Long-Term Experimental Evolution in *Escherichia coli*. IV. Targets of Selection and the Specificity of Adaptation

Michael Travisano and Richard E. Lenski

Center for Microbial Ecology, Michigan State University, East Lansing, Michigan 48824-1325

Manuscript received February 14, 1995

Accepted for publication February 2, 1996

ABSTRACT

This study investigates the physiological manifestation of adaptive evolutionary change in 12 replicate populations of *Escherichia coli* that were propagated for 2000 generations in a glucose-limited environment. Representative genotypes from each population were assayed for fitness relative to their common ancestor in the experimental glucose environment and in 11 novel single-nutrient environments. After 2000 generations, the 12 derived genotypes had diverged into at least six distinct phenotypic classes. The nutrients were classified into four groups based upon their uptake physiology. All 12 derived genotypes improved in fitness by similar amounts in the glucose environment, and this pattern of parallel fitness gains was also seen in those novel environments where the limiting nutrient shared uptake mechanisms with glucose. Fitness showed little or no consistent improvement, but much greater genetic variation, in novel environments where the limiting nutrient differed from glucose in its uptake mechanisms. This pattern of fitness variation in the novel nutrient environments suggests that the independently derived genotypes adapted to the glucose environment by similar, but not identical, changes in the physiological mechanisms for moving glucose across both the inner and outer membranes.

There is an important distinction between the causes of natural selection and its effects (Dykhuizen and Dean 1990; Dykhuizen 1995). The mathematical theory of population genetics deals with the effects of natural selection; it allows quantitative predictions about the dynamics of gene frequencies, given that selection coefficients and other relevant parameters are known. However, this theory says nothing about the ecological factors and phenotypic variation that cause systematic differences in survival and reproductive success among genotypes. Thus, an explanation of the causes of natural selection requires that one identify and characterize the ecological agents and phenotypic differences that together are responsible for systematic changes in gene frequency.

In this paper, we seek to identify the phenotypic targets of natural selection in evolving populations of the bacterium *Escherichia coli*, which we have propagated for thousands of generations in a defined laboratory environment (Lenski et al. 1991; Vasi et al. 1994; Travisano et al. 1995a). The wealth of information on the physiology, biochemistry, and genetics of this organism enables us to identify environmental manipulations that can test the relevance of putative targets of selection. That is, using organisms that have adapted genetically to a particular selective environment, we ask whether (and to what extent) a derived genotype's improved fitness carries over to novel environments for which we have *a priori* information concerning their functional similarity to the selective environment.

We also ask whether independently derived genotypes show similar or dissimilar responses to the novel environments. In novel environments that are functionally similar to the historical selective environment, we expect the derived genotypes to have uniformly high fitnesses, as natural selection will have favored systematic changes in those phenotypic traits that are important for success in both the novel and selective environments. By contrast, in novel environments that are functionally dissimilar to the selective environment, the derived genotypes should have lower fitness, on average, and be more variable in their responses. That is, those phenotypic traits that are important for success in a certain novel environment, but were unimportant in the historical selective environment, may have decayed either by random genetic drift or as a pleiotropic side-effect of mutations that were beneficial in the selective environment.

Lenski et al. (1991) began a long-term selection experiment to examine the roles of natural selection and chance events in promoting adaptation and divergence. Twelve initially isogenic populations of *E. coli* B were propagated for 2000 generations in a glucose-limited medium. [Subsequent work has extended this experiment to 10,000 generations (Lenski and Travisano 1994) and to new environments (Bennett et al. 1992; Travisano et al. 1995b; Mongold et al. 1996). In this paper, we use only those genotypes obtained at generation 2000 in the original experimental environment.] Fitness, as measured in the selective environment and
relative to the common ancestor, increased by ~35%, on average, with little variation in fitness among the replicate populations (Lenski et al. 1991). The demographic changes (in maximum growth rate, lag time, etc.) that gave rise to the higher fitnesses of the derived genotypes also exhibited a high degree of parallelism (Vasi et al. 1994). However, measurements of fitness in environments where maltose or lactose was substituted for glucose indicated that derived genotypes from the 12 populations had diverged into at least four distinct groups (Travisano et al. 1995a). Genetic variation for fitness among the derived genotypes in media containing either maltose or lactose was much greater than it was in medium containing glucose. Circumstantial evidence suggested that heterogeneous pleiotropic effects associated with different mutations beneficial in glucose, rather than distinct substitutions due to genetic drift, were primarily responsible for the fitness variation in the novel maltose and lactose environments. Moreover, as glucose and maltose require the same enzymes for catabolism but have quite different mechanisms for uptake from the medium, the much greater fitness variation in maltose suggested that glucose uptake had been an important target for selection during the 2000 generations of experimental evolution (Travisano et al. 1995a).

