Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 5882–5886, July 1991
Evolution

Adaptive evolution that requires multiple spontaneous mutations: Mutations involving base substitutions

BARRY G. HALL

Biology Department, Hutchison Hall, University of Rochester, Rochester, NY 14627

Communicated by Bernard D. Davis, March 22, 1991 (received for review January 10, 1991)

ABSTRACT **A previous study has demonstrated that adaptive missense mutations occur in the *trp* operon of *Escherichia coli*. In this study it is shown that, under conditions of intense selection, a strain carrying missense mutations in both *trpA* and *trpB* reverts to Trp$^+$ 10^8 times more frequently than would be expected if the two mutations were the result of independent events. Comparison of the single mutation rates with the double mutation rate and information obtained by sequencing DNA from double revertants show that neither our classical understanding of spontaneous mutation processes nor extant models for adaptive mutations can account for all of the observations. Despite a current lack of mechanistic understanding, it is clear that adaptive mutations can permit advantageous phenotypes that require multiple mutations to arise and that they appear enormously more frequently than would be expected.**

Adaptive evolution involves an increase in the frequency of fitter phenotypes over time, through interactions between the organism and the environment that are collectively called "selection." (It is important to distinguish adaptive evolution from the evolution of total variation, much of which is neutral and thus depends strictly upon chance events.) It has been assumed that the generation of adaptive, or fitter, variants by mutation is a process that is entirely independent of selection (i.e., that mutations arise entirely without respect to their effects on fitness). The universality of that view was challenged by Cairns *et al.* (1), who suggested that bacterial cells might have mechanisms by which organisms could create or selectively retain mutations that are needed to meet a specific environmental challenge. They called such mutations "directed mutations." This term has aroused considerable controversy (2), and "adaptive mutations" might be preferable.

Cairns' challenge, if not his specific view of environmental instruction of the DNA, was supported by my study of mutations that were mediated by a mobile DNA element (3). More recently I showed (4) that base substitution mutations also occur more often when they are advantageous than when they are neutral. When Trp$^-$ mutants are allowed to form colonies on medium with a limited supply of tryptophan, the cells exhaust the tryptophan, no further cell divisions occur, and the number of viable cells per colony decreases exponentially with a first-order rate constant of -0.24 per day. Despite the absence of cell division, Trp$^+$ revertants continued to accumulate as papillae on the colonies for at least 10 days. By that time more than 90% of the revertants that were present were the result of mutations that would not have occurred in growing cultures under conditions where the mutations were not advantageous. This process of generating mutations in aging colonies was shown to be specific to the environmental challenge encountered by the cells: when a *trp*$^-$ *cys*$^-$ double mutant was starved for tryptophan it reverted only to Trp$^+$ and when starved for cysteine it reverted only to Cys$^+$.

To explain such environmental effects, Stahl (5) suggested that after prolonged starvation cells may repair mismatched bases very slowly, resulting in unusually long persistence of mismatch substitutions. When such substitutions were in the coding strand and solved the current problem of the cell, expression of the gene would allow growth, DNA replication would fix the mismatch into the daughter strand, and a successful mutant would be detected. If a mutation failed to solve the cell's current problem, no replication would ensue and the slowly acting mismatch repair system would eventually remove the substitution so that when old colonies are later tested no increase in the mutation rate at that site will be observed. As an alternative, I suggested that as colonies age some small fraction of the cells at any instant may enter a hypermutable state where mutations might occur at many sites in the genome (4). If one of those mutations solved the cell's current problem then the cell would grow, causing it to leave the hypermutable state and be recovered as a successful mutant. If no mutation solved the cell's problem, then within a short while the cell would die. As a result, when cells in old colonies are tested for mutations at sites that were not under selection, such mutations are not recovered because the only cells being tested are those that never entered the hypermutable state. This model predicts the recovery of mutations at sites that were not under selection among the successful selected mutants, whereas it is not clear that the Stahl slow-repair model makes such a prediction. Because 2% of the Trp$^+$ revertants carried additional mutations leading to unidentified auxotrophies, I expressed a preference for the hypermutable state model over the slow mismatch repair model.

