Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*

Zachary D. Blount, Christina Z. Borland, and Richard E. Lenski*

Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824

The role of historical contingency in evolution has been much debated, but rarely tested. Twelve initially identical populations of *Escherichia coli* were founded in 1988 to investigate this issue. They have since evolved in a glucose-limited medium that also contains citrate, which *E. coli* cannot use as a carbon source under oxic conditions. No population evolved the capacity to exploit citrate for >30,000 generations, although each population tested billions of mutations. A citrate-using (Cit$^+$) variant finally evolved in one population by 31,500 generations, causing an increase in population size and diversity. The long-delayed and unique evolution of this function might indicate the involvement of some extremely rare mutation. Alternately, it may involve an ordinary mutation, but one whose physical occurrence or phenotypic expression is contingent on prior mutations in that population. We tested these hypotheses in experiments that ''replayed'' evolution from different points in that population's history. We observed no Cit$^+$ mutants among 8.4×10^{12} ancestral cells, nor among 9×10^{12} cells from 60 clones sampled in the first 15,000 generations. However, we observed a significantly greater tendency for later clones to evolve Cit$^+$, indicating that some potentiating mutation arose by 20,000 generations. This potentiating change increased the mutation rate to Cit$^+$ but did not cause generalized hypermutability. Thus, the evolution of this phenotype was contingent on the particular history of that population. More generally, we suggest that historical contingency is especially important when it facilitates the evolution of key innovations that are not easily evolved by gradual, cumulative selection.

adaptation | experimental evolution | mutation | selection

At its core, evolution involves a profound tension between random and deterministic processes. Natural selection works systematically to adapt populations to their prevailing environments. However, selection requires heritable variation generated by random mutation, and even beneficial mutations may be lost by random drift. Moreover, random and deterministic processes become intertwined over time such that future alternatives may be contingent on the prior history of an evolving population. For example, multiple beneficial mutations will arise in some unpredictable order (1, 2), and those that are substituted first may differ from others in their pleiotropic effects and epistatic interactions (3, 4), thus constraining some evolutionary paths while potentiating other outcomes (5–9). These accidents of history may even determine the survival or extinction of entire lineages, given the capricious and sudden nature of some environmental changes (10–12).

Stephen Jay Gould maintained that these historical contingencies make evolution largely unpredictable. Although each change on an evolutionary path has some causal relation to the circumstances in which it arose, outcomes must eventually depend on the details of long chains of antecedent states, small changes in which may have enormous long-term repercussions (13–15). Thus, Gould argued that contingency renders evolution

fundamentally quirky and unpredictable, and he famously suggested that replaying the "tape of life" from some point in the distant past would yield a living world far different from the one we see today. Simon Conway Morris countered that natural selection constrains organisms to a relatively few highly adaptive options, so that "the evolutionary routes are many, but the destinations are limited" (16). He and others point to numerous examples of convergent evolution as evidence that selection finds the same adaptations despite the vagaries of history. Evolution may thus be broadly repeatable, and multiple replays would reveal striking similarities in important features, with contingency mostly confined to minor details (16–19).

Of course, replaying life's tape on the planetary scale is impossible, but careful experiments can examine the role of contingency in evolution on a more modest scale (15, 20, 21). To address the repeatability of evolutionary trajectories and outcomes, the long-term evolution experiment (LTEE) with *Escherichia coli* was started in 1988 with the founding of 12 populations from the same clone (2). These populations were initially identical except for a neutral marker that distinguished six lines from six others. They have since been propagated by daily 1:100 serial transfer in DM25, a minimal medium containing 25 mg/liter glucose as the limiting resource (2, 22). Environmental conditions have been controlled, constant, and identical for all 12 lines. To date, each population has evolved for >44,000 generations, and samples have been frozen every 500 generations, providing a rich "fossil record" (23). Moreover, these samples remain viable, allowing us to perform simultaneous measurements and other experiments with bacteria from different generations. The founding strain is strictly asexual, and thus populations have evolved by natural selection and genetic drift acting on variation generated solely by spontaneous mutations that occurred during the experiment. Thus, the LTEE allows us to examine the effects of contingency that are inherent to the core evolutionary processes of mutation, selection, and drift.

Previous analyses of this experiment have shown numerous examples of parallel phenotypic and genetic evolution. All twelve populations underwent rapid improvement in fitness that decelerated over time (2, 3, 22, 23). All evolved higher maximum growth rates on glucose, shorter lag phases upon transfer into fresh medium, reduced peak population densities, and larger average cell sizes relative to their ancestor (22–26). Ten populations evolved increased DNA supercoiling (27), and those populations examined to date show parallel changes in global

gene-expression profiles (4, 28, 29). At least three genes have substitutions in all 12 populations (30, 31), and several others have substitutions in many populations (27–30), even though most loci harbor no substitutions in any of them (32). At the same time, there has also been some divergence between populations. Four have evolved defects in DNA repair, causing mutator phenotypes (3, 33). There is subtle, but significant, between-population variation in mean fitness in the glucose-limited medium in which they evolved (2, 23). In media containing other carbon sources, such as maltose or lactose, the variation in performance is much greater (34). And while the same genes often harbor substitutions, the precise location and details of the mutations almost always differ between the populations (27–31).

Throughout the duration of the LTEE, there has existed an ecological opportunity in the form of an abundant, but unused, resource. DM25 medium contains not only glucose, but also citrate at a high concentration. The inability to use citrate as an energy source under oxic conditions has long been a defining characteristic of *E. coli* as a species (35, 36). Nevertheless, *E. coli* is not wholly indifferent to citrate. It uses a ferric dicitrate transport system for iron acquisition, although citrate does not enter the cell in this process (37, 38). It also has a complete tricarboxylic acid cycle, and can thus metabolize citrate internally during aerobic growth on other substrates (39). *E. coli* is able to ferment citrate under anoxic conditions if a cosubstrate is available for reducing power (40). The only known barrier to aerobic growth on citrate is its inability to transport citrate under oxic conditions (41–43). Indeed, atypical *E. coli* that grow aerobically on citrate (Cit⁺) have been isolated from agricultural and clinical settings, and were found to harbor plasmids, presumably acquired from other species, that encode citrate transporters (44, 45).