The specific aims of the present study were threefold: (1) to characterize further the derived genotypes in terms of the parallelism vs. divergence of their independent adaptations to the selective environment, (2) to use this information to identify targets of selection with greater precision, and (3) to explore the relationship between bacterial physiology and fitness in both selective and novel environments. To these ends, we present further experiments in which genotypes from the 12 populations had diverged into at least four distinct groups (TRAVISANO et al. 1995a). Genetic variation for fitness among the derived genotypes in media containing either maltose or lactose was much greater than it was in medium containing glucose. Circumstantial evidence suggested that heterogeneous pleiotropic effects associated with different mutations beneficial in glucose, rather than distinct substitutions due to genetic drift, were primarily responsible for the fitness variation in the novel maltose and lactose environments. Moreover, as glucose and maltose require the same enzymes for catabolism but have quite different mechanisms for uptake from the medium, the much greater fitness variation in maltose suggested that glucose uptake had been an important target for selection during the 2000 generations of experimental evolution (TRAVISANO et al. 1995a).

The specific aims of the present study were threefold: (1) to characterize further the derived genotypes in terms of the parallelism vs. divergence of their independent adaptations to the selective environment, (2) to use this information to identify targets of selection with greater precision, and (3) to explore the relationship between bacterial physiology and fitness in both selective and novel environments. To these ends, we present further experiments in which genotypes from the 12 glucose-adapted populations of E. coli B (LENSKI et al. 1991) were allowed to compete against their common ancestor in nine additional media containing single limiting nutrients. With the previous three media, the fitnesses of the independently derived genotypes have now been examined in 12 different nutrient environments.

As shown in Figure 1, the 12 nutrients fall into four groups based on the physiological mechanisms of their transport across the Gram-negative bacterial outer and inner membranes (Lin 1987). Fructose, glucitol, mannitol, mannosamine, and N-acetylglucosamine (NAG) are OmpF/PTS nutrients (Figure 2), like glucose, they pass through the outer membrane via the porin OmpF (E. coli B lacks the alternative porin OmpC) and then cross the inner membrane via the phosphotransferase system (PTS). Galactose, glycerol, lactose and melibiose are OmpF/non-PTS nutrients; they pass through the outer membrane via OmpF, but their transport across the inner membrane is performed by other nutrient-specific proteins that are not part of PTS (Figure 3). Conversely, trehalose is a LamB/PTS nutrient, it passes through the outer membrane primarily via the larger diameter porin LamB (Klein and Boos 1993), but its transport across the inner membrane is mediated by the PTS. Maltose is a LamB/non-PTS nutrient, and so its mechanisms of transport differ from those of glucose at both the outer and inner membranes.

Although the structural components of the PTS and non-PTS pathways are distinct, the PTS exerts some regulatory control over non-PTS pathways (SAIER 1989). Depending on the presence or absence of PTS nutrients, the glucose-IIA protein of the PTS may either inhibit or enhance the expression and function of non-PTS inner membrane transport systems. Therefore, there is the strong potential for mutations that affect the PTS to have pleiotropic effects on transport of non-PTS nutrients.

Based upon the known physiological mechanisms for nutrient uptake, we generated four a priori expectations concerning the relative performances of the derived genotypes in various nutrient environments. Expectation 1: If OmpF was an important target of selection, then the mean fitness of the derived genotypes in trehalose (where LamB is required for rapid growth) should be less than mean fitness in other PTS nutrients. Expectation 2: If the PTS was an important target of selection, then mean fitness in OmpF/non-PTS nutrients should be less than mean fitness in OmpF/PTS nutrients. Expectation 3: If the PTS was an important target of selection, then the genetic variance in fitness should higher in OmpF/non-PTS nutrients than in OmpF/PTS nutrients. Expectation 4: More generally, if the PTS was an important target of selection, then the mean

![Figure 1](https://example.com/figure1.png)

**Figure 1.**—Classification of the 12 nutrients used in this study by the mechanisms of their transport through the outer (OmpF vs. LamB) and inner (PTS vs. non-PTS) membranes. See text for details.
Ara-). Spontaneous arabinose-utilizing mutants (ha+) are formed when large numbers of cells are spread on a medium that contains arabinose (TA indicator agar). Ara' mutants form white colonies, while Ara- mutants are red.

The evolution is strictly clonal. The ancestor is prototrophic, although it is unable to use L(+)-arabinose as a nutrient source.

The general enzymes (I and HPr) form a two-step phosphate transfer pathway: enzyme I picks up phosphate from phosphoenolpyruvate, an end product of glycolysis, and HPr transfers phosphate from enzyme I to the nutrient-specific enzymes. For each nutrient-specific pathway, there are three or four specific protein domains (IIA, IIB, IIC, and IID). Class IIC and IID proteins form nutrient-specific transmembrane channels, whereas class IIA and IIB enzymes transfer phosphate from HPr to the nutrient. Arrows indicate the flow of phosphate transfer.

### Fitness Responses and Their Among-Genotype Variances

Fitness responses and their among-genotype variances, taken together, should show qualitatively different patterns in OmpF/non-PTS and OmpF/PTS nutrients. In particular, responses that are tightly coupled to physiological traits important for fitness in glucose should show both a large fitness increase and similar responses in the replicate lines.

### Materials and Methods

**Bacterial strains and culture conditions:** The derivation of the genotypes used in this study has been previously described (Lenski et al. 1991). Briefly, all strains were derived from a single genotype of *E. coli* B. The ancestral genotype does not contain any plasmids or functional phage, and therefore its evolution is strictly clonal. The ancestor is prototrophic, although it is unable to use L(+)-arabinose as a nutrient; (i.e., Ara\(^{-}\)). Spontaneous arabinose-utilizing mutants (Ara\(^{+}\)) are easily selected when large numbers of cells are spread on minimal arabinose medium. When cells are grown on tetrazolium-arabinose (TA) indicator agar, Ara\(^{+}\) mutants form white colonies while Ara\(^{-}\) colonies are red (Levin et al. 1977; Lenski 1988a).

