrong arabinose-marker state. In recent years, genome analysis methods discovered unique mutations that have arisen in each of the LTEE populations [35, 45, 52], and those loci can now also be checked for suspicious colonies. In the event of suspected or confirmed contamination, the population in question is re-started from an earlier frozen sample; as a result, some LTEE populations are 500 generations or more behind the leaders, which recently passed 70,000 generations. Various other disruptions have occurred over the years, including moving the experiment from California to Michigan, so that the LTEE as a whole runs several thousand generations behind the maximum number that could have been achieved after some 30 years.

The LTEE populations live in 50-mL glass Erlenmeyer flasks containing 10 mL of culture medium, with small glass beakers placed over the flask openings. The flasks are incubated at 37°C with orbital shaking at 120 r.p.m. The culture medium is Davis Minimal medium [9] supplemented with thiamine at 2 μg per mL and glucose at 25 μg per mL (DM25), where glucose is the limiting resource. (Note: The first paper on the LTEE [29] misstated the concentration of thiamine, but cited an earlier paper [22] with the correct concentration; the recipe itself has never varied.) The ancestral strains reach a stationary-phase density of $\sim 5 \times 10^7$ cells per mL in DM25 [29]. The evolved bacteria tend to produce larger cells, and they reach somewhat lower numerical densities at stationary phase [27, 28, 48], with one conspicuous exception. That exception is the population, called Ara–3, that evolved the capacity to use the citrate in the medium as an additional carbon source, which allows it to reach a substantially higher cell density than the ancestors or other evolved populations [5].

2.2 Assays of relative fitness

To assess relative fitness, we conducted head-to-head competition assays in the same medium and other conditions as used for the LTEE. Each pairwise assay competed a whole-population sample taken from one of the 12 populations at 2,000, 10,000, or 50,000 generations against the reciprocally marked ancestral strain, except as noted below. Thus, the Ara^+ ancestor REL607 competed against the six populations founded by the Ara^- REL606 ancestor, while REL606 competed against the six populations founded by REL607. The arabinose-utilization marker is neutral in the LTEE environment, and the ancestral strains have indistinguishable relative fitness [29]. When grown on tetrazolium arabinose (TA) indicator agar plates, Ara^- and Ara^+ cells form red and white colonies, respectively, thereby allowing the number of cells of each type in a mixture to be estimated [29]. Over time, however, some LTEE populations evolved such that

they no longer produce colonies on TA agar, or they produce many colonies that are difficult to accurately score as Ara$^-$ versus Ara$^+$. For this reason, Wiser *et al.* [50] excluded the 10,000-generation sample for population Ara+6 and the 50,000-generation samples for populations Ara+6 and Ara–2. Additionally, this previous study omitted the Ara–3 population at 50,000 generations because the utilization of a new nutrient pool by Cit$^+$ cells that is inaccessible to the Cit$^-$ ancestor complicates both the measurement and interpretation of relative fitness values. In this study, we excluded those same four samples as well as the 50,000-generation samples for populations Ara+1 and Ara+4 because we could not reliably estimate the evolved cell numbers from their colony counts. Thus, our dataset has relative fitness estimates for a total of 30 evolved population samples, with two-fold replication of each estimate.

Prior to the competitions, 0.12 mL aliquots of frozen stocks of each whole-population sample and each of the two ancestral strains were transferred to 50-ml Erlenmeyer flasks containing 10 mL Luria Broth and grown overnight. Each revived culture was then diluted 100-fold in 10 mL of sterile saline solution, before using 0.1 mL to inoculate 9.9-mL cultures of DM25 in 50-ml Erlenmeyer flasks. Two independent cultures were inoculated for each evolved strain, 28 were inoculated for the Ara$^-$ REL606 ancestor, and 32 were inoculated for the Ara$^+$ REL607 ancestor. These cultures were grown for 24 h to acclimate them to the LTEE conditions. Each competition assay was then initiated by combining 50 μL of an evolved population with 50 μL of the ancestral strain with the opposite arabinose-marker state in 9.9 mL of DM25 in a 50-ml Erlenmeyer flask. Each acclimated culture was used to inoculate only a single competition culture, providing technical independence. After vortexing the competition culture, 0.1 mL was removed, diluted into 9.9 mL of saline solution, and 40 μL was spread on a TA plate. After 24 h, 0.1 mL of culture was serially diluted twice in 9.9 mL of saline (10,000-fold total dilution), and 40 μL was spread on TA plates for the competitions involving 2,000- and 10,000-generation populations; 80 μL was used for the assays involving the 50,000-generation populations due to their lower cell density. For both the initial and final TA plates, red and white colonies were counted after incubation for 16 to 24 h at 37°C to estimate the abundances of the ancestral and evolved competitors.