Other findings suggest that *E. coli* has the potential to evolve a Cit⁺ phenotype. Hall (41) reported the only documented case of a spontaneous Cit⁺ mutant in *E. coli*. He hypothesized that some complex mutation, or multiple mutations, activated cryptic genes that jointly expressed a citrate transporter, although the genes were not identified. Pos *et al.* (43) identified an operon in *E. coli* K-12 that apparently allows anaerobic citrate fermentation, and which includes a gene, *citT*, encoding a citrate–succinate antiporter. High-level constitutive expression of this gene on a multicopy plasmid allows aerobic growth on citrate, but the native operon has a single copy that is presumably induced only under anoxic conditions.

Despite this potential, none of the 12 LTEE populations evolved the capacity to use the citrate that was present in their environment for over 30,000 generations. During that time, each population experienced billions of mutations (22), far more than the number of possible point mutations in the ≈4.6-million-bp genome. This ratio implies, to a first approximation, that each population tried every typical one-step mutation many times. It must be difficult, therefore, to evolve the Cit⁺ phenotype, despite the ecological opportunity. Here we report that a Cit⁺ variant finally evolved in one population by 31,500 generations, and its descendants later rose to numerical dominance. The new Cit⁺ function has been the most profound adaptation observed during the LTEE and has had major consequences. As we will show, the population achieved a severalfold increase in size. Moreover, a stable polymorphism emerged, with a Cit⁻ minority coexisting with the new Cit⁺ majority. Interestingly, the population that evolved the Cit⁺ function is not one that had previously become hypermutable. It is also intriguing that this key innovation evolved so late in the experiment, given that the rate of fitness improvement had declined substantially in all of the populations (3, 23).

The long-delayed and unique evolution of the Cit⁺ phenotype might indicate that it required some unusually rare mutation,
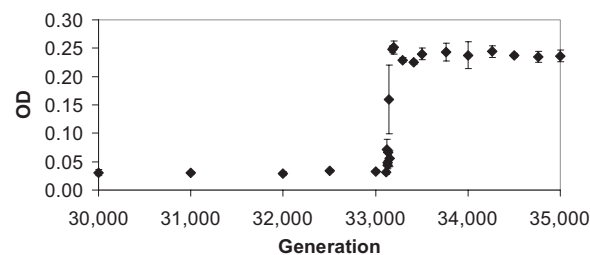

Fig. 1. Population expansion during evolution of the Cit⁺ phenotype. Samples frozen at various times in the history of population Ara-3 were revived, and three DM25 cultures were established for each generation. Optical density (OD) at 420 nm was measured for each culture at 24 h. Error bars show the range of three values measured for each generation.

such as a particular chromosomal inversion, that does not scale with typical mutation rates. Alternatively, the occurrence or phenotypic expression of the mutation that generated the Cit⁺ function might depend on one or more earlier mutations, such that its evolution was contingent on the particular history of that population. Contingent adaptations should tend to be complex and require multiple steps, some of which might not be beneficial, at least not uniquely so given other advantageous paths. Otherwise, cumulative selection would predictably favor the same steps, and the evolutionary path should be repeatable (18). Contingent adaptations should thus display two characteristics. First, independent origins should be rare, because the same historical sequences would rarely recur (19). Second, significant time-lags should occur between the presentation of ecological opportunities or challenges and the evolution of those traits that confer adaptation to those circumstances (46).

Can the hypothesis of contingent adaptation be rigorously tested in the case of the evolution of the Cit⁺ function? And can this hypothesis be formally distinguished from the alternative explanation that this new function required an unusually rare mutation, but one that was not historically contingent? The answer to both questions is yes, owing to certain features of bacteria in general and the LTEE in particular that allow us quite literally to replay evolution from various points in a population's history. We will first describe the emergence of this new function, and then present our experiments to distinguish between the hypotheses of mutational rarity and historical contingency in the origin of this key innovation.

Results

Evolution of Cit⁺ Function in Population Ara-3. The LTEE populations are transferred daily into fresh medium, and the turbidity of each is checked visually at that time. Owing to the low concentration of glucose in DM25 medium, the cultures are only slightly turbid when transferred. Occasional contaminants that grow on citrate have been seen over the 20 years of this experiment. These contaminated cultures reach much higher turbidity owing to the high concentration of citrate in the medium, which allows the contaminants to reach high density. (When contamination occurs, the affected population is restarted from the latest frozen sample.) After ≈33,127 generations, one population, designated Ara-3, displayed significantly elevated turbidity that continued to rise for several days (Fig. 1). A number of Cit⁺ clones were isolated from the population and checked for phenotypic markers characteristic of the ancestral *E. coli* strain used to start the LTEE: all were Ara⁻, T5-sensitive, and T6-resistant, as expected (2). DNA sequencing also showed that Cit⁺ clones have the same mutations in the *pykF* and *nadR* genes as do clones from earlier generations of the Ara-3 population, and each of these mutations distinguishes this population
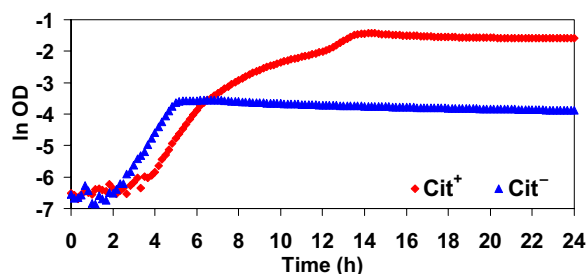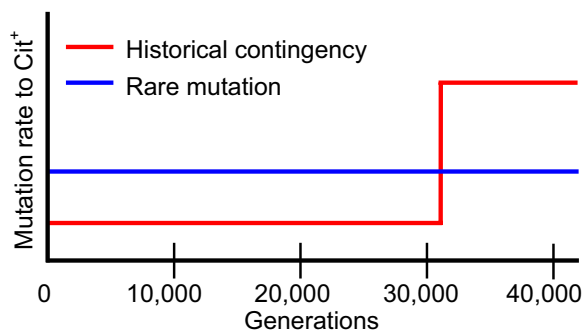
Fig. 2. Growth of Cit$^-$ (blue triangles) and Cit$^+$ (red diamonds) cells in DM25 medium. Each trajectory shows the average OD for eight replicate mixtures of three clones, all from generation 33,000 of population Ara-3.

Fig. 3. Alternative hypotheses for the origin of the Cit$^+$ function. According to the rare-mutation hypothesis, the probability of mutation from Cit$^-$ to Cit$^+$ was low but constant over time. Under the historical-contingency hypothesis, the probability of this transition increased when a mutation arose that produced a genetic background with a higher mutation rate to Cit$^+$.

from all of the others (30). Therefore, the Cit$^+$ variant arose within the LTEE and is not a contaminant.