Twelve clones, six Ara\(^{-}\) and six Ara\(^{+}\), were used to found separate populations that were then propagated for 2000 generations at 37°C in Davis minimal medium (Carlton and Brown 1981) containing 25 μg/ml of glucose as the sole usable carbon and energy source. These conditions support ~5 × 10\(^{10}\) cells/ml at stationary phase. Propagation was performed daily by transferring 0.1 ml of the previous day's culture into 9.9 ml of fresh medium (a 100-fold dilution), so that each population went through ~6.6 generations of binary fission every day. After 2000 generations (300 days), single genotypes were randomly picked from each population and stored at ~80°C in a glycerol-based suspension. All of the experiments reported in this paper were performed using these derived genotypes and their Ara\(^{-}\) and Ara\(^{+}\) progenitors.

**Measurement of selection rate constants in competition experiments:** Competition experiments were performed under the same culture conditions used to propagate the evolving populations, except that other carbon sources were substituted for glucose. The nutrient concentration was the same, 25 μg/ml (w/v), in all cases. In these experiments, two genotypes were mixed together and allowed to compete for the same pool of limiting nutrient. One competitor was one of the 12 derived genotypes and the other was the ancestral genotype bearing the reciprocal Ara marker, so that the two competitors could be readily distinguished.

The protocol for the competition experiments was as follows. Both genotypes were first grown separately for 24 hr (one complete growth cycle) in the same type of medium in which they would compete. The two genotypes were then mixed by diluting each one 200-fold into a flask containing 9.9 ml of fresh medium. Thus, the total population size was initially ~1% of the final population size, as was the case during the evolution experiment itself. The initial and final population densities for each competitor were estimated by plating the mixed culture onto TA agar.

The selection rate constant, \(r\), is a measure of relative performance and is given by

\[
r = \frac{\ln[N(1)/N(0)] - \ln[N(1)/N(0)]}{1\text{ day}}
\]

where \(N(0)\) and \(N(1)\) are the initial densities of the derived and ancestral genotypes, respectively, and \(N(1)\) and \(N(0)\) are their corresponding densities after 1 day (Nagy\&\,laki 1977; Lenski et al. 1991; Travisano et al. 1995a). The selection rate constant is thus equal to the difference in the two genotypes' realized Malthusian parameters during head-to-head competition for the same pool of limiting resources. (Any difference between the genotypes in their plating efficiencies does not affect this quantity, provided that plating efficiencies remain constant between initial and final samples.)

Previously, Lenski et al. (1991) and Vasi et al. (1994) used the ratio of Malthusian parameters as a measure of performance (=relative fitness, \(W\)). However, this ratio is very sensitive to sampling error if the two competitors have very different Malthusian parameters, as was sometimes the case in certain nutrient environments. Because the selection rate constant is a difference, it is much less sensitive to sampling error in such cases (Travisano et al. 1995a). The following formula gives an approximate conversion between a selection rate constant, \(r\), and the corresponding relative fitness, \(W\) (Lenski et al. 1991):

\[
W = 1 + (r/m),
\]

where \(m\) is the average Malthusian parameter (here, \(m = \ln 100 = 4.6\) per day).

**Experimental design:** In each nutrient environment, we obtained a minimum of three estimates of the selection rate.
constant for each of the 12 derived genotypes and five estimates for each ancestral genotype, in all cases relative to the reciprocally marked ancestral genotype. Replicate assays were performed in sets of complete blocks.

RESULTS

Selective neutrality of the genetic marker: To evaluate the effect of the arabinose-utilization marker on fitness, the Ara- and AraC variants of the common ancestor were allowed to compete in each type of medium. In every case, the Ara marker was selectively neutral within the limits of resolution and using $P < 0.05$ as the criterion for significance (Table 1). In one nutrient, mannose, the marker had a marginally nonsignificant effect ($P = 0.054$), but given that 12 tests were performed, such an outcome is probably not indicative of a meaningful difference. Although the Ara marker was neutral in the ancestral genetic background, it is conceivable that it might interact epistatically with (and be subject to selection in) the derived genetic backgrounds. However, the six Ara+-derived genotypes (as a group) are no more or less fit than the six Ara" derived genotypes (as a group) in any of these sugars, excepting mannitol (Table 2). Given that 12 such tests were performed, one marginally significant ($P = 0.044$) difference is again probably not meaningful, since we would expect one test in 20 to appear significant by chance alone. On the basis of these results, and the fact

![Figure 3. Transport across the inner membrane and subsequent catabolism of the OmpF/non-PTS nutrients used in this study. After transport through the inner membrane, the catabolic pathways rapidly converge.](image)