We calculated a relative fitness value from each competition assay as the ratio of the realized growth rates for each competitor in that assay [29], as follows:

$w = \ln(E_f / E_i) / \ln(A_f / A_i)$,

where E and A are the evolved population and ancestor, respectively, and the f and i subscripts indicate final and initial densities, respectively, as estimated from the plate counts. These fitness values integrate any and all differences in growth of the two competitors across all physiological states during the serial-transfer cycle as experienced in the LTEE itself [29, 48].

2.3 Statistical analysis

The relative fitness values and power-law model calculations from the paper by Wiser *et al.* [50] were previously deposited in the Dryad Digital Repository [51]. We deposited the new fitness data collected for this study in the Dryad Digital Repository (accession number pending) and, for convenience, we copied to this new accession the relevant information from that earlier paper.

We obtained two estimates of the relative fitness for each of 30 evolved population samples (**see section 2.2**). We chose to perform two replicates for each sample because that was the same level of technical replication as in the earlier dataset [50, 51]. However, our interest in this study is not in the statistical noise among technical replicates, but rather in the correspondence between the estimates of relative fitness obtained in the two different labs. Therefore, all statistical analyses were performed on the geometric means of the relative fitness values obtained from each set of two technical replicates. Calculations were performed in R version 3.5.0 [40].

3 Results

In this study we measured the fitness of each of the 12 LTEE populations at 2,000, 10,000, and 50,000 generations relative to the ancestral *E. coli* strains at the University of Texas at Austin (UTA). A total of 30 fitness values were obtained after some evolved population samples were omitted for technical reasons (**see section 2.2**). We compared these fitness values to a matched set of 30 fitness values obtained previously at Michigan State University (MSU) [50, 51] (**Fig. 1**). Overall, there was a strong correlation between the UTA and MSU values (Pearson's r = 0.830, $p < 0.001$). The fact that this correlation is imperfect could reflect measurement error alone, or it might reflect some consistent difference in the experimental conditions between the two locales. Therefore, we next considered whether we could detect any systematic bias between the two datasets.

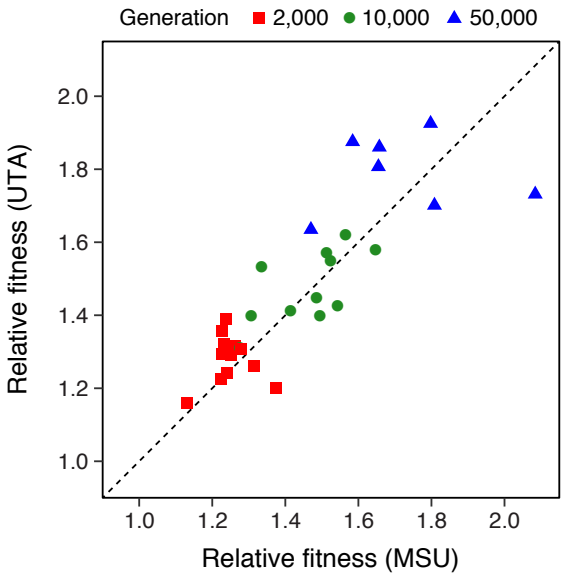

Figure 1. Comparison of fitness values for LTEE population samples from 2,000, 10,000, and 50,000 generations, measured relative to the ancestral *E. coli* strain at the University of Texas at Austin (UTA, this study) and previously at Michigan State University (MSU, [51]).