The evolved Cit$^+$ variant grows to high density in DM0 (a citrate-only medium), produces vigorous colonies on minimal citrate (MC) agar plates, and causes a positive color change on Simmon's citrate agar, all of which indicate that it can use citrate as a sole carbon source. In DM25, Cit$^+$ cells undergo a period of rapid growth on glucose that is followed by slower growth on citrate (Fig. 2). Also, growth on citrate is inhibited by the citrate analog 5-fluorocitrate (data not shown), as was observed for the one previously reported Cit$^+$ mutant of *E. coli* (42, 43).

Cit$^+$ clones could be readily isolated from the frozen sample of population Ara-3 taken at generation 33,000. To estimate the time of origin of the Cit$^+$ trait, we screened 1,280 clones randomly chosen from generations 30,000, 30,500, 31,000, 31,500, 32,000, 32,500, and 33,000 for the capacity to produce a positive reaction on Christensen's citrate agar, which provides a sensitive means to detect even weakly citrate-using cells. No Cit$^+$ cells were found in the samples taken at 30,000, 30,500, or 31,000 generations. Cit$^+$ cells constituted ≈0.5% of the population at generation 31,500, then 15% and 19% in the next two samples, but only ≈1.1% at generation 33,000. It appears that the first Cit$^+$ variant emerged between 31,000 and 31,500 generations, although we cannot exclude an earlier origin. The precipitous decline in the frequency of Cit$^+$ cells just before the massive population expansion suggests clonal interference (47), whereby the Cit$^-$ subpopulation produced a beneficial mutant that outcompeted the emerging Cit$^+$ subpopulation until the latter evolved some other beneficial mutation that finally ensured its persistence. The hypothesis of clonal interference implies that the early Cit$^+$ cells were very poor at using citrate, such that a mutation that improved competition for glucose could have provided a greater advantage than did marginal exploitation of the unused citrate.

Indeed, the Cit$^+$ clones isolated from generations 32,500 and earlier grow much more slowly on citrate than those from 33,000 generations and later. After depleting the glucose in DM25, the earliest Cit$^+$ clones grow almost imperceptibly, if at all, for many hours before they begin efficiently using the citrate (data not shown), whereas later Cit$^+$ clones switch to growth on citrate almost immediately (Fig. 2). Thus, the population expansion between generations 33,000 and 33,500 (Fig. 1) was triggered by one or more mutations that improved citrate utilization, rather than by the growth of the original Cit$^+$ mutant. This finding also raises the question of whether weak Cit$^+$ mutants might exist in any other LTEE population. We therefore screened the other 11 populations, in most cases by using samples from generation 41,500. None of 220 cultures inoculated with heterogeneous population samples grew in glucose-free DM0, nor did any of >3,500 clones show a positive reaction on Christensen's citrate agar.

Historical Contingency in the Evolution of Cit$^+$. We performed three experiments to test whether the evolution of the Cit$^+$ function required an unusually rare mutation or, alternatively, was historically contingent and depended on the prior evolution of a certain genetic background. All three experiments used clones sampled from many generations of population Ara-3 to replay evolution starting from different genetic backgrounds. The two hypotheses make distinct predictions about the propensity of these backgrounds to re-evolve the Cit$^+$ phenotype (Fig. 3). According to the rare-mutation hypothesis, Cit$^+$ variants should evolve at the same low rate regardless of the generation of origin of the clone with which a replay started. By contrast, the historical-contingency hypothesis predicts that the mutation rate to Cit$^+$ should increase after some potentiating genetic background has evolved. Thus, Cit$^+$ variants should re-evolve more often in the replays using clones sampled from later generations of the Ara-3 population.

In our first experiment, we performed the replays under the same conditions as the LTEE. We isolated three random clones from each of twelve time points, from the ancestor to 32,500 generations (Table 1), and obtained neutral Ara$^+$ mutants of each clone to embed as protection against accidental cross-contamination during the experiment (2). In total, there were 72 replay populations, six from each generation, each founded by a single clone. These populations evolved for ≈3,700 generations, and they were checked visually each day for the elevated turbidity indicative of the Cit$^+$ phenotype. We also tested samples on MC and Christensen's citrate agar plates every 250 generations, with incubation for up to a week. New Cit$^+$ variants evolved in four replay populations, all founded by clones from later generations of the original population (Table 1). These Cit$^+$ variants emerged between 750 and 3,700 generations of the replay experiment.

Our second experiment also looked for Cit$^+$ mutants derived from clones sampled at various times in the history of population Ara-3. This time, however, we incubated large populations of cells on MC plates, enabling us to test more clones and more cells of each clone. We also allowed a long incubation time to facilitate the growth and detection of very weak Cit$^+$ mutants, as well as mutations that might occur as cells sat starving on the plates. In this experiment, ≈3.9 × 10^8 cells of each of the same 68 clones used in the first replay experiment were spread on each of five MC plates, and these 340 plates were then incubated for 59 days. Five plates produced Cit$^+$ mutants, and all used clones from generations 32,000 or 32,500 of the original population (Table 1). None of the particular clones that evolved Cit$^+$ in this experiment did so in the first one, although there was overlap in the generations from which those clones were sampled.

Table 1. Summary of replay experiments

Generation	First experiment		Second experiment		Third experiment	
	Replicates	Independent Cit+ mutants	Replicates	Independent Cit+ mutants	Replicates	Independent Cit+ mutants
Ancestor	6	0	10	0	200	0
5,000	—	—	—	—	200	0
10,000	6	0	30	0	200	0
15,000	—	—	—	—	200	0
20,000	6	0	30	0	200	2
25,000	6	0	30	0	200	0
27,000	—	—	—	—	200	2
27,500	6	0	30	0	—	—
28,000	—	—	—	—	200	0
29,000	6	0	30	0	200	0
30,000	6	0	30	0	200	0
30,500	6	1	30	0	—	—
31,000	6	0	30	0	200	1
31,500	6	1	30	0	200	1
32,000	6	0	30	4	200	2
32,500	6	2	30	1	200	0
Totals	72	4	340	5	2,800	8

The third replay experiment was similar in design to the second, but on a larger scale. We isolated 20 clones from each of 13 time points in the history of population Ara-3, again through 32,500 generations. We generated and tested 10 replicate cultures of each evolved clone and 200 replicates of the ancestor. Each culture grew to ≈ 1–2×10^{10} cells, which were pelleted by centrifugation, spread on an MC plate, and incubated for 45 days. In total, $\approx 4.0 \times 10^{13}$ cells were tested for their ability to use citrate in this experiment. Cit+ mutants emerged from eight cultures representing seven clones (Table 1), with one clone yielding mutants in two replicate cultures. Four clones that produced Cit+ mutants came from generations 31,000 and later, two were from generation 27,000, and one (the one that produced two mutants) was from generation 20,000. We found no Cit+ mutants among any of the 200 ancestral cultures, nor among any of the other 600 cultures that used clones isolated before generation 20,000.