**TABLE 1**

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>No. of replicates, $n$</th>
<th>Selection rate constant (day$^{-1}$) for Ara$^+$ relative to Ara$^{-}$</th>
<th>$P^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fructose$^c$</td>
<td>7</td>
<td>0.067 ($\pm$0.093)</td>
<td>0.495</td>
</tr>
<tr>
<td>Galactose$^c$</td>
<td>9</td>
<td>0.079 ($\pm$0.091)</td>
<td>0.411</td>
</tr>
<tr>
<td>Glucitol$^c$</td>
<td>9</td>
<td>$-0.043$ ($\pm$0.115)</td>
<td>0.716</td>
</tr>
<tr>
<td>Glucose$^c$</td>
<td>5</td>
<td>0.018 ($\pm$0.058)</td>
<td>0.767</td>
</tr>
<tr>
<td>Glycerol$^c$</td>
<td>9</td>
<td>0.072 ($\pm$0.075)</td>
<td>0.364</td>
</tr>
<tr>
<td>Lactose$^c$</td>
<td>11</td>
<td>0.298 ($\pm$0.218)</td>
<td>0.202</td>
</tr>
<tr>
<td>Maltose$^c$</td>
<td>6</td>
<td>$-0.138$ ($\pm$0.086)</td>
<td>0.169</td>
</tr>
<tr>
<td>Mannitol$^c$</td>
<td>9</td>
<td>0.041 ($\pm$0.057)</td>
<td>0.490</td>
</tr>
<tr>
<td>Mannose$^c$</td>
<td>9</td>
<td>$-0.297$ ($\pm$0.132)</td>
<td>0.054</td>
</tr>
<tr>
<td>Melibiose$^c$</td>
<td>9</td>
<td>$-0.103$ ($\pm$0.083)</td>
<td>0.251</td>
</tr>
<tr>
<td>NAG$^c$</td>
<td>12</td>
<td>0.034 ($\pm$0.075)</td>
<td>0.660</td>
</tr>
<tr>
<td>Trehalose$^c$</td>
<td>12</td>
<td>0.000 ($\pm$0.043)</td>
<td>0.999</td>
</tr>
</tbody>
</table>

$^a$Mean ($\pm$SE of the mean) calculated from $n$ replicate assays.

$^b$Two-tailed probability (from the $t$ distribution with $n - 1$ degrees of freedom) of rejecting by chance the null hypothesis that the selection rate constant equals zero, indicating equal fitness for the Ara$^+$ and Ara$^{-}$ marker states.

$^c$This study.

$^d$Data from Travisano et al. (1995a).
that all experiments were fully balanced with respect to the arabinose marker, we pool the data for the Ara⁻ and Ara⁺ genotypes in subsequent analyses.

**Changes in performance of the derived genotypes in each nutrient:** We first sought to determine whether the average performance of the derived genotypes had changed in each of the 12 nutrient environments (Table 3). We applied a sequential Bonferroni correction to account for the fact that a total of 12 t-tests were performed (Rice 1989). The 12 derived genotypes had, on average, improved significantly in six nutrients: fructose, glucitol, glucose, glycerol, mannitol, and NAG. Average performances also increased in lactose and mannose, but these changes were marginally nonsignificant (0.05 < P < 0.1). The average performance of the derived genotypes had declined in two other nutrients, galactose and melibiose. Even in those nutrients where performance declined, however, all of the derived genotypes were able to grow sufficiently to establish self-sustaining populations in the absence of the ancestral competitor.

The derived genotypes fall into at least six distinct phenotypic classes: We next sought to ascertain the number of distinct phenotypic classes among the 12 independently derived genotypes. For this purpose, the 12 genotypes were treated as fixed (rather than random) factors, and we performed t-tests for multiple comparisons (Miller 1981; SAS Institute 1988; Travisano et al. 1995a). A Bonferroni-corrected multiple-range test was performed for the 12 genotypes, in each nutrient, with each test having an experimentwise type I error rate of 0.05. Then, to account for the fact that we performed 12 such tests (one for each nutrient), a further Bonferroni correction was incorporated, so that the experimentwise type I error rate for each nutrient was 0.05/12 = 0.00417. One or more derived genotypes were placed in a separate phenotypic class if they differed significantly (after the double Bonferroni correction) from all other derived genotypes in their performance in at least one nutrient. This discrimination procedure is highly conservative. Nonetheless, the procedure is statistically rigorous and indicates that there are at least six distinct phenotypic classes among the 12 independently derived genotypes. Ara⁻¹, Ara⁻², Ara⁺², Ara⁺⁴, and Ara⁺⁶ each require separate phenotypic classes, as does the pair Ara⁺⁶ and Ara⁺¹. The five remaining genotypes cannot be unambiguously assigned to any one of these phenotypic classes, nor do they necessarily require the construction of a new class.

