SELECTIVE NEUTRALITY OF 6PGD ALLOZYMES IN *E. COLI* AND THE EFFECTS OF GENETIC BACKGROUND

DANIEL DYKHUIZEN AND DANIEL L. HARTL

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907

Manuscript received July 14, 1980

ABSTRACT

We have used gluconate-limited chemostats to study selective differences between isogenic strains of *Escherichia coli* K12 into which four naturally occurring alleles coding for allozymes of 6-phosphogluconate dehydrogenase (6PGD) had been transferred. The limit of detectability of selection with our procedures is a selection coefficient of 0.5%. In the normal *E. coli* K12 genetic background, all alleles are selectively neutral or nearly neutral. The absence of detectable selection does, however, depend on genetic background and on such environmental factors as cell density. In a genetic background containing a mutation that cuts off the alternative metabolic route for 6-phosphogluconate, selection between allozymes can be detected, and the selection is in the direction expected from the measured apparent K_m values of the allozymes. Even when the alternative metabolic route is not blocked by mutation, one of the 6PGD allozymes has a detrimental, but density-dependent, interaction with a mutation conferring resistance to bacteriophage T5. In all cases, the observed selection is due to the allozymes themselves (or to associated regulatory elements), as the selection disappears when the chemostats are limited by a different carbon source (ribose plus succinate). Nevertheless, the four alleles do seem to be selectively neutral or nearly neutral in the normal *E. coli* K12 genetic background. Moreover, the distribution of allele frequencies in natural populations of *E. coli* is in accord with the expectations of selective neutrality.

I am inclined to suspect that we see, at least in some [cases], variations which are of no service to the species, and which consequently have not been seized on and rendered definite by natural selection. . . . Variations neither useful nor injurious would not be affected by natural selection, and would be left either a fluctuating element, as perhaps we see in certain polymorphic species, or would ultimately become fixed. . . . We may easily err in attributing importance to characters, and in believing that they have been developed through natural selection; . . . many structures are now of no direct use to their possessors, and may never have been of any use to their progenitors . . . . [On the other hand,] we are much too ignorant in regard to the whole economy of any organic being to say what slight modifications would be of importance or not.

CHARLES DARWIN, 1872

HERE we present evidence that, in a normal genetic background, at least four naturally occurring alleles for allozymes of 6-phosphogluconate dehydrogenase in *Escherichia coli* are selectively neutral or nearly neutral. Since the

Supported by U. S. Public Health Service Grant GM24886.

late 1960s, when the widespread occurrence of allozyme polymorphisms became evident (HARRIS 1966; LEWONTIN and HUBBY 1966), the field of population genetics has been dominated by the issue of whether these polymorphisms are subject to natural selection or whether they are selectively neutral (KIMURA 1968, 1979; KING and JUKES 1969; LEWONTIN 1974; NEI 1975; AYALA 1976; HARTL 1980). The issue is central to an understanding of the mechanisms of adaptive evolution. If observed protein polymorphisms are under selective maintenance, then they themselves are the stuff of evolution, and efforts must be made to unravel the physiological bases of their selection. If, on the other hand, such polymorphisms are selectively neutral under all conditions, then they are irrelevant to adaptive evolution and interest should be refocused on other classes of genes.

Experimental resolution of the allozyme controversy has been beyond reach. The problem has been examined from almost every conceivable angle — statistical, experimental and ecological — yet, the data point ambiguously in both directions. Allozyme clines are suggestive of selection, but clines can be produced by founder effects and, even if created by selection, may be due to selection acting on closely linked loci in linkage disequilibrium rather than on the allozymes themselves (CLEGG and ALLARD 1972; ALLARD *et al.* 1972; CHRISTIANSEN and FRYDENBERG 1974). Statistical tests of selective neutrality are conceptually elegant but lacking in power, and a distribution of allele frequencies inconsistent with neutrality may be due to misidentified alleles or to the absence of mutational equilibrium (EWENS and FELDMAN 1976; EWENS 1980). Allozymes of a particular enzyme may have different kinetic or physical properties, but it is not yet clear whether such chemical differences are indicative of differences in fitness (GIBSON 1970; HARPER and ARMSTRONG 1973; DAY, HILLIER and CLARKE 1974; KOEHN 1978; SOMERO 1978).