Interestingly, 7 of the 8 plates that yielded a mutant Cit+ colony produced multiple colonies, including one with 137 colonies. This pattern illustrates the "jackpot" effect discovered by Luria and Delbrück (48), and it implies that mutations arose during the population growth before plating on MC agar. On the other hand, the Cit+ colonies were not observed until at least 8 days of incubation and, in one case, they were first seen after 28 days. These late appearances suggest that the mutations to Cit+ occurred after plating. One possible explanation for this apparent discrepancy is that the mutants grow very slowly but, in fact, they typically produce visible colonies in only 2 days when retested on MC plates. Another potential explanation is that the high density of Cit− cells on the plates interfered with the growth and detection of emerging Cit+ colonies. To test this possibility, we seeded dense Cit− cultures with a few Cit+ cells before plating on MC agar, but the Cit+ colonies were seen after only 2–3 days. The rapid growth of Cit+ colonies occurred even when the Cit+ cells had grown on glucose, and not on citrate, before plating. These results imply that mutations to Cit+ occurred after cultures were plated on MC agar. This conclusion, taken together with the jackpot distribution of Cit+ mutants, indicates that the phenotypic change required two mutations, one of which occurred during the culture growth before plating and the other after plating.

Statistical Analysis of the Replay Experiments. All three experiments show the same tendency for Cit+ variants to evolve more often from clones sampled in later than earlier generations of population Ara-3 (Table 1). To calculate the significance of these data, we performed Monte Carlo resampling tests (shuffling without replacement) by using the Statistics101 Re-sampling Simulator version 1.0.6 (www.statistics101.net). For each experiment, we compared the observed mean generation of those clones that yielded Cit+ variants to the mean expected under the null hypothesis that clones from all generations have equal likelihood. The null thus corresponds to the rare-mutation hypothesis laid out in the Introduction. We ran one million resampling iterations for each experiment. The deviations from the null expectations range from marginally to highly significant in the three experiments, and in all cases they support the historical-contingency hypothesis, according to which clones from later generations have greater propensity to evolve the Cit+ phenotype (Table 2). Although the third experiment was the largest, it was the least significant, owing primarily to the production of two Cit+ mutants by a 20,000-generation clone. We also used the Z-transformation method (49) to combine the probabilities from our three experiments, and the result is extremely significant ($P < 0.0001$) whether or not the experiments are weighted by the number of independent Cit+ mutants observed in each one. Furthermore, the potentiation effect in later generations is underestimated by these tests, because the number of cells trended lower in later-generation cultures owing to the evolution of larger cells that reach lower population density (24, 25, 41, 50).

These analyses compel us to reject the hypothesis that a rare mutation could have produced a Cit+ variant with equal probability at any point in the LTEE. Some unusually rare mutation

Table 2. Statistical analyses of three replay experiments

	Mean generation of clones yielding Cit+		
	First experiment	Second experiment	Third experiment
Expected	24,917	28,382	22,571
Observed	31,750	32,100	27,563
Monte Carlo P value	0.0085	0.0007	0.0823
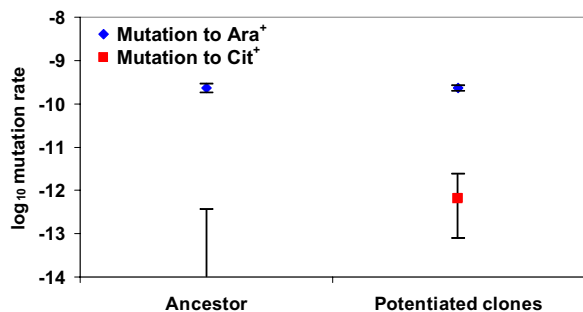
Fig. 4. Mutation rates from Ara⁻ to Ara⁺ (blue diamonds) and Cit⁻ to Cit⁺ (red squares) of the ancestor and a set of potentiated clones. Error bars are 95% confidence intervals. See text for details.

Fig. 5. Frequency-dependent selection allows stable coexistence of Cit⁻ and Cit⁺ clones. Each trajectory shows the mean of five replicate cultures for seven different initial ratios of Cit⁻ and Cit⁺ clones from generation 33,000. Error bars are 95% confidence intervals.
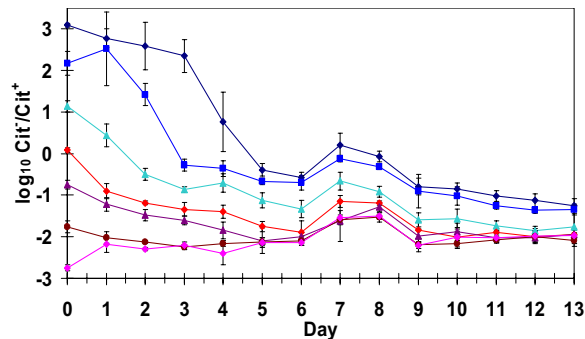
might be involved, but its rarity does not provide a sufficient explanation for the unique and exceptionally slow evolution of this new function during the LTEE. Our results instead support the hypothesis of historical contingency, in which a genetic background arose that had an increased potential to evolve the Cit⁺ phenotype.

Rates of Mutation to Cit⁺ and Quantification of the Potentiation Effect. Given the fraction of cultures that produce no mutants (P_0) and the number of cells per culture (N), one can estimate the mutation rate as $\mu = -(\ln P_0)/N$. The corresponding 95% confidence interval (CI) can also be calculated from the Poisson distribution based on the uncertainty in P_0, which is much greater than the uncertainty in N. By comparing the mutation rate to Cit⁺ between clones, one can in principle quantify the potentiation effect in the evolved genetic background. However, we cannot unambiguously identify which evolved clones remain nonpotentiated owing to the very low mutation rate to Cit⁺ even in the potentiated clones. We also cannot separate potentiated and nonpotentiated clones by generation because some generational samples may be polymorphic.

