13. C. D. Michener, *Rev. Biol. Trop.* **10**, 167 (1962).
14. S. L. Buchmann and J. P. Hurley, *J. Theor. Biol.* **72**, 639 (1978).
15. S. L. Buchmann, in *Handbook of Experimental Pollination Biology*, C. E. Jones and R. J. Little, Eds. (Scientific and Academic Editions, New York, 1983), pp. 73–113.
16. For useful comments on the manuscript, we thank H. Markl, A. Michelsen, and E. O. Wilson and two anonymous reviewers. For fruitful discussions, we thank B. Assmann, O. N. Larsen, and N. Franks. A. Rakowska made the drawings. Financial support came from the Deutsche Forschungsgemeinschaft Leibniz Prize (B.H.) and grant SFB 251/TP 18.

12 July 1994; accepted 31 October 1994

Experimental Tests of the Roles of Adaptation, Chance, and History in Evolution

Michael Travisano, Judith A. Mongold, Albert F. Bennett, Richard E. Lenski*

The contributions of adaptation, chance, and history to the evolution of fitness and cell size were measured in two separate experiments using bacteria. In both experiments, populations propagated in identical environments achieved similar fitnesses, regardless of prior history or subsequent chance events. In contrast, the evolution of cell size, a trait weakly correlated with fitness, was more strongly influenced by history and chance.

The diversity of organisms is the product of three fundamental evolutionary influences: adaptation, chance, and history. Their relative contributions to evolutionary change have been the subject of intense debate (*1*). Adaptation has sometimes been regarded as the sole influence on evolution, and some biologists have invoked natural selection to explain almost any phenotypic difference. Unsubstantiated claims that adaptation is the cause of all biological diversity have prompted critics to offer two alternative causes, chance and history, that might account for any particular phenotypic difference. Chance effects include mutation and genetic drift, which govern the stochastic appearance and subsequent loss or fixation of new traits. Chance is usually invoked in the context of molecular genetic traits that are selectively neutral; however, chance is also important for phenotypic evolution, because beneficial mutations arise at random and may be lost soon after they appear, even in large populations. Other evolutionists have emphasized the effects of history, which may constrain or promote particular evolutionary outcomes according to the genetic and developmental integration of the ancestral phenotype. In this view, the set of potential adaptations is severely limited by inherited constitution, so that at every moment the course of evolution is contingent on prior (historical) events.

S. J. Gould (*2*) has argued for the great importance of historical contingency. He has presented a gedanken experiment of "replaying life's tape" to test the repeatability of evolution and thereby evaluate the roles of adaptation, chance, and history. Of course, one cannot perform an actual experiment on the grand scale envisioned by Gould, but one can perform rigorous experiments, of shorter duration and in simpler environments, to quantify the roles of adaptation, chance, and history in evolution. Instead of replaying life's tape sequentially, one can achieve the same objective by doing an experiment in which replicate populations are propagated simultaneously.

Imagine, first, that a single ancestral genotype is used to create a set of initially identical populations that will be propagated in identical environments. If one measures the initial mean value of some trait for each population, one should find that they are identical within statistical limits of measurement error. If one subsequently measures the mean value of this trait for each of the derived populations, one may still find that none of the derived populations has changed significantly relative either to their common ancestor or to one another (Fig. 1A). In that case, one would conclude that the trait had not evolved. Alternatively, one might find that, although there was no significant change in the grand mean (over all populations) from the ancestral value, there was significant variation among the derived populations (Fig. 1B). One would attribute this among-population variation to chance, because the derived populations had identical ancestors and were subject to identical environments. This chance divergence might reflect mutation or drift or their interactions with other evolutionary processes; attributing this variation to chance makes no specific claims in that regard. A third possible outcome is that the grand mean of the derived populations changed significantly from the value for the ancestor but without significant variation among the replicate populations (Fig. 1C). One would attribute this systematic change in mean value of a trait to adaptation. By invoking adaptation, we do not necessarily mean that the trait was the actual target of selection; it might instead be correlated with some other trait that was selected. Nor do we mean that stochastic processes were not involved; for example, adaptation may depend on random mutations, but similar mutations may be common enough to permit parallel evolution in the replicate populations. A fourth possibility is that both chance and adaptation contribute significantly to the trait's evolution (Fig. 1D).