Davis (6) proposed an entirely different class of model, based on the hypothesis that during prolonged starvation transcription might be mutagenic. That model is very attractive because it would permit the environment to target genes for mutation by regulation of transcription and because transcription requires temporary unwinding to yield mutationally vulnerable single-stranded DNA. Several of the genes studied up to now have been maximally expressed under selective conditions, and there has not yet been a strong test of the Davis hypothesis.

The data are not really sufficient to make a serious distinction among the models, and additional detailed biochemical information is required to determine which model, if any, has any validity.

The prediction that multiple mutations could be recovered from cells exposed to prolonged intense selection led me to wonder whether it might be possible, under similar conditions, to recover mutants in which two mutations were required to produce the advantageous phenotype. This is a specific case of a very general problem of molecular adaptive evolution: How is an advantageous phenotype selected when it requires multiple mutations, none of which are advantageous until all are present (i.e., only the last mutation to occur is actually selected)? If cells had a means of specifically

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

increasing the rate of advantageous multiple mutations, they might be able to circumvent a barrier that would appear to be difficult when two independent random mutations are required to improve fitness and insuperable when more than two are required. Here I present evidence that *Escherichia coli* cells do, indeed, possess such a mechanism.

MATERIALS AND METHODS

Strains. All bacterial strains were *Escherichia coli* K-2. W3110 is a wild-type F^- strain. Strains FCY2B (*trpA46*), FCY21B (*trpB9578*), and FCY6B (*trpA46 trpB9578*) are isogenic derivatives of the wild-type F^- strain W3110 (4) and carry *tna*::Tn*10* and Δ(*bgl–pho*).

Media and Growth Conditions. Minimal medium consisted of 423 mg of sodium citrate, 100 mg of $MgSO_4 \cdot 7H_2O$, 1 g of $(NH_4)_2SO_4$, 540 μg of $FeCl_3$, 1 mg of thiamine, 3 g of KH_2PO_4, 7 g of K_2HPO_4, and 2 g of carbon source per liter.

TAD medium contained, in addition to the above ingredients, 5 g of vitamin-free Casamino acids (Difco), 20 mg of adenosine, 30 mg of cytosine, 20 mg of guanosine, 20 mg of uridine, 30 mg of thymidine, 20 mg of $MnSO_4$, 200 μg of pantothenic acid, 400 μg of pyridoxine, 400 μg of riboflavin, 200 μg of *p*-aminobenzoic acid, 200 μg of niacin, and 1 μg of biotin per liter. Plates for selecting *trp*$^+$ revertants were incubated at 30°C in a humidified chamber. For all other purposes plates were incubated and liquid cultures were shaken at 37°C. For selection of Trp^+ revertants 5 μM tryptophan was added to TAD plates.

The rich complete medium was LB (Luria broth) (7).

Growth Rates. Growth rates were measured by following the increase in A_{600} of 50-ml cultures in TAD medium in 250-ml flasks shaken at 37°C. Growth rates (hr^{-1}) are expressed as the first-order growth rate constant as determined from a least squares fit of $\ln(A_{600})$ vs. time. Each rate is the mean of three determinations. Death rates (day^{-1}) were determined from viable counts.

PCR Amplification and DNA Sequencing. *trpB* DNA was amplified using primers corresponding to base pairs (bp) 5551–5580 and the complement of bp 6121–6150 of the *trp* operon (GenBank: accession name ECOTGP and accession numbers J01714, M12471, and M12472) as described (8). *trpA* DNA was amplified using primers corresponding to bp 5921–5950 and the complement of bp 6886–6915 of the *trp* operon. The amplified DNA was purified of oligonucleotides and proteins by using the Geneclean II kit (Bio 101, La Jolla, CA) according to manufacturer's instructions. The purified amplified DNA was sequenced directly using the T7 sequencing kit (Pharmacia) according to manufacturer's instructions, except that the primers were hybridized to template DNA by boiling the template–primer mixture, quick freezing in dry ice for 5 min, and allowing the mixture to warm slowly to room temperature. The same oligonucleotides used for amplification were used for sequencing, and in addition a primer corresponding to the complement of bp 6821–6850 was used to sequence *trpA*. An average of about 175 bases were read from a sequencing reaction mixture.