We therefore performed an independent series of fluctuation tests using seven clones that yielded Cit⁺ mutants in at least one replay experiment. These additional tests permit an unbiased estimate of the mutation rate to Cit⁺ in the potentiated background. We prepared 40 replicate cultures for each potentiated clone, and 280 for the ancestor. Potentiated and ancestral cultures had, on average, 1.1 and 1.5×10^{10} cells, respectively, which were harvested and incubated on MC agar plates for 45 days.

None of the ancestral cultures yielded any Cit⁺ mutants. We can nonetheless calculate an upper limit to the mutation rate by noting that the Poisson distribution has a 5% probability of yielding zero events when the expectation is three. With no more than three mutations among the 8.4×10^{12} cells tested here and in the third replay experiment, the upper bound on the ancestral mutation rate to Cit⁺ is 3.6×10^{-13} per cell per generation (Fig. 4). To the best of our knowledge, this value is the lowest upper bound ever reported for a mutation rate that has been experimentally measured. It is also probably far too high because no mutations were actually observed for the ancestor, nor were any found among another 9.0×10^{12} cells of 60 clones sampled through 15,000 generations; and because some cell turnover and other DNA activity probably occurred during the many days that plates were incubated.

Even among the potentiated clones, the rate of mutation to Cit⁺ is extremely low. Cit⁺ mutants arose in 2 of the 280 new cultures, giving an estimate of 6.6×10^{-13} for the mutation rate, with the 95% CI extending from 7.9×10^{-14} to 2.4×10^{-12} (Fig. 4). (Although the upper bound of the CI for the ancestor overlaps the lower bound of the CI for the potentiated clones,

that upper bound does not overlap the point estimate for the potentiated clones, indicating a significant difference and adding further support to the replay experiments.) The potentiated genetic background thus increases the mutation rate to Cit⁺ at least 2-fold, and probably much more. However, even this potentiated value represents an unusually low mutation rate. A typical mutation rate in *E. coli* is $\approx 5 \times 10^{-10}$ per base pair per generation (51). Such a low rate suggests that the final mutation to Cit⁺ is not a point mutation but instead involves some rarer class of mutation or perhaps multiple mutations. The possibility of multiple mutations is especially relevant, given our evidence that the emergence of Cit⁺ colonies on MC plates involved events both during the growth of cultures before plating and during prolonged incubation on the plates.

Another issue is whether the potentiating effect might indicate the evolution of a generalized hypermutability in population Ara-3. Previous surveys of mutation rates in the 12 LTEE lines found that 4 had become mutators by generation 20,000, although population Ara-3 retained the low ancestral mutability (3, 33). To investigate this issue further, we performed another series of fluctuation tests using the same seven potentiated clones to determine their rate of mutation from Ara⁻ to Ara⁺, which serves as a proxy for the background mutation rate (33). This mutation reverts a phenotype that was knocked out during the derivation of a predecessor to the ancestor of the LTEE (2, 52). Sequence analysis indicates that most Ara⁺ mutants have a mutation from GAC to GGC at codon 92 of the *araA* gene, restoring that codon to its distant ancestral state. Forty cultures of each clone were grown in DM1000 (1 g of glucose per liter), and the cells were spread on minimal agar plates containing arabinose. At the same time, 280 cultures of the ancestor were tested in the same way. The ancestral mutation rate to Ara⁺ is 2.3×10^{-10} per cell per generation calculated by using the P_0 method (95% CI from 1.8×10^{-10} to 2.9×10^{-10}), a typical rate for point mutations. The rate for potentiated clones is also 2.3×10^{-10} (95% CI from 2.0×10^{-10} to 2.7×10^{-10}); in fact, these estimates are within 1% of each other (Fig. 4). We conclude that general hypermutability is not responsible for the elevated mutation rate to Cit⁺ in the potentiated clones.

Frequency-Dependent Selection Maintains Ecological Diversity. To this point, we have examined the evolutionary origin of the Cit⁺ function in population Ara-3. We now turn to its ecological consequence. The Cit⁺ phenotype did not achieve fixation during the population expansion but, instead, Cit⁻ cells persisted as a minority. When we mixed Cit⁻ and Cit⁺ clones at different initial frequencies, they stably coexisted over many serial transfers (Fig. 5). In these mixtures, the Cit⁻ cells gradually approached an equilibrium frequency of roughly 1%, regardless of

their initial frequency. We saw a transient jump in the frequency of the Cit$^-$ subpopulation on days 7 and 8, which was probably caused by accidentally using a glucose-only medium on those days. After that perturbation, the populations resumed their previous trajectories. Negative frequency-dependent selection thus maintains the polymorphism.

This stable coexistence suggests that the Cit$^-$ cells are superior to the Cit$^+$ cells in competition for glucose, allowing the former to persist as glucose specialists. Indeed, the Cit$^-$ cells have a shorter lag phase and higher growth rate on glucose than do the Cit$^+$ cells (Fig. 2). These differences were also evident when we monitored the intraday dynamics of mixtures of Cit$^-$ and Cit$^+$ cells (data not shown).

Discussion and Future Directions

E. coli cells cannot grow on citrate under oxic conditions, and that inability has long been viewed as a defining characteristic of this important, diverse, and widespread species. In a long-term experiment, we propagated 12 populations of *E. coli*, all founded from the same ancestral strain, in a medium containing glucose, which is the limiting resource, and abundant citrate. For more than 30,000 generations, none of them evolved the capacity to use the citrate, although billions of mutations occurred in each population, such that any typical base pair mutation would have been tested many times in each one. It is clearly very difficult for *E. coli* to evolve this function. In fact, the mutation rate of the ancestral strain from Cit$^-$ to Cit$^+$ is immeasurably low; even the upper bound is 3.6×10^{-13} per cell generation, which is three orders of magnitude below the typical base pair mutation rate. Nevertheless, one population eventually evolved the Cit$^+$ function, whereas all of the others remain Cit$^-$ after more than 40,000 generations.

We demonstrated that the evolution of this new function was contingent on the history of the population in which it arose. In particular, we showed that one or more earlier mutations potentiated the evolution of this function by increasing the mutation rate to Cit$^+$, although even the elevated rate is much lower than a typical mutation rate. The potentiated cells are not generally hypermutable. Rather, their potentiation appears to be specific to the Cit$^+$ function, which suggests two possible mechanisms. One mechanism is epistasis, whereby the functional expression of the mutation that finally yielded the Cit$^+$ phenotype requires interaction with one or more mutations that evolved earlier. A second possibility is that the physical production of the mutation that produced the Cit$^+$ phenotype requires some previous mutation that allows the final sequence to be generated. For example, the insertion of a mobile genetic element creates new sequences at its junctures, and one of these new sequences might then undergo a mutation that generates a final sequence that could not have occurred without the insertion. The *E. coli* genome has many insertion-sequence elements (53), some of which have been active in the LTEE (54–56). Whatever the mechanism, this potentiation made the Cit$^+$ function mutationally accessible, and a weak Cit$^+$ variant emerged by 31,500 generations.