**Expectation 1—derived genotypes are less fit in trehalose than in OmpF/PTS nutrients:** If OmpF was an important target of selection during the 2000 generations of experimental evolution, then the average performance of the derived genotypes should be significantly lower on trehalose (which is a LamB/PTS nutrient) than on the five OmpF/PTS nutrients, excluding glucose. (We excluded glucose from this and all subsequent tests because it was the one nutrient for which enhanced performance was directly selected, whereas changes in performance for the other 11 nutrients constitute correlated responses. All of the test results would become even more significant if glucose was included as one of the OmpF/PTS nutrients.) Table 4 summarizes the five paired t-tests comparing the selec-

**TABLE 2**

| Nutrient | Ara⁻³-derived relative to Ara⁺ ancestor | Ara⁺-derived relative to Ara⁻ ancestor | Difference | P*
|-----------|----------------------------------------|---------------------------------------|------------|----
| Fructose ¹ | 1.338                                  | 0.993                                 | 0.345 ± 0.240 | 0.182
| Galactose ² | -0.442                                 | -0.633                                | 0.191 ± 0.229 | 0.423
| Glucitol ³ | 0.953                                  | 1.340                                 | -0.388 ± 0.348 | 0.292
| Glucose ⁴ | 1.300                                  | 1.355                                 | -0.035 ± 0.078 | 0.663
| Glycerol ⁵ | 0.905                                  | 1.637                                 | -0.732 ± 0.786 | 0.373
| Lactose ⁶ | 1.102                                  | 2.615                                 | -1.513 ± 1.278 | 0.264
| Maltose ⁷ | -0.244                                 | 0.244                                 | -0.487 ± 0.468 | 0.322
| Mannitol ⁸ | 1.490                                  | 1.746                                 | -0.256 ± 0.123 | 0.044
| Mannose ⁹ | 0.227                                  | 0.629                                 | -0.403 ± 0.333 | 0.255
| Melibiose ¹⁰ | -2.751                                 | -1.772                                | -0.979 ± 0.535 | 0.097
| NAG ¹¹ | 1.566                                  | 1.442                                 | 0.123 ± 0.096 | 0.277
| Trehalose ¹² | 0.410                                  | -0.084                                | 0.494 ± 0.687 | 0.489

*Difference (±standard error of the difference) in the selection rate constants based on six independently derived genotypes in each marker class.

¹ Two-tailed probability (from the t-distribution with 6 + 6 - 2 = 10 degrees of freedom) of rejecting by chance the null hypothesis that the difference in the selection rate constants for the two marker classes of derived genotype equals zero.

² This study.

³ Data from Travisano et al. (1995a).
Selection rate constants estimated in trehalose and each of the OmpF/PTS nutrients. To adjust for multiple comparisons with a single treatment group, Dunnett’s method was employed to insure a 0.05 maximum experimentwise error rate (Dunnett 1964; SAS Institute 1988). The derived genotypes are less fit, on average, in trehalose than in all five of the OmpF/PTS nutrients; in four cases, this difference is significant even after adjusting for multiple comparisons (Table 4). These differences between a LamB/PTS and several OmpF/PTS nutrients support the hypothesis that OmpF was an important target of selection during the long-term propagation in glucose minimal medium.

Expectation 2—derived genotypes are more fit in OmpF/PTS than in OmpF/non-PTS nutrients: If the PTS was also an important target of selection during the experimental evolution, then the derived genotypes should perform significantly better in the five OmpF/PTS nutrients (excluding glucose) than in four OmpF/non-PTS nutrients. To test this hypothesis, we ranked the nine OmpF nutrients by the proportion of derived genotypes (out of 12) that had improved performance relative to the ancestor (i.e., a positive value for the selection rate constant). A one-tailed Mann-Whitney U-test was then performed on these rankings. (We did not perform a parametric test because of the highly divergent genetic variances in performance on the different nutrients, as reported below). As seen in Table 5 (second column), the derived genotypes were in fact more likely to have improved fitness in OmpF/PTS nutrients than in OmpF/non-PTS nutrients, although this difference was only marginally significant \( P = 0.0317 \). Thus, it appears that changes in the PTS were also responsible for some of the evolutionary adaptation to the glucose environment.

Expectation 3—derived genotypes are more variable
in OmpF/non-PTS nutrients than in OmpF/PTS nutrients: As another possible indication that the PTS was a target of selection during the experimental evolution, we predicted that the derived genotypes would be more variable in their performances in OmpF/non-PTS nutrients than in OmpF/PTS nutrients. The increased genetic variation in performance on OmpF/non-PTS nutrients is expected because of the lack of selection to improve or even maintain non-PTS transport functions. This variation might be generated by either random genetic drift causing damage to genes that encode non-PTS functions or heterogeneity among alleles that improve performance on PTS nutrients in terms of their pleiotropic effects on non-PTS functions. For each nutrient, we performed a two-way analysis of variance (with derived genotype and block as random effects) of the selection rate constants. The genetic variance component, $V_c$, was then estimated as the difference in the genotype and error mean-squares, divided by the number of replicate assays (blocks) per genotype. We then ranked the nine OmpF nutrients by their corresponding genetic variances and performed a one-tailed Mann-Whitney $U$-test. As shown in Table 5 (fourth column), this index gives completely nonoverlapping values for the OmpF/PTS and OmpF/non-PTS nutrients, and the separation is highly significant ($P = 0.0079$). This result is consistent with our expectation that functions that are important in the selective environment should exhibit both improved performance and reduced genetic variability relative to traits that are unimportant in that environment. In other words, the 12 independently derived genotypes exhibited significantly parallel adaptive evolution with respect to the PTS, indicating that it was an important target of selection.