First, there might be some difference between the laboratories that affected all measurements of fitness in the same direction. In this case, one set of values would consistently overestimate fitness relative to the other. We found no evidence for this type of bias. The UTA fitness values were slightly higher, on average, than the corresponding MSU values, but this difference was not

significant ($p = 0.169$, two-tailed paired *t*-test). The 95% confidence interval for this difference ranged from –0.015 to 0.080.

Second, there might be a bias that somehow depends on the generational time point tested. There has been widespread parallel evolution in the LTEE [26, 45, 52], and it is possible that the populations accumulated mutations over time in a way that led to generation-specific changes in the sensitivity of fitness measurements to the two different lab environments. However, we saw no systematic difference in fitness between the UTA and MSU measurements at any of the generations assayed ($p = 0.246$, $p = 0.645$, and $p = 0.446$ for the samples from 2,000, 10,000, and 50,000 generations, respectively, by two-tailed paired *t*-tests).

We then considered an alternative way of evaluating the UTA data. The true fitness values of the LTEE populations are not known, of course, but models of fitness trajectories that integrate information from fitness measurements across many generations should provide a more accurate estimate of the actual fitness at any generation than the values used above. In particular, a power-law model (PLM) has been shown to describe and even predict the fitness trajectories for each LTEE population over time much better than an alternative hyperbolic model. The PLM was first evaluated by Wiser *et al.* [50] using a dataset of fitness measurements for each population at 41 time points through 50,000 generations of the LTEE. This model was subsequently extended and supported through 60,000 generations [30].

If we assume the fitness value predicted by the PLM for each population at each time point represents the true value (or at least is closer to the true value than the measurements from that generation alone), then we would expect there to be a better correlation between the UTA data and the PLM predictions than there is between the UTA and MSU data. Indeed, this is the case. The correlation between the UTA data and PLM predictions ($r = 0.882$) is somewhat stronger than the correlation between the UTA and MSU data ($r = 0.830$). However, the improvement in the strength of this correlation is not significant ($p = 0.125$, one-tailed paired *t*-test comparing the squared residuals).

Under the assumption that the PLM predicts the true fitness values with reasonable precision, we can also perform a regression to test whether we can discriminate between the measured data and the predicted values. A linear regression of the MSU data against the PLM predictions that is forced to pass through a relative fitness of one on both axes gives a slope that is not significantly different from one (**Fig. 2A**, $p = 0.769$). Regressing the UTA data against the PLM predictions in this way also gives a slope that is not significantly different from one (**Fig. 2B**, $p = 0.113$).

Taken together, these analyses indicate that the estimates of relative fitness for the evolved population samples from the LTEE are reproducible, within statistical limits, when measured in different laboratories, by different people, and at different times.