To visualize the effects of history, imagine that a similar experiment is done using several different ancestral genotypes. One might observe that any initial variation in the value of some trait among ancestral genotypes was eliminated from the derived populations because of the effects of adaptation or chance or both (Fig. 1E). That is, the statistical contribution of initial genetic composition to the value of the derived trait was lost, so that one cannot reconstruct a derived genotype's ancestry using that trait. Alternatively, one might observe

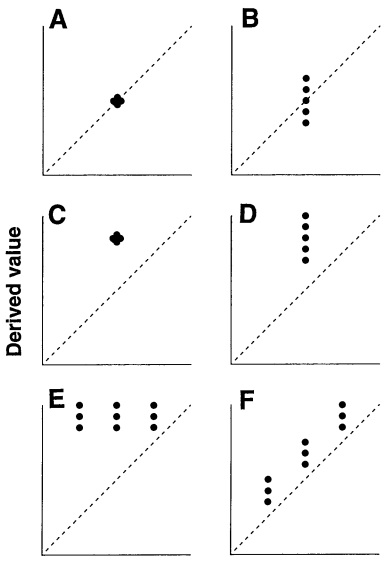

Fig. 1. Schematic representation of effects due to adaptation, chance, and history on evolutionary change and diversification. (**A**) No initial variation and no evolutionary change and hence no effects. (**B**) An effect due to chance only. (**C**) An effect due to adaptation only. (**D**) Effects due to both chance and adaptation. (**E**) An initial effect due to history is eliminated by subsequent effects due to chance and adaptation. (**F**) An initial effect due to history is maintained, with subsequent effects due to chance and adaptation superimposed. See text for further explanation.

M. Travisano, J. A. Mongold, R. E. Lenski, Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824, USA.
A. F. Bennett, Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, USA.
*To whom correspondence should be addressed.

that the variation among sets of derived populations was statistically significant and was similar in magnitude to the initial variation among the several ancestral genotypes (Fig. 1F). Thus, although the value of the trait in a derived population may have changed, it still reflects some contribution from its ancestral state. The effect of history may also be amplified over time; that is, genetic differences having no discernible effect on the initial value of some trait may constrain subsequent evolution so that the effect of ancestry becomes evident only later. The effects of chance, history, and adaptation are not mutually exclusive; all three may simultaneously influence a particular lineage. As we will now show, one can rigorously quantify the contributions of these different influences.

Bacteria have several properties that make them well suited for evolution experiments (3). Their rapid growth allows evolving populations to be tracked for hundreds of generations. They can be frozen indefinitely and then revived, which allows ancestral and derived genotypes to be compared directly, including measurement of their relative fitness in competition. Because bacteria reproduce asexually, one can initiate replicated populations that are identical, consisting of a single genotype; evolutionary change in these populations thus depends entirely on mutations that occur during the course of the experiment. Hence, experiments may encompass the origin, as well as the fate, of genetic variation and phenotypic novelties.

We analyzed the contributions of adaptation, chance, and history in two experiments with *Escherichia coli*. The first examined evolution in a novel nutrient environment, and the second examined evolution in a novel thermal environment. In each experiment, we measured (with independent replication to permit estimation of the pure measurement error) the mean value of two traits in each population. One of these traits was fitness, which is the most important trait in evolutionary theory (4). Mean fitness of a derived population was assayed by allowing it to compete against its ancestor in the same environment used for the experimental evolution; relative fitness is expressed as the ratio of Malthusian parameters (5). We also measured average cell size for each population (6). Size is a morphological characteristic that influences almost all functional processes (7), and the sizes of whole organisms and their component parts are widely used in paleontological (as well as neontological) research, where other traits are difficult or impossible to measure. An increment in fitness might be achieved by many different phenotypic changes. Therefore, we expect that fitness, as a trait, is likely to exhibit parallel and even convergent evolution. In contrast, we expect traits that are only weakly correlated with fitness to be more prone to chance divergence as well as to retention of ancestral differences.