We also computed this index of parallelism for the three other nutrients in our study. Glucose yielded a value of 200.00, while maltose gave 0.90 and trehalose 0.44. It is quite striking that this index was highest for glucose, followed by the five OmpF/PTS nutrients that share transport mechanisms across both the outer and inner membrane (Figure 4). The lowest values of the composite statistic were for the six nutrients that differ from glucose in one or both of these transport mechanisms.

### DISCUSSION

We studied genotypes from 12 populations of *E. coli* that had been propagated for 2000 generations in a glucose-limited environment. LENSKI et al. (1991) found relatively little genetic variation for mean fitness (measured in the glucose environment) among the replicate populations. Measurements of the demographic parameters that conferred the fitness gains also showed a high degree of parallelism, suggesting that the derived genotypes had independently achieved similar adaptations to the common selective regime (VASSI et al. 1994).

In contrast, our results indicate that the indepen-
were usually more fit than the ancestor and they tended to have no effect on fitness in the glucose-limited environment used in the evolution experiment. The random fixation of mutations in these nonexpressed genes would reduce average performance and increase genetic variation for performance in nutrients where LamB and/or non-PTS transport mechanisms are used. (2) Pleiotropy hypothesis: alleles that were selected in the replicate lines because they caused similar improvements in performance on glucose could nonetheless differ in terms of their pleiotropic effects on other transport functions. Although the proteins used to transport OmpF/PTS and other nutrients are distinct, such pleiotropy could arise via the regulation of gene expression; indeed, the PTS system is known to exert regulatory control over certain non-PTS functions.

Travisano et al. (1995a) argued previously that pleiotropy was primarily responsible for the heterogeneity of the derived genotypes with respect to their fitnesses in maltose medium. Their reasoning was based upon the dynamics of genetic drift, given the relatively few gene loci involved in the specific transport and subsequent utilization of maltose but not glucose. Here, we address the relative likelihood of drift vs. pleiotropy in a similar manner, by comparing the evolutionary responses of the derived genotypes within and between the OmpF/PTS and OmpF/non-PTS nutrient groups.

We first consider the explanation for the changes in performance in those media where another OmpF/PTS nutrient was substituted for glucose. When competing for these OmpF/PTS nutrients, the derived genotypes usually have improved fitness and exhibit relatively little genetic variation for fitness (Figure 4). Only a few genes are required for uptake and catabolism of the five OmpF/PTS nutrients used in this study that are not also required when cells are grown on glucose, and these nutrient-specific genes are encoded by only \( \sim 5000 \) base pairs (Table 6 and Figure 2). We assume that mutations in these nutrient-specific genes are selectively neutral in the glucose environment. We can then use this information to examine the possibility that random genetic drift might have caused the observed changes in fitness on OmpF/PTS nutrients other than glucose. The expected number of nucleotide substitutions caused by random drift is given by

\[
\text{number of substitutions} = (m)(g)(b),
\]

where \( m \) is the mutation rate per nucleotide per generation, \( g \) is generations, and \( b \) is the number of potentially mutable nucleotide bases (Kimura 1983). Given an average mutation rate of \( \sim 5 \times 10^{-10} \) per nucleotide base pair per generation for \( E. coli \) (Drake 1991), then the expected number of mutations fixed by drift that might affect fitness in any one of the five OmpF/PTS nutrients (but not in glucose) in each population during 2000 generations is only \( \sim 5.5 \times 10^{-5} \). This estimate is independent of effective population size (and hence of selection at other loci), and it is also liberal (high) because it assumes that any mutation in a nutrient-spe-
TABLE 6

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Genes</th>
<th>Nucleotide bases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fructose</td>
<td>fruFKA,fruR</td>
<td>6,254</td>
</tr>
<tr>
<td>Glucitol</td>
<td>gudOPABDMR</td>
<td>4,270</td>
</tr>
<tr>
<td>Mannitol</td>
<td>muiOPADR</td>
<td>4,141</td>
</tr>
<tr>
<td>Mannose</td>
<td>manA, manXYZ, nagC</td>
<td>4,128</td>
</tr>
<tr>
<td>NAG</td>
<td>nagDCABE</td>
<td>6,483</td>
</tr>
</tbody>
</table>

* Number of nucleotide bases in the indicated genes and operons as determined by counting from the upstream regulatory region to the termination site (including intergenic regions).


Data from Yamada and Saier (1987, 1988).

Data from Lee and Saier (1983), Davis et al. (1988), and Figs. et al. (1994).

The manXYZ operon is regulated by the NagC protein. Data from Miles and Guest (1984), Erni et al. (1987), and Plumbridge and Kolb (1991).

Data from Peri and Waygood (1988) and Peri et al. (1990).

Specific gene affects fitness in that nutrient but had no effect in the selective environment (glucose). Thus, it is unlikely that drift contributed much to the changes in performance of the derived genotypes in the other OmpF/PTS nutrients. Moreover, fitness had systematically increased in these other OmpF/PTS nutrients, which is a very unlikely outcome of random drift. Clearly, the simplest explanation is that the mutations that were selected in the glucose environment have beneficial pleiotropic effects on performance in other OmpF/PTS nutrients.