Direct measurement of fitness components in such eukaryotes as Drosophila is limited to the detection of large differences because of the substantial sampling variance of such estimates (PROUT 1965, 1969, 1971a,b). Experimental populations can theoretically be used to detect selection coefficients as small as about 5% (YAMAZAKI 1971), but interpretations are sometimes uncertain in the face of possible linkage disequilibrium between the allozyme locus and nearby loci that are unrecognized, but subject to selection. Failure to find selection in experimental populations leaves a nagging ambiguity because fitness depends on an interplay between genotype and environment, and the particular environment of an experimental population may be unconducive to the expression of real differences in fitness between genotypes. In other words, experimental populations often lack a suitable positive control that demonstrates selection against as obviously defective allele. Such positive control experiments are difficult to devise in organisms such as Drosophila because the usual growth medium is complex and chemically undefined and because, in many instances, the role of the allozyme in metabolism is unknown.

In those instances in which the metabolic role of the allozyme is known, more rigorous experiments can be carried out. In the case of the alcohol dehydrogenase

locus in *Drosophila melanogaster*, for example, systematic changes in allele frequency occur in experimental populations provided with alcohol-supplemented medium, but not in populations provided with glycerol-supplemented medium (VAN DELDEN, KAMPING and VAN DIJK 1975; VAN DELDEN, BOERMA and KAMPING 1978; CAVENER and CLEGG 1978); this finding strongly implicates the allozymes themselves or associated regulatory elements as being responsible for the observed selection. Similarly, in the case of the amylase locus in *D. melanogaster*, systematic changes in allele frequency occur in the presence of medium containing starch, but not in the presence of medium containing glucose (DE JONG and SCHARLOO 1976; YARDLEY 1978; HICKEY 1978; SCHARLOO and DE JONG 1980). On the other hand, although the results for alcohol dehydrogenase and amylase provide irrefutable evidence of selection, negative results in such experiments are of uncertain significance because selection coefficients can be as large as several percent and nevertheless remain undetected with present techniques. Experiments in Drosophila are therefore one-sided; selection, when it occurs, can be proven, but the apparent lack of selection cannot be rigorously interpreted because of insufficient resolving power of the techniques.

In the prokaryotic system reported here, the limit of resolution of the technique is a selection coefficient of 0.5%. With suitable positive controls to show that selection can be detected and suitable negative controls to show that the locus under scrutiny is the one responsible for the selection, we have studied four allozyme-associated alleles of 6-phosphogluconate dehydrogenase in isogenic strains of *Escherichia coli* in chemostats. In the normal *E. coli* K12 genetic background, selection, if it occurs at all, leads to an overall fitness difference of less than 0.5%. Thus, the alleles are selectively neutral or nearly neutral. However, in certain other genetic backgrounds — in one case involving a laboratory-derived mutation that cuts off alternative metabolic routes for 6-phosphogluconate; in another case involving a laboratory-derived mutation that confers resistance to bacteriophage T5 — selection has been detected. We interpret these results as meaning that allozymes have a latent potential for producing selective differences in certain genetic backgrounds or perhaps under particular environmental circumstances, but that, in the normal genetic background, the alleles are nearly neutral.

BACKGROUND

Bacterial strains for our study were generously provided by ROGER MILKMAN, who was first to demonstrate the occurrence of surprisingly high levels of allozyme polymorphism in natural populations of *E. coli* (MILKMAN 1973). More recently, SELANDER and LEVIN (1980) surveyed 20 enzymes among 109 independent clones, including many in MILKMAN's original collection, and found a level of polymorphism substantially greater than that observed in natural populations of eukaryotes. The significance of the finding of such high levels of polymorphism in a haploid species is at present unclear.

The overall design of our experiments is conceptually quite simple and consists of four phases. First, standard genetic methods are used to transfer all rele-

vant alleles into the genetic background of *E. coli* K12. Second, a growth medium is employed that demonstrates the occurrence and detectability of selection against an obviously defective allele at the locus in question; this growth medium constitutes the positive control. Third, a different growth medium is employed that eliminates selection against the defective allele; this growth medium constitutes the negative control; its purpose is to verify that observed selective differences are due to the locus under study. Fourth, isogenic strains carrying various alleles of the allozyme locus are studied in pairwise competition in both types of growth media in order to estimate their selection coefficients.