The origin of the Cit$^+$ function also had profound consequences for the ecology and subsequent evolution of that population. This new capacity was refined over the next 2,000 generations, leading to a massive population expansion as the Cit$^+$ cells evolved to exploit more efficiently the abundant citrate in their environment. Although the Cit$^+$ cells continued to use glucose, they did not drive the Cit$^-$ subpopulation extinct because the Cit$^-$ cells were superior competitors for glucose. Thus, the overall diversity increased as one population gave rise evolutionarily to an ecological community with two members, one a resource specialist and the other a generalist.

The evolution of the new Cit$^+$ function represents a key innovation that involves multiple steps, and it provides an explicit demonstration of the importance of historical contingency in evolution. It also transcends the phenotypic boundaries of a diverse and well studied species, and led to an ecological transition from a single population to a two-member community. Our future research on this fascinating case of evolution in action will revolve around four themes: genetics, physiology, ecology, and speciation.

What is the genetic basis of this evolutionary innovation? The emergence of the Cit$^+$ phenotype in population Ara-3 indicates at least two important genetic events: the origin of the function in its weak form, and its subsequent refinement for efficient use of citrate. The replay experiments indicate an even more complex picture that must involve, at a minimum, three important genetic events. At least one mutation in the LTEE was necessary to produce a genetic background with the potential to generate Cit$^+$ variants, while the distribution and dynamics of Cit$^+$ mutants in fluctuation tests indicate at least two additional mutations are involved. To find the relevant mutations, we will perform whole-genome resequencing, which has become a powerful approach that is well suited to experimental evolution (57–59). We expect to find dozens of mutations relative to the ancestor (22), which will complicate identification of those changes that were important specifically for the origin of the Cit$^+$ function. However, some of the key changes should become apparent if we also resequence a Cit$^-$ clone from the same population around the time that the Cit$^+$ variants first emerged. Once candidate genes and mutations have been identified, we can examine the other 19 Cit$^+$ variants from our replay and mutation-rate experiments for parallel changes.

We are especially eager to find the potentiating mutation or mutations. We want to know whether the potentiating mutation interacts epistatically with a later mutation to allow expression of the Cit$^+$ function or, alternatively, whether it was physically required for the later mutation to occur. We also want to test whether the potentiating mutation was itself beneficial or, alternatively, a neutral or deleterious change that fortuitously hitchhiked to high frequency. We anticipate that identifying the potentiating mutation will be especially challenging, however, because its only known phenotype is to increase the rate of production of certain mutants that are themselves extremely rare.

Once we have identified all of the relevant mutations, it might be interesting to model the population dynamics that govern the emergence of this new function. Such a model would require not only all of the relevant mutation rates but also the ecological phenotypes of the mutants, including their growth rates on glucose and citrate as well as their abilities to transition between the two resources. A satisfactory model should also reflect the stochastic origin of mutations, the role of random drift, and the possibility of alternative mutational paths to the phenotype of interest (9, 60).

What physiological mechanism has evolved that allows aerobic growth on citrate? *E. coli* should be able to use citrate as an energy source after it enters the cell, but it lacks a citrate transporter that functions in an oxygen-rich environment. One possibility is that the Cit$^+$ lineage activated a "cryptic" transporter (41), that is, some once-functional gene that has been silenced by mutation accumulation. This explanation seems unlikely to us because the Cit$^-$ phenotype is characteristic of the entire species, one that is very diverse and therefore very old. We would expect a cryptic gene to be degraded beyond recovery after millions of years of disuse. A more likely possibility, in our view, is that an existing transporter has been coopted for citrate transport under oxic conditions. This transporter may previously have transported citrate under anoxic conditions (43) or, alternatively, it may have transported another substrate in the presence of oxygen. The evolved changes might involve gene regulation, protein structure, or both (61).

What will be the long-term fates of the coexisting Cit⁻ and Cit⁺ subpopulations? We showed they stably coexist owing to the inferiority of the Cit⁺ cells in competition for glucose. However, the Cit⁺ lineage might eventually acquire mutations that compensate for its inferior performance on glucose, thus undermining the coexistence. Compensation for maladaptive side-effects of adaptations, including resistance to phages and antibiotics, has been observed in many other experiments with bacteria (62–66). Moreover, the Cit⁺ subpopulation is much larger than the Cit⁻ subpopulation, so it should experience more beneficial mutations even without compensation. On the other hand, coexistence would be strengthened if selection in the Cit⁺ subpopulation favors specialization on citrate. In the same way that we established multiple populations for retrospective re-plays in this study, we can establish multiple communities to examine their prospective evolution. We can also vary environmental factors, such as the presence or absence of glucose, and the presence or absence of the Cit⁻ subpopulation, to investigate how they influence the future evolution of the Cit⁺ lineage.

Will the Cit⁺ and Cit⁻ lineages eventually become distinct species? According to the biological species concept widely used for animals and plants, species are recognized by reproductive continuity within species and reproductive barriers leading to genetic isolation between species (67). Although the bacteria in the LTEE are strictly asexual, we can nonetheless imagine testing this criterion by producing recombinant genotypes. In particular, we could move mutations that are substituted in the evolving Cit⁺ lineage into a Cit⁻ background to test whether they reduce fitness in their ancestral context. One could also perform the reciprocal experiment, although we anticipate more rapid evolution in the Cit⁺ lineage because it has acquired a key innovation that substantially changed its ecological niche. Such experiments would require, of course, controls to examine the fitness effects of the same mutations in the lineage where they arose. If the Cit⁺ lineage is indeed evolving into a new species, then we expect, with time, that more and more of the beneficial mutations substituted in that lineage would be detrimental in the ecological and genetic context of its Cit⁻ progenitor.

In any case, our study shows that historical contingency can have a profound and lasting impact under the simplest, and thus most stringent, conditions in which initially identical populations evolve in identical environments. Even from so simple a beginning, small happenstances of history may lead populations along different evolutionary paths. A potentiated cell took the one less traveled by, and that has made all the difference.