RESULTS

Strains FCY2B (*trpA46*), FCY21B (*trpB9578*), and FCY6B (*trpAB*) were plated onto TAD/5 μM tryptophan plates at a density of ≈100 cells per plate. Under these conditions *trp*$^-$ colonies exhaust the tryptophan and reach a maximum number of viable cells in 3 days, and Trp^+ revertants appear as papillae on those colonies within 24 hr of the occurrence of the mutations (4). In each experiment the number of viable cells per colony was monitored daily by resuspending individual colonies, diluting, and plating onto rich medium. As reported (4), cells in these nongrowing colonies die exponentially. Colonies of the *trpA46* and of the *trpB9578* strains produced revertant papillae at an approximately constant rate for at least 4 weeks after the tryptophan was exhausted (Fig. 1*A*), whereas the *trpA trpB* double mutant reverted very slowly. An expansion of the *y*-axis scale (Fig. 1*B*) shows that one double-revertant papilla appeared on one of the 9254 colonies on day 12, and then others appeared from day 20 onward. As each revertant papilla appeared cells were streaked out and the resulting Trp^+ isolate was stored at −80°C as a permanent stock.

To calculate the rate at which Trp^+ double revertants from the *trpAB* strain are expected, it is necessary to know the rates at which the individual *trp* mutations revert. In nongrowing colonies the mutation rate is time-dependent and increases as the colonies age (4). The mutation rate each day was calculated as the number of revertants divided by the number of viable cells on the previous day (Fig. 2). The expected double mutation rate each day (i.e., the product of the *trpA* and *trpB* mutation rates) increased from 5×10^{-22} per cell at the beginning of the experiment to 4×10^{-19} per cell at the end. The observed double-mutation rates increased from 5.5×10^{-12} at day 20 to 9.9×10^{-11} by day 30. Double mutations thus occurred more than eight orders of magnitude more frequently than would be expected if the two mutations were independent events. At the end of the experiment, in which 9254 FCY6B colonies had been incubated for 1 month, a total of 37 independent double revertants had been isolated.