A primary experimental tool for the study of nutrient-limited competition is the bacterial chemostat (KUBITSCHEK 1970; DYKHUIZEN 1978). A chemostat is basically a culture vessel having an inflow aperture for the influx of sterile nutrient medium and an outflow aperture for the efflux of exhausted medium, living cells and cellular debris. In practice, the device is complicated by various attachments for aeration of the medium and for the prevention of contamination. The rate of division of cells within the chemostat can be varied by a factor of 10 by appropriate adjustment of the rate of inflow, but an equilibrium will eventually be reached in which the number of new cells created by division is exactly balanced by the number of cells washed out in the overflow. Thus, chemostats provide an environment in which cell division is continuous, but population size is held constant.

The dynamics of competition for substrate in chemostats is approximated reasonably well by equations known as the MONOD equations (DYKHUIZEN and HARTL 1981). For present purposes, the most important implication of the MONOD equations is as follows: if $p(t)$ and $q(t)$ [with $p(t) + q(t) = 1$] represent the relative proportions of two competing strains after t hours of competition, and if, for a fixed rate of inflow of nutrient medium, the relative "fitnesses" of the two strains, measured in terms of Malthusian parameters, are $1 + s$ and 1, respectively, then $\ln[p(t)/q(t)] = \ln[p(0)/q(0)] + st$. Consequently, the slope of the linear regression of $\ln[p(t)/q(t)]$ against time provides an estimate of the selection coefficient, s.

The absence of genetic recombination between strains in chemostats can be used to advantage in order to avoid extravagant amounts of electrophoresis. Our isogenic strains are initially produced in genetic backgrounds that are sensitive to bacteriophage T5 ($T5^S$). From each strain, we select a spontaneous mutation at the *tonA* locus that confers resistance to the bacteriophage ($T5^R$). In the experiments, a strain bearing one allozyme and $T5^R$ is always placed in competition with a strain bearing another allozyme and $T5^S$, and the change in the allozyme frequencies is monitored through changes in the frequencies of $T5^R$ and $T5^S$. That the allozyme-$T5$ association remains intact in the chemostat has been verified by electrophoresis, and, within the range of $T5^R$ and $T5^S$ frequencies employed, changes in frequency due to mutation pressure are inconsequential. Our use of $T5^R$ and $T5^S$ is valid only insofar as $T5^R$ and $T5^S$ are themselves selectively neutral in the experimental chemostats. This requirement has been verified by appropriate controls discussed below.

In order to evaluate our statistical procedures, we have carried out extensive computer simulations of the sampling process and have examined the adequacy of our estimates of $ln[p(t)/q(t)]$ and the validity of the linear regression. Our normal sampling procedure is as follows: A sample from the chemostat overflow is collected and the cells counted in a Petroff-Hausser counting chamber in order to estimate total cell density. The sample is diluted with buffer to a density of 700 cells/ml, and aliquots are added to equal volumes of molten agar to bring the density to 350 cells/ml. Bacteriophage T5, which is present in excess in half the molten-agar samples, kills the $T5^S$ cells in these samples. From both the T5-containing and the T5-free samples, 30 aliquots of 0.1 ml each are dispensed in a single petri dish by means of a Colworth Droplette. The agar is solidified and incubated overnight, and the small colonies that arise are magnified and counted the next day. All colonies within any drop examined are counted, and the counts continue until the colony count exceeds 325 or until all 30 drops have been counted, whichever comes first. In this manner we obtain two estimates: (A) an estimate of the total cell density, which is obtained from the T5-free samples, and (B) an estimate of the $T5^R$ cell density, which is obtained from the T5-containing samples. Reconstruction experiments have demonstrated that the sampling procedure is valid for estimating A and B. The quantity $(A\text{-}B)$ estimates the density of $T5^S$ cells, and $\ln[B/(A\text{-}B)]$ is our estimate of $\ln(p/q)$, which is the quantity used in the linear regression.