In the first experiment, 36 populations were derived from a common ancestor as follows: A single genotype of *E. coli* strain B was cloned and used to found 12 replicate populations (8), which were serially propagated for 2000 generations at 37°C in glucose-limited medium (5). The 12 derived populations had similar fitnesses to one another when glucose was the carbon source, but they were very heterogeneous when their fitnesses were assayed in a maltose-limited environment (9). One genotype from each of the 12 replicate populations was cloned and used to found 3 new replicate populations. These 36 populations were then propagated for another 1000 generations under the ancestral conditions, except that an equal concentration (m/v) of maltose replaced glucose. Thus, the experiment began with replicated points in genotypic space that subsequently experienced identical environments. We could therefore assess the extent to which the populations' final phenotypic states (after 1000 generations in maltose) depended on their unique history (differences that arose during the preceding 2000 generations in glucose). Moreover, we could quantitatively compare the historical effect with the effects due to chance (divergence among lines with the same initial genotype) and adaptation (systematic changes in mean value irrespective of initial genotype). Thus, we could determine whether evolution followed the pattern in Fig. 1E or Fig. 1F or some other variant.

The mean fitnesses, obtained before and after 1000 generations in maltose, are shown in Fig. 2A for each of the populations (*10*). The pattern is similar to that in Fig. 1E, in which the effect of history is diminished (shown by compression of the variance along the ordinate relative to the abscissa) whereas the effect of adaptation is pronounced (shown by elevation of points above the isocline). We can formalize the contributions of adaptation, chance, and history to fitness by estimating the change in grand mean, which reflects adaptation, and by doing a nested analysis of variance (ANOVA) to estimate variance components corresponding to chance and history (*11*). Figure 2B shows the relative contributions of adaptation, chance, and history to fitness before and after evolution in maltose. By design, the initial effects of adap-

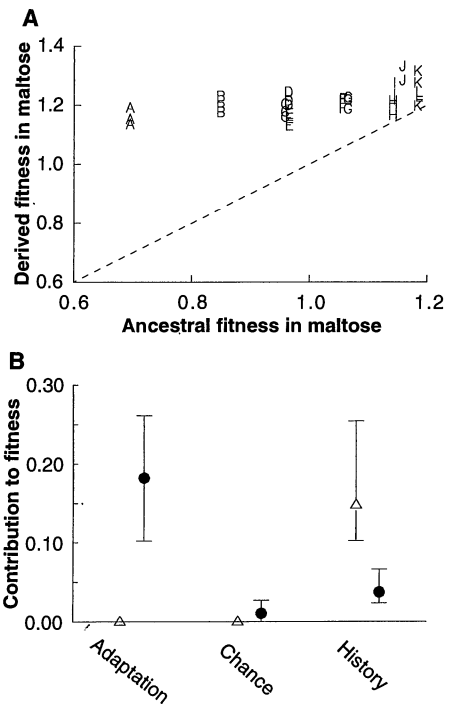

Fig. 2. Evolution of fitness during 1000 generations in maltose. (**A**) Derived versus ancestral values for mean fitness in the 36 experimental populations. Symbols A to L indicate 12 different progenitor genotypes. (**B**) Relative contributions of adaptation, chance, and history to mean fitness before (△) and after (●) 1000 generations in maltose. Error bars represent 95% confidence intervals.

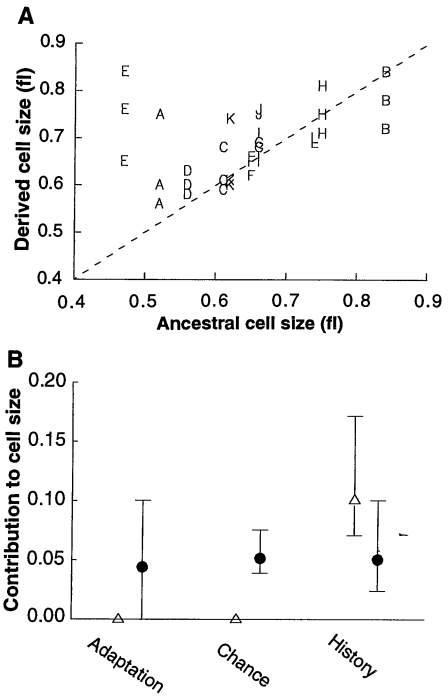

Fig. 3. Evolution of cell size during 1000 generations in maltose. (**A**) Derived versus ancestral values for average cell volume in the 36 experimental populations. Symbols A to L indicate the 12 progenitor genotypes from Fig. 2; fl, femtoliters. (**B**) Relative contributions of adaptation, chance, and history to average cell volume before (△) and after (●) 1000 generations in maltose. Error bars represent 95% confidence intervals.