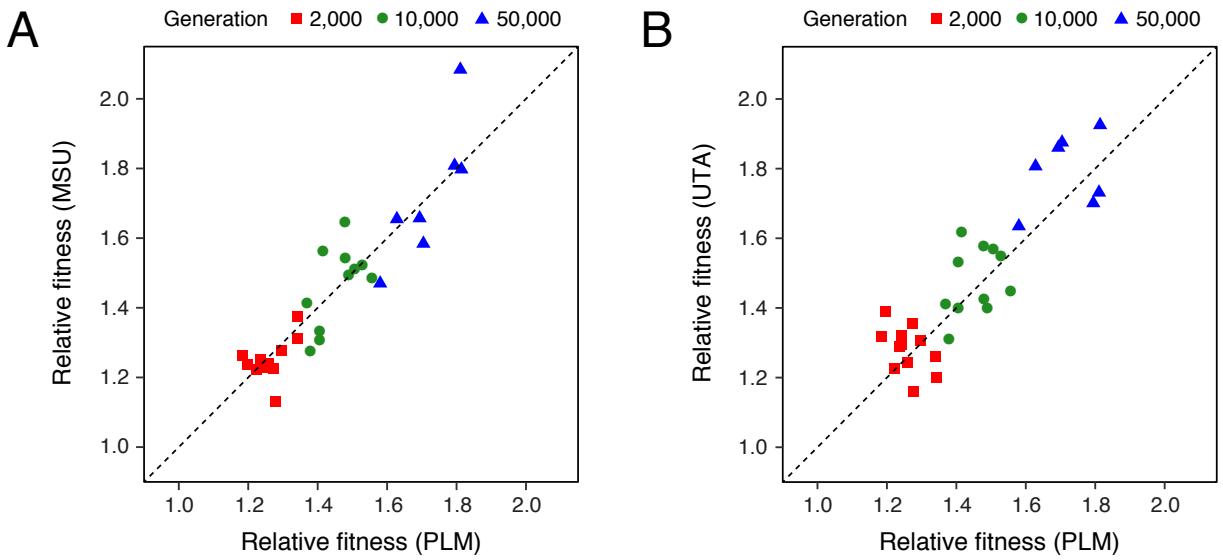

Figure 2. Comparisons of (**A**) MSU and (**B**) UTA relative fitness measurements to predictions of a power-law model for the fitness trajectories in the LTEE (PLM). PLM parameters fit separately to each population were used to predict fitness values at 2,000, 10,000, and 50,000 generations.

4 Discussion

The LTEE has been running for more than 30 years and 70,000 generations, and we hope it will continue far into the future. More than 80 papers have been published using the bacteria and data derived from this one experiment; several other evolution experiments have been spun off from it; and countless other experiments have been influenced by it. Arguably, no other evolution experiment has been studied in as many ways and with as much quantitative rigor as the LTEE. Much of the value of any model system, including the LTEE, rests on its reproducibility. The LTEE was designed to be simple in terms of the culture medium and other environmental conditions used, and such simplicity undoubtedly helps promote reproducibility. The fact that the bacteria can be stored frozen and later revived is also critical, because it allows samples from one time and location to be analyzed in new ways or re-analyzed at different times and places.

In evolutionary biology, the relative fitness of different genotypes, including the ancestral and derived bacteria in the LTEE, is a quantity of central interest and importance. It effectively integrates everything about the organisms' genomes and their phenotypes (at least those that are relevant for performance in a given environment) into a single measure of reproductive success. However, that integrative aspect also raises the possibility that measurements of relative fitness might be especially sensitive to subtle variations in the test conditions. The simplicity of the conditions used in the LTEE means that many possible sources of variation are well controlled. Nonetheless, there is the potential for small fluctuations and unintended perturbations to have an outsized impact, especially on an integrative and quantitative trait like fitness.

To address this issue, we performed competition assays at the University of Texas at Austin to measure the fitness of population samples from generations 2,000, 10,000, and 50,000 of the LTEE relative to their common ancestor. We compared these new estimates to values obtained for the same populations several years earlier at Michigan State University using the same

methods. In short, the new and old data agreed well, with deviations fully consistent with ordinary sampling error. It is highly encouraging for the future of the LTEE that such a potentially sensitive metric as relative fitness can be reliably estimated even after population samples have been stored frozen for years, 'copied' by re-culturing the samples, and then shipped to and analyzed in another laboratory. This consistency is important because it means that the experiment can be passed to future generations of researchers, who can continue to monitor and study the evolution of the populations to analyze both the long-term trends that are in common to all of them and any surprises that may be in store for particular populations.