Assuming binomial sampling of $T5^R$ and $T5^S$ cells and a Poisson distribution of colonies per drop, we have simulated 500 replicates of the sampling procedure for various values of p, the frequency of $T5^R$ cells. For values of p ranging between 0.01 and 0.65, the estimator $\ln[B/(A\text{-}B)]$, has been found to be unbiased and to have an approximately constant variance. In our experiments, therefore, we have strived to remain within this range of p.

As for the power of the procedure to detect selection, we have carried out simulations of the sampling procedure for seven equally spaced time points over 100 hours and have performed a linear regression analysis of $\ln[B/(A\text{-}B)]$ against time. The change in frequency in the chemostat was assumed to follow $ln\ (p/q) = C + st$, where s was allowed to assume values between 0 and -0.05 per hour. The results for 512 simulations when the initial frequency of $T5^R$ was $p = 0.5$ are shown in Table 1. The numbers in the table give the proportion of cases (out of 512) in which the null hypothesis $s = 0$ was rejected at the 5% level. As expected, rejection occurs 5% of the time when the true value of s equals 0. Rejection always occurs when the absolute value of s is as large as 0.05. When the absolute value of s is 0.005 per hour, we have a 50% chance of rejection in any one chemostat; thus, we regard this as the realistic limit of resolution of our present procedures.

The allozymes that we have studied are those of 6-phosphogluconate dehydrogenase (EC 1.1.1.43), which are due to alleles at the *gnd* locus at 44 minutes on the standard *E. coli* genetic map (BACHMANN and Low 1980). Construction of isogenic strains was accomplished through the use of the nearby histidine operon (*his*, at 44 minutes) and by the inability of *edd*$^-$ *gnd*$^-$ cells to grow on

TABLE 1

Proportion of 512 simulated cases in which the hypothesis that $s = 0$ *could be rejected at the 5% level*

True $\|s\|$	Probability of rejection (percent)	True $\|s\|$	Probability of rejection (percent)
0	5.08	0.006	59.6
0.001	5.66	0.007	77.7
0.002	11.9	0.008	87.1
0.003	19.9	0.009	96.5
0.004	31.0	0.01	97.1
0.005	47.8	0.05	100.0

Calculations based on an analysis of variance of the regression of $\ln(p/q)$ against time, p initially assumed to equal 0.5 and $s \leq 0$. Regression based on seven equally spaced time points over 100 hours.

gluconate as their sole carbon source. (*edd*, at 41 minutes, codes for phosphogluconate dehydratase [EC 4.2.1.12], which converts 6-phosphogluconate into 2-keto-3-hydroxygluconate-6-phosphate, thus bypassing 6-phosphogluconate dehydrogenase by providing an alternative metabolic route for 6-phosphogluconate.) The principal genetic tool in strain construction was transduction mediated by bacteriophage P1(*cml clr100*), which carries a chloramphenicol-resistance gene permitting the selection of extremely rare lysogens and also a temperature-sensitive repressor leading to high-titer lysates when induced by high temperature. Unless otherwise specified, all genetic manipulations followed the procedures of MILLER (1972).

A key strain in our constructions is designated *DD715*, which is genotypically *edd*$^-$ *gnd*$^-$ *his*$^-$ *rpsL*. (*rpsL*, at 72 minutes, confers resistance to streptomycin.) *DD715* was constructed as follows: *DF1070* (*HfrC edd1 gnd1 relA1 ton A22 T2*^{R}), obtained from B. BACHMANN, was mated with *AM3100* (*his rpsL*), and the His$^+$ Str^R phenotype was selected. *edd1 gnd1 T5*^{S} *T2*^{S} exconjugants were tested for *relA1* by the leucine step-down test of ROGERSON and FREUNDLICH (1970). All exconjugants, and *AM3100* itself, proved to be *relA*$^-$. One of the *edd1 gnd1 T5*^{S} *T2*^{S} *relA*$^-$ exconjugants was exposed to penicillin selection for His$^-$, and a spontaneous, revertible *his* mutation was selected. To eliminate the *relA* mutation in this strain, the strain was first rendered *thyA*$^-$ by trimethoprin selection, then transduced to *fuc*$^-$ *thyA*$^+$ with a P1 lysate from a *fuc*$^-$ *thyA*$^+$ strain provided by T. NEWMAN; the *fuc*$^-$ *thy*$^+$ strain (still *relA*$^-$) was transduced to *fuc*$^+$ *thyA*$^+$ *relA*$^+$ using a P1 lysate from strain *DD323* (DYKHUIZEN and DAVIES 1980). The resulting strain — *edd1 gnd1 his*$^-$ *rpsL* — was designated *DD715*. Another strain of importance in our constructions is *DD722* (*edd1 gnd1 rpsL*), which was obtained as a P1-mediated *his*$^+$ transductant of *DD715*.