tation and chance were zero, as there had not yet been any time to adapt or diverge; the differences among starting genotypes reflect the historical effect of independent phylogeny. An ANOVA indicates that the initial contribution of history (starting genotype) was highly significant (*12*). After 1000 generations in maltose, the grand mean fitness of the 36 populations had significantly increased (*13*). The effect of history was still significant (*14*); however, its magnitude was reduced to only ~25% of its initial value, which indicates convergence (Fig. 2B). This remaining historical contribution was much smaller than that of adaptation. Thus, adaptation that occurred during 1000 generations in identical environments largely, but not entirely, eliminated the historical effect on fitness of contingencies during the preceding 2000 generations. Chance divergence was not significant (*15*).

During evolution in maltose, changes in cell size (Fig. 3A) (*16*) were not significantly correlated with changes in fitness, which indicates that size was effectively an independent trait (*17*). Cell size showed no significant overall change (*18*), as shown by the confidence limits for the contribution of adaptation (Fig. 3B). However, chance (shown by divergence among populations with the same starting genotype) contributed significantly to the final distribution of cell size (*19*). The effect of history (starting genotype) was highly significant at both the beginning (*20*) and end (*21*) of the 1000 generations in maltose. In fact, a large fraction of the initial contribution of history to cell size was maintained, and the final historical contribution was at least comparable to that of adaptation (Fig. 3B).

In the second experiment, a single genotype from 1 of the 12 populations propagated for 2000 generations in glucose at 37°C became the common ancestor (and competitor). This genotype was cloned to found 24 populations, 6 of which were propagated in the same glucose medium under each of four thermal regimes (*22*): constant 32°, 37°, and 42°C and daily alternation between 32° and 42°C. After 2000 generations, fitnesses at the temperatures at which each group had evolved were systematically greater than fitnesses at the other experimental temperatures, which indicates temperature-specific genetic adaptation (*22*). One genotype from each of these 24 populations was then used to found a new population, which was propagated in the same medium at 20°C for an additional 1000 generations. Thus, this experiment examined evolution at a novel temperature of populations adapted to an array of other thermal regimes. The beginning and end of the 1000 generations at 20°C provide the initial and final time points of our analysis of changes in fitness and cell size. The effects of adaptation are again defined by changes in mean value for a trait. The effects of chance and history are estimated by ANOVA, but are confounded because genotypes were not replicated at the start of the period at 20°C; we therefore estimate their combined effects in this experiment.

Mean fitnesses before and after 1000 generations at 20°C are shown in Fig. 4A for each of the 24 populations (*23*). The predominant effect on fitness at 20°C was adaptation (Fig. 4B), which went from nonexistent (by design) at the start of the 20°C experiment to highly significant after 1000 generations (*24*). The effect of chance plus history also went from statistically insignificant (*25*) to highly significant (*26*). We asked whether populations whose ancestors had adapted to lower temperatures (32°C and alternating between 32° and 42°C) eventually became more fit at 20°C than did populations whose ancestors had adapted to higher temperatures (37° and 42°C). In fact, this effect of ancestral selection history on fitness was significant (*27*). Even so, the effect of adaptation on fitness was significantly greater than that of chance plus history (*28*).

As in the previous experiment, changes in cell size (Fig. 5A) (*29*) were not significantly correlated with changes in fitness, so that size is considered an independent trait (*30*). The effects of chance plus history were highly significant at both the beginning (*31*) and end (*32*) of 1000 generations at 20°C; the magnitude of these effects was essentially unchanged (Fig. 5B). Moreover, although adaptation had a significant effect on cell size at the experiment's end (*33*), its magnitude was significantly less than that of chance plus history.

We have shown that the contributions of adaptation, chance, and even history to phenotypic evolution can be disentangled and rigorously quantified by appropriately designed experiments. Bacterial populations showed parallel and even convergent evolution in fitness. In contrast, the effects of chance and history were more important for the evolution of cell size, a trait only weakly correlated with fitness. These results are consistent with the view that the footprint of history may be obliterated for traits that are subject to strong selection, whereas the effect of history is preserved in traits that are less important. However, experiments can span only short stretches of time. Over much longer periods, the footprint of history might eventually become too deep to be obscured even by intense selection (*2*).