It is also interesting that, while the methods and resulting estimates are reproducible, the bacteria themselves have sometimes evolved in ways that make obtaining those estimates more challenging. In particular, several populations have evolved such that cells no longer produce colonies that can be reliably counted on the indicator agar used to distinguish competitors in the assays of relative fitness. This colony feature is not a phenotype that was selected or contributes to fitness in the LTEE; on the contrary, this trait can decay precisely because it has no bearing on fitness in that environment. Other challenges that have evolved include crossfeeding, in which some cells secrete metabolites that others can use as resources, and the ability of one lineage to use citrate. As a consequence of these changes, the fitness of some evolved lineages relative to their ancestor or to one another depends (albeit only subtly in most cases) on the frequency of the competitors at the start of a competition assay. It is now becoming possible to decouple making fitness measurements from the requirement that cells from the LTEE form colonies on indicator agar by using DNA sequencing to read out the frequencies of different alleles [8] and to multiplex fitness measurements using DNA barcoding techniques [32]. These and other new technologies will continue to enhance our understanding of evolutionary dynamics in the LTEE. The phenotypic changes in the LTEE may complicate some experimental procedures, but they also enrich the LTEE by revealing the interesting ways that the bacteria can evolve and adapt, and thus the experiment's lasting potential to generate new insights and discoveries.

5 Acknowledgments

We thank Erik Goodman for his inspirational and effective leadership of the BEACON Center, Mike Wiser for sharing data, and Caroline Turner for catching the error in the published recipe for the DM25 medium. This work has been supported by an NSF grant (DEB-1451740), the BEACON Center for the Study of Evolution in Action (NSF Cooperative Agreement DBI-0939454), and the John Hannah endowment at Michigan State University. Any opinions, findings, and conclusions or recommendations in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Atwood, K. C., Schneider, L. K., Ryan, F. J.: Periodic selection in *Escherichia coli*. Proc Natl Acad Sci USA 37, 146–155 (1951)
2. Barrick, J. E., Lenski, R. E.: Genome dynamics during experimental evolution. Nature Rev Gen 14, 827–839 (2013)
3. Barrick, J. E., Yu, D. S., Yoon, S. H, Jeong, H., Oh, T. K., Schneider, D., Lenski, R. E., Kim, J. F.: Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature 461, 1243–1247 (2009)
4. Blount, Z. D., Barrick, J. E., Davidson, C. J., Lenski, R. E.: Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature 489, 513–518 (2012)
5. Blount, Z. D., Borland, C. Z., Lenski, R. E.: Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. Proc Natl Acad Sci USA 105, 7899–7906 (2008)
6. Bouma, J. E., Lenski, R. E.: Evolution of a bacteria/plasmid association. Nature 335, 351–352 (1988)
7. Bruger, E. L., Marx, C. J.: A decade of genome sequencing has revolutionized studies of experimental evolution. *Curr. Opin. Microbiol.* **45**, 149–155 (2018)
8. Chubiz, L. M., Lee, M.-C., Delaney, N. F., Marx, C. J.: FREQ-Seq: a rapid, cost-effective, sequencing-based method to determine allele frequencies directly from mixed populations. PLoS ONE 7, e47959 (2012)
9. Carlton, B. C., Brown, B. J.: Gene mutation. Pp. 222-242 *in* Gerhardt, P. (ed.), Manual of Methods for General Bacteriology. American Society for Microbiology, Washington (1981)
10. Conrad, T. M., Lewis, N. E., Palsson, B. O.: Microbial laboratory evolution in the era of genome-scale science. Mol. Syst. Biol. 7, 509 (2011)
11. De Visser, J. A. G. M., Krug, J.: Empirical fitness landscapes and the predictability of evolution. Nature Rev Gen 15, 480–490 (2014)
12. De Visser, J. A. G. M., Elena, S. F., Fragat, I., Matuszewski, S.: The utility of fitness landscapes and big data for predicting evolution. Heredity, preprint https://doi.org/10.1038/s41437-018-0128-4 (2018)
13. Egli, T.: Microbial growth and physiology: A call for better craftsmanship. Front Microbiol 6, 287 (2015)
14. Elena, S. F., Lenski, R. E.: Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat Rev Genet 4, 457–469 (2003)
15. Fox, J. W., Lenski, R. E.: From here to eternity–the theory and practice of a really long experiment. PLoS Biol 13, e1002185 (2015)
16. Good, B. H., Desai, M. M.: The impact of macroscopic epistasis on long-term evolutionary dynamics. Genetics 199, 177–190 (2015)
17. Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., Desai, M. M.: The dynamics of molecular evolution over 60,000 generations. Nature 551, 45–50 (2017)
18. Haas, J. W., Jr.: The Reverend Dr William Henry Dallinger, F.R.S. (1839–1909). Notes Rec R Soc Lond 54, 53–65 (2001)
19. Jeong, H., Barbe, V., Lee, C. H., Vallenet, D., Yu, D. S., Choi, S. H., Couloux, A., Lee, S. W., Yoon, S. H., Cattolico, L., Hur, C. G., Park, H. S., Ségurens, B., Kim, S. C., Oh, T. K., Lenski, R. E., Studier, F. W., Daegelen, P., Kim, J. F.: Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). J Mol Biol 394, 644–652 (2009)
20. Kryazhimskiy, S., Tkacik, G., Plotkin, J. B.: The dynamics of adaptation on correlated fitness landscapes. Proc Natl Acad Sci USA 106, 18638–18643 (2009)