Compared with the other OmpF/PTS nutrients, the OmpF/non-PTS nutrients have many more differences in their uptake and subsequent utilization from glucose (Figure 3). The OmpF/non-PTS nutrients differ not only in the nutrient-specific transport enzymes but also in the energetic force that drives transport. However, we do not include the differences in energetic force in our evaluation of the potential role of random genetic drift, as any genetic changes that affect ATP and/or ion pools are likely to have widespread pleiotropic effects (given their fundamental importance to the cell) and therefore seem unlikely to be the result of drift. We present two lines of evidence that pleiotropy, and not drift, was also the more important population genetic process causing changes in fitness in the OmpF/non-PTS nutrients. First, consider the very different fitness responses of the derived genotypes to melibiose and lactose. In melibiose, the derived genotypes are systematically less fit than the ancestor, whereas in lactose the derived genotypes tend to be more fit (Table 3), and yet the pathways for degradation of these two nutrients are quite similar (Figure 3). The amount of melibiose-specific sequence (melA, melB, and melR) are encoded by ~5000 bases; Yazyu et al. 1984; Liljestrom and Liljestrom 1987; Webster et al. 1989) that is distinct from that required for growth on lactose (lacI, lacY, and lacZ are encoded by ~6000 bases; Farabaugh 1978; Buechel et al. 1980; Kalnins et al. 1983) is quite small, and, for the reasons stated above, these genes are unlikely to have been subject to the effects of drift. Therefore, it seems more likely that the dissimilar responses of the derived genotypes to melibiose and lactose are caused by pleiotropic effects of mutations that are beneficial in glucose and systematically hinder performance in melibiose but not lactose. Second, the amount of glycerol-specific sequence (glpD, glpF, glpK, and glpR) are encoded by ~3300 base pairs; Ye and Larson 1988; Muramatsu and Mizuno 1989; Weissenborn et al. 1992), distinct from that required for the uptake and utilization of glucose, is again no more than that for the other OmpF/PTS nutrients; but the correlated fitness responses of the derived genotypes in glycerol medium are quite different from those in the OmpF/PTS nutrients, in this case being far more variable (Table 5) rather than differing in sign as was the case for melibiose. Unless mutation rates vary substantially among different nutrient-specific pathways, then it seems that both the similarity of the fitness responses to OmpF/PTS nutrients and the heterogeneity of these responses to other nutrients must be caused by pleiotropic effects of mutations selected for their beneficial effects in glucose, rather than random genetic drift of mutations at other loci.

Specific targets of selection: Maltose and glucose differ in their utilization by E. coli primarily in having distinct transport mechanisms. Travisano et al. (1995a) demonstrated that glucose transport was an important target of selection by showing that the derived genotypes had significantly lower fitness in maltose than in glucose. From those earlier results, however, it was unclear what aspect(s) of glucose transport had improved. The results presented here imply that genetic changes affecting transport across both the outer and inner membranes were responsible for the improved fitness in glucose.

Trehalose is a PTS substrate that requires the outer membrane protein LamB for effective uptake (Klein and Boos 1993). The finding that the derived genotypes are significantly more fit in several OmpF/PTS nutrients than in this LamB/PTS nutrient indicates that changes in OmpF or in its expression were responsible for some of the observed fitness gains. LamB and OmpF are both transmembrane proteins that allow solutes to pass from the external medium to the periplasmic space located between the two membranes (Hancock 1987), but LamB forms a large maltodextrin-specific channel (Ferenci et al. 1980) that is highly expressed only dur-
ing growth in media containing trehalose or maltodextrin. Because glucose is not sterically hindered in passing through OmpF, increasing its channel size is not expected to appreciably increase the rate of diffusion (Schindler and Rosenbusch 1978; Cowan et al. 1992). It seems more likely that there has been an increase in the production of OmpF, thereby yielding a higher density of OmpF channels in the outer membrane.

Improved transport of glucose across the inner membrane is indicated by the more general improvement and greater uniformity of the derived genotypes in OmpF/PTS nutrients than in OmpF/non-PTS nutrients. The PTS transfers phosphate from phosphoenolpyruvate, an intermediate of central metabolism, to a molecule of nutrient concomitant with that molecule's transport into the cytoplasm. This process is carried out by a combination of general proteins (enzymes I and HPr, except for fructose which uses a specialized enzyme FPr instead of HPr) and nutrient-specific permease complexes (Figure 2) (Saier et al. 1970; Kornberg 1990). At the relatively low nutrient concentrations used in this study, only enzyme I is required by all of the PTS pathways (Kornberg 1990; Postma et al. 1993). The improvement of the derived genotypes on most or all of the OmpF/PTS nutrients therefore suggests that changes in enzyme I or its expression are probably responsible for at least some of the adaptation to glucose.