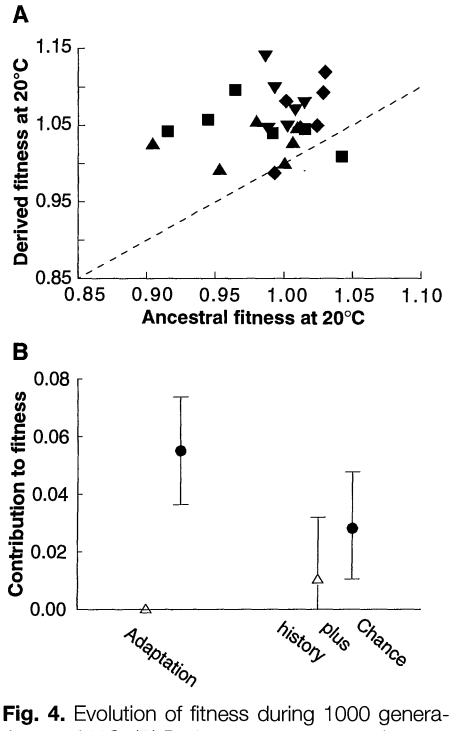

Fig. 4. Evolution of fitness during 1000 generations at 20°C. (**A**) Derived versus ancestral values for mean fitness in the 24 experimental populations. Symbols indicate four different ancestral selection regimes: 32°C (▼), 37°C (■), 42°C (▲), alternating 32° and 42°C (◆). (**B**) Relative contributions of adaptation and chance-plus-history before (△) and after (●) 1000 generations at 20°C. Error bars represent 95% confidence intervals.

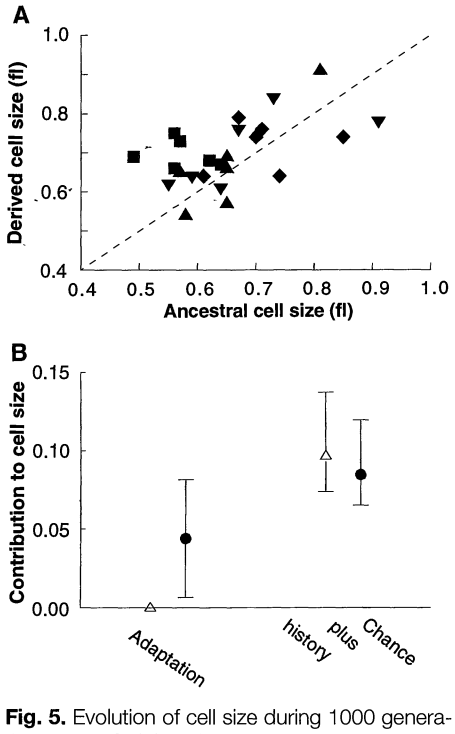

Fig. 5. Evolution of cell size during 1000 generations at 20°C. (**A**) Derived versus ancestral values for average cell volume in the 24 experimental populations. Symbols indicate the four ancestral selection regimes from Fig. 4; fl, femtoliters. (**B**) Relative contributions of adaptation and chance-plus-history before (△) and after (●) 1000 generations at 20°C. Error bars represent 95% confidence intervals.