21. Le Gac, M., Plucain, J., Hindré, T., Lenski, R. E., Schneider, D.: Ecological and evolutionary dynamics of coexisting lineages during a long-term experiment with *Escherichia coli*. Proc Natl Acad Sci USA 109, 9487–9492 (2012)
22. Lenski, R. E.: Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. Evolution 42, 425–432 (1988)
23. Lenski, R. E.: Experimental studies of pleiotropy and epistasis in *Escherichia coli*. II. Compensation for maladaptive pleiotropic effects associated with resistance to virus T4. Evolution 42, 433–440 (1988)
24. Lenski, R. E.: What is adaptation by natural selection? Perspectives of an experimental microbiologist. PLoS Genet 13, e1006668 (2017)
25. Lenski, R. E.: Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J 11, 2181–2194 (2017)
26. Lenski, R. E.: Convergence and divergence in a long-term experiment with bacteria. Am Nat 190, S57–S68 (2017)
27. Lenski, R. E., Mongold, J. A.: Cell size, shape, and fitness in evolving populations of bacteria. Pp. 221-235 *in* Brown, J. H., West, G. B. (eds.), Scaling in Biology. Oxford Univ Press, Oxford (2000)
28. Lenski, R. E., Travisano, M.: Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc Natl Acad Sci USA 91, 6808–6814 (1994)
29. Lenski, R. E., Rose, M. R., Simpson, S. C., Tadler, S. C.: Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. Am Nat 138, 1315–1341 (1991)
30. Lenski, R. E., Wiser, M. J., Ribeck, N., Blount, Z. D., Nahum, J. R., Morris, J. J., Zaman, L., Turner, C. B., Wade, B. D., Maddamsetti, R., Burmeister, A. R., Baird, E. J., Bundy, J., Grant, N. A., Card, K. J., Rowles, M., Weatherspoon, K., Papoulis, S. E., Sullivan, R., Clark, C., Mulka, J. S., Hajela, N.: Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with *Escherichia coli*. Proc R Soc B 282, 20152292 (2015)
31. Leon, D., D'Alton, S., Quandt, E. M., Barrick, J. E.: Innovation in an *E. coli* evolution experiment is contingent on maintaining adaptive potential until competition subsides. PLoS Genet 14, e1007348 (2018)
32. Levy, S. F., Blundell, J. R., Venkataram, S., Petrov, D. A., Fisher, D. S., Sherlock, G.: Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature 519, 181–186 (2015)
33. Novick, A., Szilard, L.: Experiments with the chemostat on spontaneous mutations of bacteria. Proc Natl Acad Sci USA 36, 708–719 (1950)
34. O'Keefe, K. J., Morales, N. M., Ernstberger, H., Benoit, G., Turner, P. E.: Laboratory-dependent bacterial ecology: A cautionary tale. Appl Environ Microbiol 72, 3032–3035 (2006)
35. Papadopoulos, D., Schneider, D., Meier-Eiss, J., Arber, W., Lenski, R. E., Blot, M.: Genomic evolution during a 10,000-generation experiment with bacteria. Proc Natl Acad Sci USA 96, 3807–3812 (1999)
36. Passagem-Santos, D., Zacarias, S., Perfeito, L.: Power law fitness landscapes and their ability to predict fitness. Heredity, preprint https://doi.org/10.1038/s41437-018-0143-5 (2018)