Each of the MILKMAN strains to be used in the subsequent analysis was made lysogenic for P1 (*cml clr100*), and lysates were used to transduce *DD715*, selection being for *his*$^+$ and subsequent scoring for *gnd*$^+$. This cycle of transduction was carried out again, and the resulting *edd1 gnd*$^+$ *his*$^+$ *rpsL* strain was made

lysogenic for P1, lysates then being used to transduce the isogenic *DD722*, selection in this case being for *gnd*$^+$ At this point, a spontaneous *tonA* mutation conferring resistance to phage T5 was selected in each strain and confirmed by co-transduction of the T5 resistance phenotype with *pan6*, and each strain was then transduced to *edd*$^+$ with P1 phage from an isogenic lysogen (*DF1071* = *HfrC gnd1 relA1 tonA22 T2*^{R}, obtained from B. BACHMANN). The *gnd* allozyme in each strain was verified electrophoretically after each transduction of *gnd*$^+$, using the methods of MILKMAN (1973). Altogether, three consecutive transductions into isogenic strains were carried out for each *gnd* allele to be studied. Rates of transduction in the first two transfers were abnormally low (of order 10^{-7} to 10^{-9}), but by the third transduction the rate was in the normal range for *E. coli* K12.

An odd and unexpected feature of the transductions should be mentioned here, as it is potentially important in light of experimental results to be discussed below. All of the MILKMAN strains that we used were resistant to bacteriophage T5, and B. R. LEVIN (personal communication) has found that 97 of 109 natural clones of *E. coli* are resistant. However, among the *gnd*$^+$ *his*$^+$ transductants produced from P1 lysates from the original strains, roughly half were still resistant to bacteriophage T5. The genetic and physiological basis of the resistance is unknown. Since the T5-resistance factor co-transduces with *his*, however, it is evidently not due to the *tonA* locus, which is at three minutes. T5 resistance among wild strains of *E. coli* may therefore involve loci other than *tonA*. A possible reason for a low frequency of *tonA* in natural populations is suggested by the work of MOSER (1958), who found that *tonA* cells are selectively neutral in energy-limited chemostats, but disfavored when nitrate or phosphate is limiting. In any case, in the study at hand, the *his*-linked resistance was a mere nuisance factor that was eliminated by choosing the T5-sensitive transductants for subsequent genetic manipulation.

RESULTS

Results for four naturally occurring *gnd* alleles (*i.e.*, *W*$^+$, *F2*, *S4* and *S8*) in the *edd*$^+$ genetic background are summarized in Table 2. Each row in the table provides the genotypes of the competing strains, the carbon source in the chemostat medium, the number of experiments, the selection coefficient [estimated as the slope of $\ln(T5^R/T5^S)$ against time], the statistical significance of the slope based on the null hypothesis $s = 0$, the F value obtained in the analysis of variance of the regression of $ln(T5^R/T5^S)$ against time, and the number of degrees of freedom. In those cases with two replicate experiments, the results were invariably homogeneous, and the entries in the last four columns of the table pertain to the pooled results. However, replicate experiments were carried out with different initial frequencies of $T5^R$ to guard against the possibility of frequency-dependent selection; the "2" marked with an asterisk denotes that the replicates had initial frequencies of 0.50 and 0.10, and in all other cases the replicates had initial frequencies of 0.50 and 0.01. No evidence of frequency-dependent selection was found.