REFERENCES AND NOTES

1. M. Kimura, *Nature* **217**, 624 (1968); J. L. King and T. H. Jukes, *Science* **164**, 788 (1969); S. J. Gould and R. C. Lewontin, *Proc. R. Soc. London Ser. B* **205**, 581 (1979); G. Oster and P. Alberch, *Evolution* **36**, 444 (1982); M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, 1983); E. Mayr, *Am. Nat.* **121**, 324 (1983); J. Maynard Smith *et al.*, *Q. Rev. Biol.* **60**, 265 (1985); D. W. Stephens and J. R. Krebs, *Foraging Theory* (Princeton Univ. Press, Princeton, NJ, 1986); G. A. Parker and J. Maynard Smith, *Nature* **348**, 27 (1990); G. C. Williams, *Natural Selection: Domain, Levels and Challenges* (Oxford Univ. Press, Oxford, 1992).
2. S. J. Gould, *Wonderful Life: The Burgess Shale and the Nature of History* (Norton, New York, 1989): "Any replay of the tape would lead evolution down a pathway radically different from the road actually taken. . . . Each step proceeds for cause, but no finale can be specified at the start, and none would ever occur a second time in the same way, because any pathway proceeds through thousands of improbable stages. Alter any early event, ever so slightly and without apparent importance at the time, and evolution cascades into a radically different channel."
3. D. E. Dykhuizen and D. L. Hartl, *Microbiol. Rev.* **47**, 150 (1983); R. E. Lenski, in *Encyclopedia of Microbiology*, J. Lederberg, Ed. (Academic Press, San Diego, CA), vol. 2, pp. 125–140.
4. R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford, 1930; reprinted by Dover, New York, 1958); S. Wright, *The Theory of Gene Frequencies*, vol. 2 of *Evolution and the Genetics of Populations* (Univ. of Chicago Press, Chicago, IL, 1969); J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory* (Harper and Row, New York, 1970).
5. R. E. Lenski, M. R. Rose, S. C. Simpson, S. C. Tadler, *Am. Nat.* **138**, 1315 (1991).
6. Cell sizes were obtained with the use of a Coulter particle counter and channelizer, which measures the volume (1 fl = 10^{-12} ml) displaced by a particle [see F. Vasi, M. Travisano, R. E. Lenski, *Am. Nat.* **144**, 432 (1994)].
7. R. H. Peters, *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, 1983); W. A. Calder III, *Size, Function, and Life History* (Harvard Univ. Press, Cambridge, MA, 1984); K. Schmidt-Nielsen, *Scaling: Why is Animal Size So Important?* (Cambridge Univ. Press, Cambridge, 1984).
8. The progenitors were isogenic, except for an arabinose-use marker that enabled us to monitor possible cross-contamination and to score competitors in assays of relative fitness (*5*). Half of the populations were Ara$^-$ and the other half were Ara$^+$. The marker itself had no discernible effect on fitness in the environments used in this study.
9. M. Travisano, F. Vasi, R. E. Lenski, *Evolution*, in press.
10. Fitnesses were measured with fivefold replication for the 12 ancestral genotypes and with 10-fold replication for the 36 derived populations. These assays and all those reported later were run in sets of complete blocks.
11. The variance component due to measurement error is the mean square error. The variance component due to chance indicates variation among populations with the same ancestor that is above and beyond that due to measurement error. The variance component due to history reflects variation among ancestor groups that is above and beyond that due to chance divergence and measurement error. Approximate confidence limits can be placed on these variance components, and they are asymmetric (*34*). We report the square roots of the variance components for chance and history, so that magnitudes and units are directly comparable to the mean change due to adaptation. The t distribution with $n - 1$ df was used to obtain confidence limits for the change in grand mean, where $n = 12$ is the number of independent sets of populations.
12. $F = 24.3$, 11 and 44 df, $P < 0.001$.
13. $t = 5.06$, 11 df, $P < 0.001$.
14. $F = 7.31$, 11 and 24 df, $P < 0.001$.
15. $F = 1.20$, 24 and 314 df, $P = 0.240$.
16. Cell sizes were measured with fivefold replication for both the 12 ancestral genotypes and the 36 derived populations.
17. $r^2 = 0.042$, $n = 12$, $P = 0.525$.
18. $t = 1.71$, 11 df, $P = 0.114$ (*35*).
19. $F = 10.2$, 24 and 139 df, $P < 0.001$; one value was missing.
20. $F = 99.7$, 11 and 44 df, $P < 0.001$.
21. $F = 3.59$, 11 and 24 df, $P = 0.004$.
22. A. F. Bennett, R. E. Lenski, J. E. Mittler, *Evolution* **46**, 16 (1992). Again, half the populations were Ara$^-$ and the other half were Ara$^+$ (*8*).
23. Fitnesses were measured with sevenfold replication for both the 24 ancestral genotypes and the 24 derived populations.
24. $t = 6.10$, 23 df, $P < 0.001$.
25. $F = 1.13$, 23 and 138 df, $P = 0.325$.
26. $F = 2.09$, 23 and 138 df, $P = 0.005$. However, the confidence interval for the initial variance component includes the point estimate of the final variance component. Hence, we cannot confidently claim that the effect of chance-plus-history was amplified during 1000 generations in identical environments, nor can we exclude this possibility.
27. $t = 2.57$, 22 df, $P = 0.017$.
28. Although the confidence intervals for chance-plus-history and for adaptation overlap, neither includes the point estimate of the other. Hence, the difference is significant.
29. Cell sizes were measured with fourfold replication for both the 24 ancestral genotypes and the 24 derived populations.
30. $r^2 = 0.092$, $n = 24$, $P = 0.149$.
31. $F = 19.4$, 23 and 69 df, $P < 0.001$.
32. $F = 32.0$, 23 and 69 df, $P < 0.001$.
33. $t = 2.43$, 23 df, $P = 0.023$ (*35*).
34. R. R. Sokal and F. J. Rohlf, *Biometry* (Freeman, San Francisco, CA, ed. 2, 1981), pp. 217–218.
35. The effect due to adaptation is based on the change in grand mean and thus reflects the directional component of selection on cell size. Stabilizing selection on cell size might be suggested by a reduction in the combined effects of chance and history, but this was not observed.