Materials and Methods

The Long-term Evolution Experiment. The LTEE is described in detail elsewhere (2, 22). Briefly, two ancestral clones of *E. coli* B were each used to found six populations. The ancestors differ by a single mutation that allows one of them to use arabinose (Ara⁺). Ara⁻ and Ara⁺ cells make red and white colonies, respectively, on tetrazolium–arabinose (TA) plates, but the mutation is neutral in the environment of the LTEE (2). The twelve populations have been propagated for almost 20 years by daily serial dilution in DM25, a minimal salts medium that has 139 μM glucose and 1,700 μM citrate (2). Given 1:100 dilution and regrowth, the populations achieve $\approx$6.64 generations per day, and they have evolved for over 40,000 generations in this experiment to date. Every 500 generations, population samples are frozen at -80°C with glycerol added as a cryoprotectant. These samples contain all of the diversity present in a

population at that generation. Except as otherwise noted, all bacteria used in this study come from samples of population Ara-3.

Media and Culture Conditions. DM0 is identical to DM25 except that it contains no glucose. MC plates have the same formulation as liquid DM0, except that agar is added to solidify the medium and the concentration of citrate is higher to allow large colonies. Before starting an experiment, bacteria were preconditioned in the relevant experimental medium unless otherwise stated. For additional details, see supporting information (SI) *Materials and Methods*.

Analysis of the Population Expansion. When the increased turbidity was observed in population Ara-3, we froze samples of the two previous days' cultures, which had been temporarily stored at 4°C. We also froze samples after every transfer for several weeks. We later revived these and other samples taken between generations 30,000 and 35,000 to quantify the population expansion. For details, see *SI Materials and Methods*.

Searches for Cit⁺ Variants. To determine when the earliest Cit⁺ variant arose in population Ara-3, we randomly chose 1,280 clones from samples frozen at generations 30,000, 30,500, 31,000, 31,500, 32,000, 32,500, and 33,000. We transferred colonies to Christensen's citrate agar plates, which were incubated for 10 days and examined daily for evidence of positive reactions. All putative Cit⁺ clones were retested on Christensen's agar for confirmation, and a subset was tested on MC agar. All clones that had positive reactions on Christensen's plates also grew on MC plates. To find Cit⁺ cells that might be present at low frequency in the other 11 LTEE populations, we tested their genetically heterogeneous full-population samples for growth in DM0. We allowed the top of each sample to thaw and inoculated 50 μl into LB broth. After overnight growth, five 10-ml DM0 cultures were inoculated with 100 μl of 100-fold dilutions of each LB culture. The DM0 cultures were incubated for 28 days with periodic checks for turbidity.

Growth Curves. Three representative Cit⁻ clones and three representative Cit⁺ clones from generation 33,000 were separately preconditioned in DM25. The Cit⁻ clones were then combined, as were the Cit⁺ clones, and the mixtures were diluted in DM25 to $\approx$3.4 × 10⁴ cells per ml. Eight 200-μl aliquots of each mixture were dispensed into randomly assigned wells in a 96-well plate, and optical density (OD) at 420-nm wavelength was measured periodically by using a VersaMax automated plate reader (Molecular Devices).

Replay Experiments. In the first replay experiment, populations evolved under the same conditions as the LTEE. In the second and third replay experiments, cells were incubated on MC plates. We started the first replay experiment on the 3rd anniversary of Stephen Jay Gould's death; we ended it on the 66th anniversary of his birth. For further details, see *SI Materials and Methods*.

Confirmation of Cit⁺ Variants. Each putative Cit⁺ variant was streaked on MC agar and Christensen's citrate agar to confirm its phenotype. Additional tests are described in the *SI Materials and Methods*.

Test for Frequency-Dependent Interaction. We constructed mixtures of clones from generation 33,000 to test whether Cit⁻ and Cit⁺ clones coexisted and, if so, whether they interacted in a frequency-dependent manner. For details, see *SI Materials and Methods*.

1. Mani GS, Clarke BC (1990) Mutational order: A major stochastic process in evolution. *Proc R Soc London B* 240:29–37.
2. Lenski RE, Rose MR, Simpson SC, Tadler SC (1991) Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am Nat* 138:1315–1341.
3. Cooper VS, Lenski RE (2000) The population genetics of ecological specialization in evolving *E. coli* populations. *Nature* 407:736–739.
4. Cooper TF, Remold SK, Lenski RE, Schneider D (2008) Expression profiles reveal parallel evolution of epistatic interactions involving the CRP regulon in *Escherichia coli*. *PLoS Genet* 4:e35.
5. Jacob F (1977) Evolution and tinkering. *Science* 196:1161–1166.
6. Wright S (1988) Surfaces of selective value revisited. *Am Nat* 131:115–123.
7. Yedid G, Bell G (2002) Macroevolution simulated with autonomously replicating computer programs. *Nature* 420:810–812.
8. Weinreich DM, Watson RA, Chao L (2005) Sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* 59:1165–1174.
9. Weinreich DM, Delaney NF, DePristo MA, Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312:111–114.
10. Lewontin RC (1966) Is nature probable or capricious? *BioScience* 16:25–27.