TABLE 2

Results of competition between isogenic strains of E. coli *K12 carrying various alleles at the* gnd *locus*

	Genotype $T5^S$	Genotype $T5^R$	Medium	No. expts.	s	Stat. sign.	F	d.f.
System controls:	+	—	gluconate	1	−0.0876 ± 0.0072	***	148.1	4
	+	—	ribose + succinate	1	−0.0012 ± 0.0027	ns	0.2	6
	—	+	gluconate	1	0.0931 ± 0.0032	***	827.0	5
	—	+	ribose + succinate	1	0.0159 ± 0.0029	**	29.9	6
Strain controls:	*F2*	*F2*	gluconate	2	−0.0017 ± 0.0013	ns	2.0	11
	W+	*W+*	gluconate	2	0.0001 ± 0.0023	ns	0.1	10
	S4	*S4*	gluconate	2	−0.0018 ± 0.0018	ns	0.5	11
	S8	*S8*	gluconate	2	−0.0165 ± 0.0022	***	29.5	10
	S8	*S8*	ribose + succinate	1	0.0019 ± 0.0022	ns	0.7	6
Competitions:	*F2*	*S4*	gluconate	2	−0.0026 ± 0.0013	ns	2.2	10
	S4	*F2*	gluconate	2	−0.0005 ± 0.0013	ns	0.9	8
	S8	*F2*	gluconate	2	−0.0019 ± 0.0034	ns	0.3	9
	S4	*W+*	gluconate	2*	0.0018 ± 0.0016	ns	0.7	10
	S8	*W+*	gluconate	2	−0.0034 ± 0.0020	ns	2.2	10
	W+	*S8*	gluconate	2	−0.0170 ± 0.0024	***	26.7	10
	W+	*S8*	ribose + succinate	1	−0.0038 ± 0.0039	ns	1.0	6

Each row in the table gives the strains involved in the competition, the chemostat medium, the number of experiments, the selection coefficient s [defined as the slope of $\ln(T5^R/T5^S)$ against time], the F value in the analysis of variance of the regression and the number of degrees of freedom. Gluconate medium consists of 0.4 gm/l gluconate and Davis salts; ribose + succinate medium consists of 0.1 gm/l ribose, 0.3 gm/l succinate and Davis salts; flow rates of the chemostats were so adjusted to produce a doubling time of approximately two hours (see DYKHUIZEN 1978 for details about procedures). In the number-of-experiments column, the asterisk denotes replicates with initial $T5^R$ frequencies of 0.50 and 0.10; otherwise, replicates had initial frequencies of 0.50 and 0.01. In the statistical significance column, the symbols *ns*, ** and *** denote nonsignificance, significance at the 1% level and significance at the 0.1% level, respectively. The selection coefficients are calculated per hour; those wishing an approximate per-generation selection coefficient can multiply the tabulated values by 2.

Source of alleles: *gnd+* = *gnd* allele in *E. coli* K12; *gnd⁻* = *gnd1* allele in strain *DD715*; *gndW+* = *gnd* allele in MILKMAN strain *RM43A*, electrophoretically indistinguishable from *gnd+* in *E. coli* K12; *gndF2* = *gnd* allele in MILKMAN strain *RM202F*; *gndS4* = *gnd* allele in MILKMAN strain RM72B; *gndS8* = *gnd* allele in MILKMAN strain *RM200Q*. *gndS8* scores as a null allele with electrophoresis in our hands, but enzyme activity is normal in extracts. MILKMAN (1973) gives the frequencies of *gndW+*, *gndF2*, *gndS4*, and *gndS8* among 825 wild clones as 0.72, 0.035, 0.025 and 0.011, respectively.

Many of the entries in Table 2 are controls of various sorts, and these warrant a brief discussion. In the first two entries, a strain carrying *gnd*$^+$ (the *gnd* allele in normal *E. coli* K12) with *T5*^{S} was grown in competition with an isogenic strain carrying a *gnd*$^-$ point mutation (*gnd1*) with *T5*^{R}. Selection on gluconate medium is significant; whereas, no selection is observed on medium consisting of a mixture of ribose and succinate. Thus, the positive control in the experimental design consists of gluconate medium, because selection against a defective allele is observed, and the negative control in the experiments consists of ribose-succinate medium, because this medium eliminates the selection against the *gnd*$^-$ allele. The selection noted in the first entry of Table 2 is therefore specific to growth on gluconate. In the second two entries in Table 2, a *gnd*$^-$ *T5*^{S} strain is grown in competition with a *gnd*$^+$ *T5*^{R} strain. There is again selection against *gnd*$^-$ in gluconate medium, but in this case there is also significant selection against the *gnd*$^-$ *T5*^{S} strain in ribose-succinate medium. We interpret the selection in the latter case as being due to inadvertent differences in genetic background, and it is noteworthy that the selection on gluconate is some six-fold greater than it is on ribose-succinate, which indicates that the greater part of the selection is specific to gluconate. It should be emphasized that the selection against *gnd*$^-$ cells in the upper part of Table 2 is very strong; it is about 25% of the theoretical maximum amount of selection that could be obtained at a two-hour doubling time, the maximum being obtained when one strain cannot grow at all.