36. R.E.L. dedicates this paper to the memory of his mother, Jean Lenski. We thank A. Cullum, L. Forney, R. Hudson, S. Kalisz, M. Rose, D. Straney, J. Tiedje, and S. Tonsor for discussions; A. Inouye and B. Korona for assistance in the lab; and two reviewers for comments. Supported by NSF grants DEB-9249916 (to R.E.L.) and IBN-9208662 (to A.F.B. and R.E.L.) and by the NSF Center for Microbial Ecology (BIR-9120006).

1 July 1994; accepted 3 October 1994

A Proficient Enzyme

Anna Radzicka and Richard Wolfenden

Orotic acid is decarboxylated with a half-time ($t_{1/2}$) of 78 million years in neutral aqueous solution at room temperature, as indicated by reactions in quartz tubes at elevated temperatures. Spontaneous hydrolysis of phosphodiester bonds, such as those present in the backbone of DNA, proceeds even more slowly at high temperatures, but the heat of activation is less positive, so that dimethyl phosphate is hydrolyzed with a $t_{1/2}$ of 130,000 years in neutral solution at room temperature. These values extend the known range of spontaneous rate constants for reactions that are also susceptible to catalysis by enzymes to more than 14 orders of magnitude. Values of the second-order rate constant k_{cat}/K_m for the corresponding enzyme reactions are confined to a range of only 600-fold, in contrast. Orotidine 5′-phosphate decarboxylase, an extremely proficient enzyme, enhances the rate of reaction by a factor of 10^{17} and is estimated to bind the altered substrate in the transition state with a dissociation constant of less than 5×10^{-24} M.

Many enzymes act on substrates at rates that approach the rate of encounter between enzyme and substrate in solution (*1*). Biological reactions vary greatly, however, in their spontaneous rates and in the consequent burden that they place on an efficient catalyst. For example, carbon dioxide is hydrated nonenzymatically within several seconds in neutral aqueous solution at room temperature (*2*); peptide bonds undergo cis/trans isomerization within a few minutes (*3*); chorismic acid is converted to prephenic acid over a period of several hours (*4*); and the hydrolytic deamination of cytidine or adenosine approaches completion only after several centuries (*5*). Observations of reactions in sealed quartz tubes at elevated temperatures, described below, indicate that under physiological conditions in the absence of enzymes, the decarboxylation of orotic acid and the hydrolysis of phosphodiesters proceed on a time scale that dwarfs the life-span of living organisms. Enzymes catalyzing these reactions are correspondingly proficient at lowering the activation barrier for reaction and exhibit a high degree of discrimination between the substrates in the ground state and in the altered form that is present in the transition state, binding the latter species very tightly. Enzymes of this type are expected to be especially sensitive to inhibition by stable analogs of the altered substrate in the transition state.

1,3-Dimethylorotic acid has been shown to undergo decarboxylation slowly in sulfolane at 206°C (*6*). In order to compare this reaction with the enzymatic decarboxylation of orotidine 5′-phosphate (OMP) by

Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599, USA.