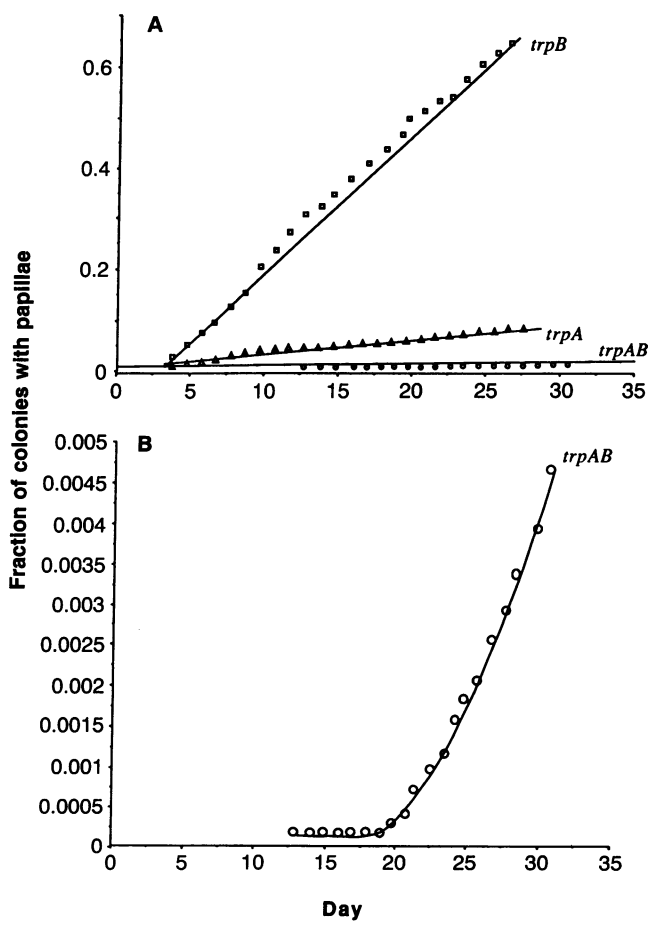

FIG. 1. Accumulation of Trp^+ revertants as papillae on colonies during tryptophan starvation. (*A*) *trpB* is strain FCY21B, *trpA* is strain FCY2B, and *trpAB* is strain FCY6B. (*B*) *trpAB*, same data as in *A* except that *y*-axis scale has been expanded 160 times. At the beginning of the experiment there were 2010 *trpA* colonies, 2630 *trpB* colonies, and 8980 *trpAB* colonies. In each case about 10% of the colonies had been used for estimation of cell death rate by the end of the experiment.

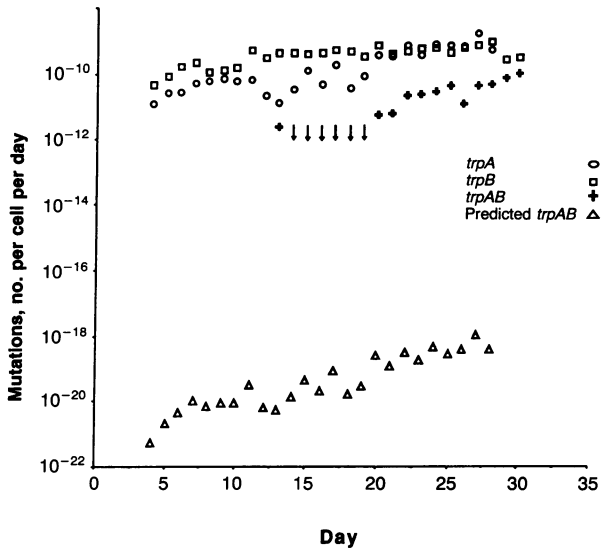

FIG. 2. Mutation rates in tryptophan-starved colonies. The rates are given as the number of Trp^+ revertants that appeared on a given day divided by the number of viable cells on the previous day. The predicted *trpAB* values are the product of the *trpA* and *trpB* values on that day. Arrows indicate that the observed rate was $<5 \times 10^{-12}$ on that day.

The calculation of the expected double mutation rate in Fig. 2 is based on the assumption that the mutations are independent but simultaneous. An alternative is that the mutations might occur independently, but sequentially. Since neither mutation alone permits growth and since each single revertant died exponentially, the number of double revertants that would be expected by sequential mutations is easily calculated. The cumulative number of *trpA* and *trpB* revertants per cell on each day was calculated from the data in Fig. 1*A* as revertant cells (R) per colony according to the Poisson distribution where $R = -\ln(P_0)$, and P_0 is the proportion of colonies with zero papillae on that day. Those values were multiplied by the number of FCY6B colonies and the number of cells per colony each day to estimate the number of *trpA* and *trpB* revertants in the FCY6B experiment each day. To estimate the number of double revertants that would be produced by successive mutation each day, the number of $trpA^+$ revertants was multiplied by the $trpB \rightarrow trpB^+$ mutation rate on that day as estimated in Fig. 2, and similarly the number of $trpB^+$ revertants was multiplied by the $trpA \rightarrow trpA^+$ mutation rate on that day; the two numbers were added. Fig. 3 shows that the cumulative number of Trp^+ revertants that could be accounted for by successive mutations is more than 10^5 times lower than the number of Trp^+ double revertants that was observed.