Controls for the selective equivalence of *T5*^{S} and *T5*^{R} are shown in the second part of Table 2. Since one of the *gnd* alleles (*S8*) interacts curiously with *T5*^{R}, discussion of this allele will be taken up below. For the others (*W*$^+$, *F2* and *S4*), no selection involving the bacteriophage resistance is observed.

Competition experiments involving different allozyme alleles are shown in the third part of Table 2. Again postponing consideration of the *S8* allele, it is evident that *F2*, *S4* and *W*$^+$ are selectively equivalent within the limit of resolution of the technique. Results for four chemostats — a positive control, a negative control and two experimental — are shown graphically in Figure 1.

Since selective differences between *F2*, *S4* and *W*$^+$ would involve selection coefficients of at most 0.5%, we regard these alleles as being selectively neutral or nearly neutral. Allele *S8* is an exception, however, and it illustrates the potential importance of interaction effects between an allozyme allele and other alleles in the genetic background. Pertinent data are shown in the middle region of Table 2, which indicate that, in gluconate medium, selection does occur favoring *S8 T5*^{S} over *S8 T5*^{R}. Suspecting that the selection might be due to inadvertent lack of isogenicity or to an odd *tonA* (*T5*^{R}) allele, we again selected a spontaneous *tonA* mutation in the *S8 T5*^{S} strain, but selection of the same magnitude and direction as that found previously was observed. That the selection involves a detrimental interaction between *S8* and *tonA* is indicated by the finding that selection against *S8 T5*^{R} occurs only in gluconate medium, not in the ribose-succinate control. We do not understand the metabolic basis of the *S8-tonA* interaction, but it is interesting that the effect is density dependent. When *S8 T5*^{S} is