37. Plucain, J., Hindré, T., Le Gac, M., Tenaillon, O., Cruveiller, S., Médigue, C., Leiby, N., Harcombe, W. R., Marx, C. J., Lenski, R. E., Schneider, D.: Epistasis and allele specificity in the emergence of a stable polymorphism in *Escherichia coli*. Science 343, 1366–1369 (2014)
38. Quandt, E. M., Deatherage, D. E., Ellington, A. D., Georgiou, G., Barrick, J. E.: Recursive genomewide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in *Escherichia coli*. Proc Natl Acad Sci USA 111, 2217–22 (2014)
39. Quandt, E. M., Gollihar, J., Blount, Z. D., Ellington, A. D., Georgiou, G., Barrick, J. E.: Fine-tuning citrate synthase flux potentiates and refines metabolic innovation in the Lenski evolution experiment. eLife 4, 1–22 (2015)
40. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2018)
41. Rozen, D. E., Lenski, R. E.: Long-term experimental evolution in *Escherichia coli*. VIII. Dynamics of a balanced polymorphism. Am Nat 155, 24–35 (2000)
42. Rozen, D. E., Schneider, D., Lenski, R. E.: Long-term experimental evolution in *Escherichia coli*. XIII. Phylogenetic history of a balanced polymorphism. J Mol Evol 61, 171–180 (2005)
43. Sniegowski, P. D., Gerrish, P. J., Lenski, R. E.: Evolution of high mutation rates in experimental populations of *E. coli*. Nature 387, 703–705 (1997)
44. Studier, F. W., Daegelen, P., Lenski, R. E., Maslov, S., Kim, J. F.: Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3) and comparison of the *E. coli* B and K-12 genomes. J Mol Biol 394, 653–680 (2009)
45. Tenaillon, O., Barrick, J. E., Ribeck, N., Deatherage, D. E., Blanchard, J. L., Dasgupta, A., Wu, G. C., Wielgoss, S., Cruveiller, S., Médigue, C., Schneider, D., Lenski, R. E.: Tempo and mode of genome evolution in a 50,000-generation experiment. Nature 536, 165–170 (2016)
46. Turner, C. B., Blount, Z. D., Lenski, R. E.: Replaying evolution to test the cause of extinction of one ecotype in an experimentally evolved population. PLoS ONE 10, e0142050 (2015)
47. Van den Bergh, B., Swings, T., Fauvart, M., Michiels, J.: Experimental design, population dynamics, and diversity in microbial experimental evolution. Microbiol Mol Biol Rev 82, e00008-18 (2018)
48. Vasi, F., Travisano, M., Lenski, R. E.: Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. Am Nat 144, 432–456 (1994)
49. Wiser, M. J., Dolson, E. L., Vostinar, A., Lenski, R. E., Ofria, C.: The boundedness illusion: asymptotic projections from early evolution underestimate evolutionary potential. PeerJ Preprints 6, e27246v1 (2018)
50. Wiser, M. J., Ribeck, N., Lenski, R. E.: Long-term dynamics of adaptation in asexual populations. Science 342, 1364–1367 (2013)
51. Wiser, M. J., Ribeck, N., Lenski, R. E.: Data from: Long-term dynamics of adaptation in asexual populations. Dryad Digital Repository. https://doi.org/10.5061/dryad.0hc2m (2013)
52. Woods, R., Schneider, D., Winkworth, C. L., Riley, M. A., Lenski, R. E.: Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. Proc Natl Acad Sci USA 103, 9107–9112 (2006)