Properties of the Double Revertants. All revertants were tetracycline resistant, showing that they carried the parental Tn*10* and that they were unlikely to have been contaminants. The growth rates of all 37 FCY6B *trpAB* double revertants were determined on the rich TAD medium (without tryptophan) (Fig. 4) and on glucose minimal medium (data not shown). All revertants were able to grow on glucose minimal medium and thus did not carry additional auxotrophic mutations, as had previously been observed for some revertants of FCY2B and FCY21B (4). The revertants fell into three distinct growth rate classes. Class I strains grew at 30% of the wild-type rate, class II strains grew at 60% of that rate, and class III strains grew at the wild-type rate. Clearly, not all of the reversions involved mutations only to the wild-type bases at both sites.

Sequencing *trp* DNA. The regions in *trpA* and *trpB* that were expected to contain the *trpA46* and *trpB9578* mutations were amplified by PCR and sequenced from wild-type strain W3110, the double-mutant parental strain FCY6B, and 11 *trpAB* revertants. These included four class I, three class II, and four class III revertants.

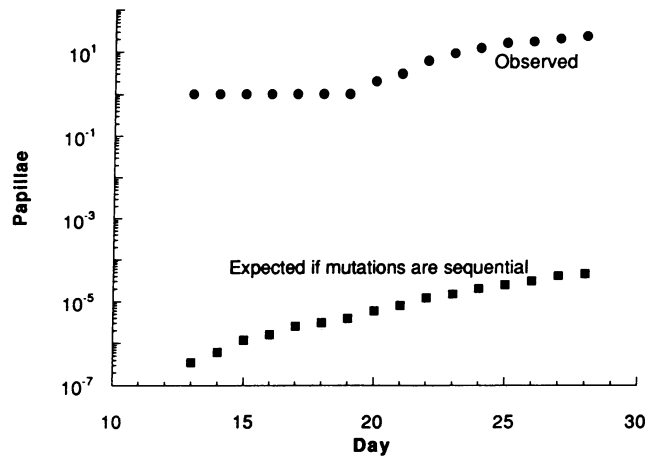

FIG. 3. Accumulation of Trp^+ revertants from strain FCY6B (*trpAB*) if mutations were sequential. Circles show the observed number of Trp^+ papillae from strain FCY6B in the experiment shown in Fig. 1. Squares show the number of papillae that would have accumulated in that experiment if the $trpA^+$ and $trpB^+$ mutations were independent and sequential.

In the *trpB* gene all 11 revertants had the wild-type guanosine instead of the *trpB9578*-specified adenosine at bp 5717, converting the mutant glutamic acid back to the wild-type glycine.

In the *trpA* gene, three of the class III revertants had the wild-type guanosine instead of the *trpA46*-specified adenosine, at bp 6634, again converting the mutant glutamic acid back to the wild-type glycine. The remaining class III mutant had a cytidine at that position, resulting in alanine instead of the wild-type glycine. The substitution of alanine for glycine apparently has too little effect on activity to affect the growth rate. The class II mutants all had a thymidine at that position, resulting in a valine instead of the wild-type glycine. The class I strains all retain the *trpA46*-specified adenosine at that position, thus implying either a second-site mutation elsewhere within *trpA* or an extragenic missense suppressor. It is therefore not surprising that their growth rate is reduced to 30% of wild type.

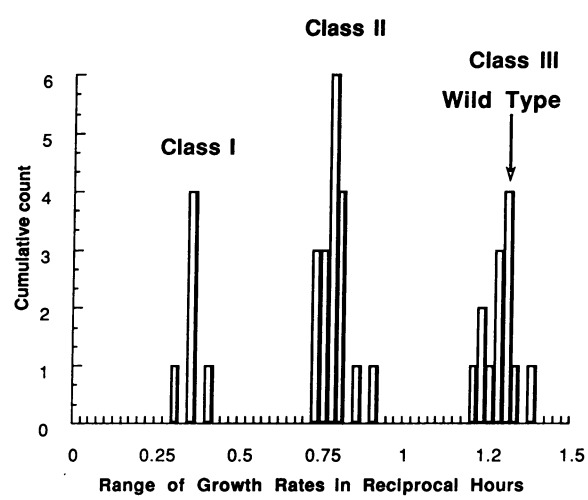

FIG. 4. Growth rates of Trp^+ revertants of strain FCY6B (*trpAB*) on TAD medium. Arrow indicates the growth rate of the wild-type strain W3110. The *y* axis is the number of revertants whose growth rate was in the interval indicated on the *x* axis.