11. Gould SJ (1985) The paradox of the first tier: An agenda for paleobiology. *Paleobiology* 11:2–12.
12. Jablonski D (1986) Background and mass extinctions: The alternation of macroevolutionary regimes. *Science* 231:129–133.
13. Gould SJ (1989) *Wonderful Life* (Norton, New York).
14. Gould, SJ (2002) *The Structure of Evolutionary Theory* (Belknap, Cambridge, MA).
15. Beatty J (2006) Replaying life's tape. *J Philos* 7:336–362.
16. Conway Morris S (2003) *Life's Solution* (Cambridge Univ Press, Cambridge, UK).
17. Van Valen LM (1991) How far does contingency rule? *Evol Theory* 10:47–52.
18. Dawkins R (1996) *The Blind Watchmaker* (Norton, New York).
19. Vermeij GJ (2006) Historical contingency and the purported uniqueness of evolutionary innovations. *Proc Natl Acad Sci USA* 103:1804–1809.
20. Travisano M, Mongold JA, Bennett AF, Lenski RE (1995) Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* 267:87–90.
21. Mongold JA, Bennett AF, Lenski RE (1999) Evolutionary adaptation to temperature. VII. Extension of the upper thermal limit of *Escherichia coli*. *Evolution* 53:386–394.
22. Lenski RE (2004) Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*. *Plant Breed Rev* 24:225–265.
23. 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.
24. Vasi F, Travisano M, Lenski RE (1994) Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. *Am Nat* 144:432–456.
25. Lenski RE, Mongold JA (2000) in *Scaling in Biology*, eds Brown J, West G (Oxford Univ Press, Oxford, UK), pp 221–235.
26. Novak M, Pfeiffer T, Lenski RE, Sauer U, Bonhoeffer S (2006) Experimental tests for an evolutionary trade-off between growth rate and yield in *E. coli*. *Am Nat* 168:242–251.
27. 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.
28. 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.
29. Pelosi L, *et al.* (2006) Parallel changes in global protein profiles during long-term experimental evolution in *Escherichia coli*. *Genetics* 173:1851–1869.
30. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE (2006) Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc Natl Acad Sci USA* 103:9107–9122.
31. Cooper VS, Schneider S, Blot M, Lenski RE (2001) Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of *E. coli* B. *J Bacteriol* 183:2834–2841.
32. Lenski RE, Winkworth CL, Riley MA (2003) Rates of DNA sequence evolution in experimental populations of *Escherichia coli* during 20,000 generations. *J Mol Evol* 56:498–508.
33. Sniegowski PD, Gerrish PJ, Lenski RE (1997) Evolution of high mutation rates in experimental populations of *Escherichia coli*. *Nature* 387:703–705.
34. Travisano M, Vasi F, Lenski RE (1995) Long-term experimental evolution in Escherichia coli. III. Variation among replicate populations in correlated responses to novel environments. *Evolution* 49:189–200.
35. Koser SA (1924) Correlation of citrate-utilization by members of the colon-aerogenes group with other differential characteristics and with habitat. *J Bacteriol* 9:59–77.
36. Scheutz F, Strockbine NA (2005) Genus I. *Escherichia*, Castellani and Chalmers 1919, *Bergey's Manual of Systematic Bacteriology, Volume Two: The Proteobacteria*, eds Garrity GM, Brenner DJ, Kreig NR, Staley JR (Springer, New York), pp 607–624.
37. Frost GE, Rosenberg H (1973) The inducible citrate-dependent iron transport system in Escherichia coli K-12. *Biochim Biophys Acta* 330:90–101.
38. Hussein S, Hantke K, Braun V (1981) Citrate-dependent iron transport system in *Escherichia coli* K-12. *Eur J Biochem* 117:431–437.
39. Lara FJS, Stokes JL (1952) Oxidation of citrate by *Escherichia coli*. *J Bacteriol* 63:415–420.
40. Lutgens M, Gottschalk G (1980) Why a co-substrate is required for anaerobic growth of *Escherichia coli* on citrate. *J Gen Microbiol* 199:63–70.
41. Hall BG (1982) Chromosomal mutation for citrate utilization by *Escherichia coli* K-12. *J Bacteriol* 151:269–273.
42. Reynolds CH, Silver S (1983) Citrate utilization by *Escherichia coli*: Plasmid- and chromosome-encoded systems. *J Bacteriol* 156:1019–1024.
43. Pos KM, Dimroth P, Bott M (1998) The *Escherichia coli* citrate carrier CitT: A member of a novel eubacterial transporter family related to the 2-oxoglutarate-malate translocator from spinach chloroplasts. *J Bacteriol* 180:4160–4165.
44. Ishiguro N, Oka C, Sato G (1978) Isolation of citrate positive variants of *Escherichia coli* from domestic pigeons, pigs, cattle, and horses. *Appl Environ Microbiol* 36:217–222.
45. Ishiguro N, Oka C, Hanazawa Y, Sato G (1979) Plasmids in *Escherichia coli* controlling citrate-utilizing ability. *Appl Environ Microbiol* 38:956–964.
46. Foote MJ (1998) Contingency and convergence. *Science* 280:2068–2069.
47. Gerrish PJ, Lenski RE (1998) The fate of competing beneficial mutations in an asexual population. *Genetica* 102/103:127–144.
48. Luria SE, Delbrück M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28:491–511.
49. Whitlock MC (2005) Combining probability from independent tests: The weighted Z-method is superior to Fisher's approach. *J Evol Biol* 18:1368–1373.
50. Elena SF, Cooper VS, Lenski RE (1996) Punctuated evolution caused by selection of rare beneficial mutations. *Science* 272:1802–1804.
51. Drake JW (1991) A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA* 88:7160–7164.
52. Lederberg S (1966) Genetics of host-controlled restriction and modification of deoxyribonucleic acid in *Escherichia coli*. *J Bacteriol* 91:1029–1036.
53. Schneider D, *et al.* (2002) Genomic comparisons among *Escherichia coli* strains B, K-12, and O157:H7 using IS elements as molecular markers. *BMC Microbiol* 2:18.
54. Papadopoulos D, *et al.* (1999) Genomic evolution during a 10,000-generation experiment with bacteria. *Proc Natl Acad Sci USA* 96:3807–3812.
55. Schneider D, Duperchy E, Coursange E, Lenski RE, Blot M (2000) Long-term experimental evolution in *Escherichia coli*. IX. Characterization of IS-mediated mutations and rearrangements. *Genetics* 156:477–488.
56. Schneider D, Lenski RE (2004) Dynamics of insertion sequence elements during experimental evolution of bacteria. *Res Microbiol* 155:319–327.
57. Velicer GJ, *et al.* (2006) Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. *Proc Natl Acad Sci USA* 103:8107–8112.
58. Herring CD, *et al.* (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* 38:1406–1412.
59. Hegreness M, Kishony R (2007) Analysis of genetic systems using experimental evolution and whole-genome sequencing. *Genome Biol* 8:201.
60. Lenski RE, Ofria C, Pennock RT, Adami C (2003) The evolutionary origin of complex features. *Nature* 423:139–144.
61. Mortlock RP, ed (1984) *Microorganisms as Model Systems for Studying Evolution* (Plenum, New York).
62. Lenski RE (1988) 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.
63. Bouma JE, Lenski RE (1988) Evolution of a bacteria/plasmid association. *Nature* 335:351–352.
64. Schrag SJ, Perrot V, Levin BR (1997) Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. *Proc R Soc London B* 264:1287–1291.
65. Lenski RE (1997) The cost of antibiotic resistance—from the perspective of a bacterium. *Ciba Found Symp* 207:131–140.
66. Bjorkman J, Hughes D, Andersson DI (1998) Virulence of antibiotic-resistant *Salmonella typhimurium*. *Proc Natl Acad Sci USA* 95:3949–3953.
67. Mayr E (1942) *Systematics and the Origin of Species from the Viewpoint of a Zoologist* (Columbia Univ Press, New York).