While our findings indicate that changes in the PTS are partly responsible for the pattern of fitness responses to the different nutrients, a more precise explanation for the particular responses is a matter of speculation. We present three plausible explanations, all of which focus on the integral role of enzyme I in PTS function. (1) The pts operon, which encodes enzyme I (ptsI), also encodes the glucose-IIA protein, which serves as both a glucose-specific transport enzyme and exerts regulatory control of the PTS over non-PTS functions. Thus, certain genetic changes that affect expression of enzyme I may also affect expression of glucose-IIA protein and, concomitantly, may have strong pleiotropic effects on the transport of non-PTS nutrients (De Reuse and Danchin 1988; van der Vlag et al. 1994). Depending on the particular changes in the pts operon in different derived genotypes, these pleiotropic effects could yield the pattern of higher genetic variance for fitness in non-PTS than PTS nutrients. (2) Transcription of ptsI is positively regulated by glucose permease (IIBC), but not by mannose permease (IIC and IID) (De Reuse and Danchin 1991). Of the OmpF/PTS nutrients included in this study, the average fitness improvement in mannose was less than half that in any other OmpF/PTS nutrient (Table 3). An analysis of the relationships among PTS permeases indicates that mannose permease is strongly divergent in nucleotide sequence from the other sequenced permeases (Saier et al. 1992), suggesting that the interaction of glucose permease and enzyme I may have been a target of selection during the experimental evolution. (3) Enzyme I occurs as a monomer but can also form a dimer via a reversible dimerization (Kukuruzinska et al. 1982), and the dimer to monomer transition may be required for PTS-mediated phosphate transfer (Licalsi et al. 1991). However, the enzyme I monomer-to-dimer transition is several orders of magnitude slower than the other steps in PTS-mediated phosphate transfer (Chauvin et al. 1994a,b), suggesting that the very slow rate of enzyme I dimerization might also have been a target of selection.

Population divergence and mechanism: On the basis of significant differences in their fitness responses to novel nutrient environments, the 12 independently derived genotypes formed at least six physiologically distinct groups. This divergence of populations derived from the same ancestral state did not require either environmental differences or small population sizes, which are the two usual explanations for phenotypic divergence. Rather, it seems that substantial phenotypic divergence occurred because adaptation depended on the stochastic appearance and subsequent fixation of beneficial mutations.

Unfortunately, it is quite difficult to infer the specific physiological and genetic differences among the independently derived genotypes, despite the fact that a great deal is known about the relevant mechanisms of nutrient utilization. For example, genotype Ara+6 is distinguished by being more fit in maltose and lactose, and less fit in glucitol, than any of the other 11 derived genotypes. But if Ara+6 is so fit in maltose, then why is it not especially fit in trehalose, which (like maltose) also requires the LamB outer membrane protein for efficient transport? And if Ara+6 is so fit in lactose, then should it not also be quite fit in galactose, given the overlap in the two pathways (see Figure 3)? Similar difficulties arise in interpreting each of the six distinct physiological groups. These difficulties may arise because each derived genotype acquired several beneficial mutations during the 2000 generations of experimental evolution (Lenski et al. 1991), which may have had distinct pleiotropic effects and may even have interacted epistatically, thus obscuring the effect of any particular mutation. In essence, these difficulties may reflect the lack of quantitative information on the performance and functioning of E. coli as a whole organism, rather than as an atomized set of its various traits (Gould and Lewontin 1979; see also Maddox 1992).

At the other extreme, evolutionary models that emphasize the effects of natural selection and ecological interactions can often predict some phenotypic outcomes of selection, without incorporating much mechanistic detail (Rose 1984; Lenski 1988b; Mueller 1988; Bull et al. 1991). However, such models may miss important aspects of evolutionary change, and it is a matter of debate whether models that ignore underlying mechanistic details are sufficient to predict evolutionary change (Lewontin 1974; Rose et al. 1987; Dykhuizen et al. 1992).
and DEAN 1990; LENSKI et al. 1991, 1994; BULL and MOLINEUX 1992). For example, BULL and MOLINEUX (1992, p. 92) state that simple models of selection cannot predict "correlated responses to selection or the multiplicity of genetic states satisfying the selected phenotypic criterion." Consistent with this view, the replicative genotypes in this study exhibited a variety of correlated responses when challenged with novel nutrient environments. While the overall pattern and variety of correlated responses makes sense in the light of mechanistic understanding of the transport and subsequent catabolism of nutrients by E. coli, the particular responses of any single genotype were not easily reconciled with known mechanisms. This discrepancy suggests that the ability to predict phenotypic outcomes of adaptive evolution may be severely limited by physiological and genetic complexity, even in systems that are relatively simple and well understood.

In summary, the pattern of direct and correlated fitness responses in populations of E. coli indicates that mechanisms for the transport of glucose across both the outer and inner membranes were important targets of selection during 2000 generations of experimental evolution. Having thus narrowed the search for candidate loci, we hope to identify specific genetic changes that occurred. Once these changes have been identified, it should be possible to examine their effects, both individually and in concert, on fitness in the various nutrient environments.

We thank A. F. BENNETT, L. J. FORSEY, S. A. FRANK, S. KALISZ, J. A. MONGOLD, M. S. ROBERTS, M. R. ROSE, M. H. SAIER JR., D. O. STRANEY, D. R. TAYLOR, J. M. TIEDE, S. J. TONSON, P. E. TURNER, and F. VASI for many valuable discussions. The editorial suggestions of A. A. HOFFMANN and two anonymous reviewers were very helpful. This research was supported by a grant from the National Science Foundation (DEB-9421237) to R.E.L. and by the National Science Foundation (BIR-9120006).}

LITERATURE CITED


LENISKI, R. E., V. SOUZA, L. P. DUONG, Q. G. PHAN, T. N. M. NGUYEN et al., 1994 Epistatic effects of promoter and repressor functions


Communicating editor: A. A. HOFFMANN