The sequencing results confirm that the revertants are not the result of contamination or, in most strains, of simple missense suppressor mutations.

Tests for Mutator Alleles. When double mutations are selected, the resulting strains often turn out to carry mutator mutations (J. W. Drake, personal communication). It was possible that mutator mutations occurred during the growth of a few colonies so that in those colonies a few cells would have very high spontaneous mutation rates and thus give rise to double revertants at unexpectedly high rates. All 37 of the FCY6B double revertants were, therefore, tested for the presence of mutator alleles, both those that might act in growing cells and those that might increase adaptive mutations in stationary cells.

To test for mutator alleles that affect growing cells, entire duplicate 1-ml cultures containing 4.5×10^8 cells were plated onto LB medium containing the antibiotic rifampicin (200 μg/ml). The 37 mutants produced 1.7 ± 3.8 colonies per culture (mean ± SD), while the FCY6B parent control produced 1.5 ± 2.3 ($n = 4$) colonies per plate. Evidently none of these revertants carry mutator alleles that affected growing cells.

To test for mutator alleles that might influence specifically the rate of adaptive mutations (e.g., advantageous mutations in stationary cells in old colonies), six of the revertant strains whose DNA was sequenced, plus the parent strain FCY6B, were plated onto MacConkey cellobiose medium. Cellobiose-utilizing mutants can arise as a consequence of spontaneous base substitution mutations in *celD* that make the *cel* repressor sensitive to cellobiose as an inducer (9). The base substitutions in *celD* are A → G and A → C, precisely the substitutions that were responsible for the reversions in *trpAB*. Wild-type *E. coli* are unable to utilize cellobiose, but they grow on other resources in MacConkey plates. When colonies have exhausted those alternative resources, there is strong selection for cellobiose utilization and Cel^+ papillae appear on the colonies, a situation directly comparable to selection for Trp^+ mutants on TAD plates with limiting tryptophan. After 24 days of incubation, no difference between the double revertants and the parent strain FCY6B, in terms of the number or timing of appearance of Cel^+ papillae, could be detected. The double revertants, therefore, do not appear to carry mutator mutations that specifically increase the rate of base substitutions under conditions of prolonged intense selection.

The high rate of double mutations is, therefore, very unlikely to depend on a prior mutation that alters the rate of *trp* reversions, either during growth or on medium lacking tryptophan.

DISCUSSION

How can one account for double revertants arising at rates 10^8 higher than would be expected on the basis of independent mutations? The data rule out mutator mutations or missense suppressors as trivial explanations. The double revertants also cannot be explained on the basis of sequential mutations.

Both the hypermutable state model and the Stahl slow-repair model (see above and refs. 4 and 5) seem to offer attractive explanations for the double revertants. The hypermutable state model, in particular, predicts multiple mutations arising nonindependently from a common cause, entry into the hypermutable state. Nevertheless, the data are not consistent with those simple models.

If the probability per day of mutating is the product of entering into the hypermutable state and then having the specific required mutation while in that state, then the observed mutation rate μ_A, for *trpA* → trp^+, is *HA* where *H* is the probability of entering the hypermutable state and *A* is the conditional probability of reverting the *trpA* mutation while in that state. Similarly μ_B, the mutation rate for *trpA* → trp^+, is *HB*, and μ_{AB}, the mutation rate for *trpAB* → trp^+, is *HAB*. Given the three mutation rates μ_A, μ_B, and μ_{AB}, then *H*, *A*, and *B* are obtained by solving the three simultaneous equations. Using the observed values of μ_A, μ_B, and μ_{AB} on day 20, $H = 4.5 \times 10^{-8}$ per cell, $A = 8 \times 10^{-3}$ per cell in the hypermutable state, and $B = 1.6 \times 10^{-2}$ per cell in the hypermutable state. The mean values for days 20–28 are $H = 1.9 \times 10^{-8}$, $A = 0.044$, and $B = 0.042$. Since *A* and *B* are, respectively, the probabilities of the two specific base substitutions, if those values apply to the entire genome this implies that while a cell is in the hypermutable state each base has a probability of about 0.04 of experiencing a base substitution. This is obviously implausible when applied to the entire genome, since that base substitution rate would certainly be lethal. The cell might, however, be able to tolerate that substitution rate if it applied only to a small local region such as might result from an error-prone repair tract in the region of the *trp* operon. That would imply that a high substitution rate ought to be observed in the region surrounding the reversion mutations in the double revertants. Sequencing a little over 700 bp in *trpA* and *trpB* in each FCY6B double revertant, however, detected no substitutions other than those responsible for the reversions themselves. Those 7700 sequenced bases (11 strains × 700 bases per strain) included the equivalent of 1958 fourfold redundant silent sites, thus the probability that the substitution rate was ≥0.04 is 2×10^{-35}, and the probability that the substitution rate was even 0.005 is $<10^{-4}$.

The above analysis applies to all models that account for adaptive mutations on the basis of some triggering event or state followed by random mutations. Thus, none of the simple models proposed to date, including the Davis model (6) in which mutagenic transcription is the common causal event, are consistent with the observations presented here.

The simple models also fail to account for beginning to produce double mutants continuously only on and after day 20. There are no dramatic differences in the individual *trpA* and *trpB* reversion rates prior to day 20 or after that time (Fig. 2). Why do double revertants appear only after about 3 weeks of incubation? [This same delay has been observed in additional experiments with the same double mutant (data not shown).]

The answer, then, is that present knowledge cannot account for double revertants occurring 10^8 times more often than expected. The double revertants are certainly not the result of independent events, but neither the Hall "hypermutable state" model nor the Stahl "slow repair" model accounts for all of the data. Although an explanation for the phenomenon remains obscure, its reality is clear. The potential for adaptive mutations speeding up adaptive microevolution is clear. Because adaptive mutations can permit phenotypes that require multiple mutations to arise enormously more frequently than would be expected, cells have a means of producing new *advantageous* phenotypes that would be unlikely to arise at all by unrelated random mutations.

I am grateful to L. Xu, S. Bayoumy, and D. Doty for skilled technical assistance, to M. Balas for contributions to the early stage of this project, to C. Yanofsky for providing *trp* mutant strains and helpful suggestions, to I. Crawford for helpful information on *trpB*, and to W. Gilbert, F. Stahl, J. Cairns, H. Echols, and an anonymous reviewer for very helpful suggestions. This work was supported by National Science Foundation Grant DMB 8903311.

1. Cairns, J., Overbaugh, J. & Miller, S. (1988) *Nature (London)* **335**, 142–145.
2. Charlesworth, D., Charlesworth, B., Bull, J. J., Graffen, A.,

Holliday, R., Rosenberger, R. F., Valen, L. M. V., Danchin, A., Tessman, I. & Cairns, J. (1988) *Nature (London)* **336,** 525–528.
3. Hall, B. G. (1988) *Genetics* **120,** 887–897.
4. Hall, B. G. (1990) *Genetics* **126,** 5–16.
5. Stahl, F. W. (1988) *Nature (London)* **335,** 112–113.
6. Davis, B. D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5005–5009.
7. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
8. Hall, B. G., Betts, P. W. & Wootton, J. C. (1989) *Genetics* **123,** 635–648.
9. Parker, L. L. & Hall, B. G. (1990) *Genetics* **124,** 473–482.