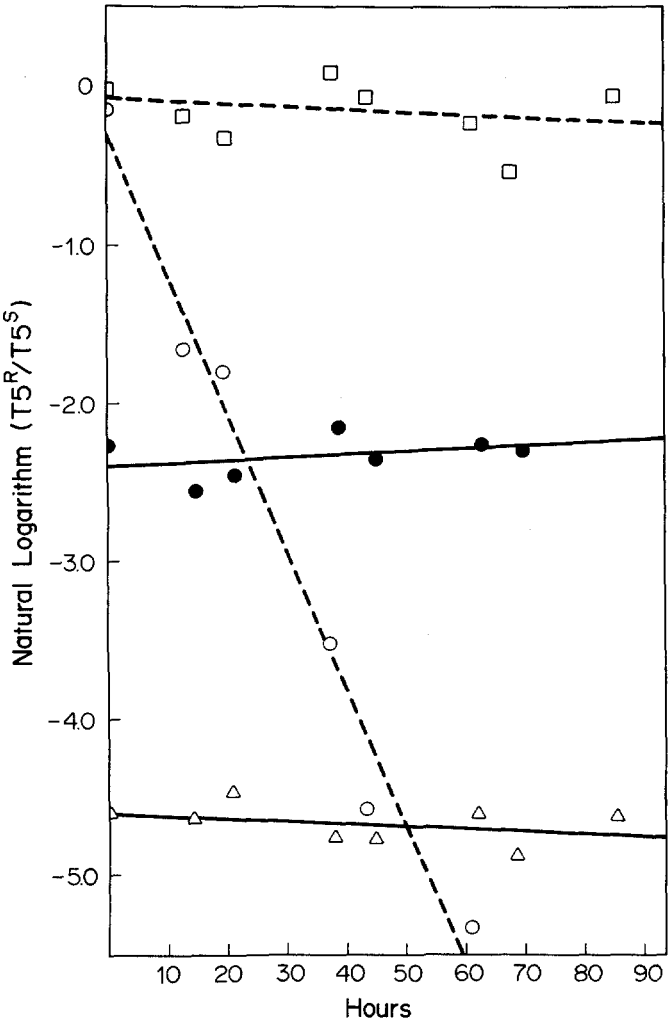

FIGURE 1.—Results of some of the experiments summarized in Table 2. Broken lines indicate controls, solid lines experimental chemostats. The positive control (open circles) is *gnd*$^-$ *T5*^{R} *vs. gnd*+ *T5*^{S} grown in gluconate medium. The negative control (squares) is *gnd*$^-$ *T5*^{R} *vs. gnd*+ *T5*^{S} grown on ribose-succinate medium. Experimental results pertain to *gndF2 T5*^{S} *vs. gndS4 T5*^{R} (triangles) and to *gndS4 T5*^{S} *vs. gndW*+ *T5*^{R} (closed circles), both types of experimental chemostats grown on gluconate.

grown in competition with *S8 T5*^{R} in medium containing only 0.04 gm/l gluconate (in contrast to the usual 0.4 gm/l), the equilibrium cell density is reduced by a factor of about 10, and the observed selection coefficient becomes $s = 0.0009 \pm 0.0024$, which does not differ significantly from zero.

Other results relevant to the *S8-tonA* interaction are shown in the third part of Table 2. In a *T5*^{S} genetic background, allele *S8* is selectively equivalent to the others. Selection occurs only in the *T5*^{R} genetic background and is specific to

growth on gluconate. We do not know whether the *S8-tonA* interaction is important in natural populations. One could argue that it is important on grounds that naturally occurring strains are resistant to bacteriophage T5. One could argue with equal justification that the interaction is not important because, as note earlier, T5 resistance of naturally occurring strains does not seem to be due to the *tonA* locus.

One possible physiological reason for the selective equivalence of *F2*, *S4*, W^+ and *S8* $T5^s$ might be the existence of the edd^+ gene product, which provides an alternative metabolic pathway for 6-phosphogluconate. Small differences in the functional efficiency of the allozymes might be compensated metabolically by small changes in the rate of flow of substrate through the alternative pathway. The pathway involving edd^+ is apparently the only alternative for 6-phosphogluconate, as edd^- gnd^- cells are unable to grow on gluconate. To examine what may be called the physiological compensation hypothesis, we undertook a study of *gnd* allozymes in an edd^- genetic background. Construction of the relevant strains has been described previously, but the final transduction of edd^+ was not carried out.

The study of selection in the edd^- genetic background is difficult because edd^- cells do not grow well in gluconate; with excess gluconate, the doubling time for edd^+ cells is about 60 minutes, that for edd^- cells is about 120 minutes. This difference is reflected in the time required for edd^- cells to attain their equilibrium cell density in gluconate-limited chemostats. With edd^+ cells, the equilibrium density is reached in a matter of about 10 hours; with edd^- cells, attainment of the equilibrium cell density requires about three days. Part of this three-day adaptation involves genetic changes in the strains, as cells removed from the chemostat and grown in excess gluconate have a doubling time of about 75 minutes. Moreover, some of the genetic adaptation involves changes in the *gnd* region itself because, when the *gnd* region from adapted strains is transduced into nonadapted cells, the doubling time of the edd^- gnd^+ transductants in excess gluconate is now about 90 minutes, not the expected 120 minutes.

Because adaptation involves genetic changes in the *gnd* region, it is unwarranted to attribute selective differences between adapted strains to the original *gnd* allele that the strains carried. In one case, however, we are confident that selection due to the 6-phosphogluconate dehydrogenase allozymes does occur. This case involves *F2* and *S4*, which reproducibly yields selection in favor of *S4*, whether or not the original strains have been preadapted to growth on gluconate medium before the experiment is carried out. Results for nonadapted and preadapted strains are shown in Figure 2. Note again that the selection is specific for gluconate. When the cells are grown on ribose medium, the selection disappears. However, selection in gluconate chemostats continues for only about 50 to 60 hours, after which there is no further change in frequency. The constancy in frequency after 50 to 60 hours is not due to growth of cells on the chemostat walls, as such cultures maintain their constant frequency even when transferred to fresh chemostats.

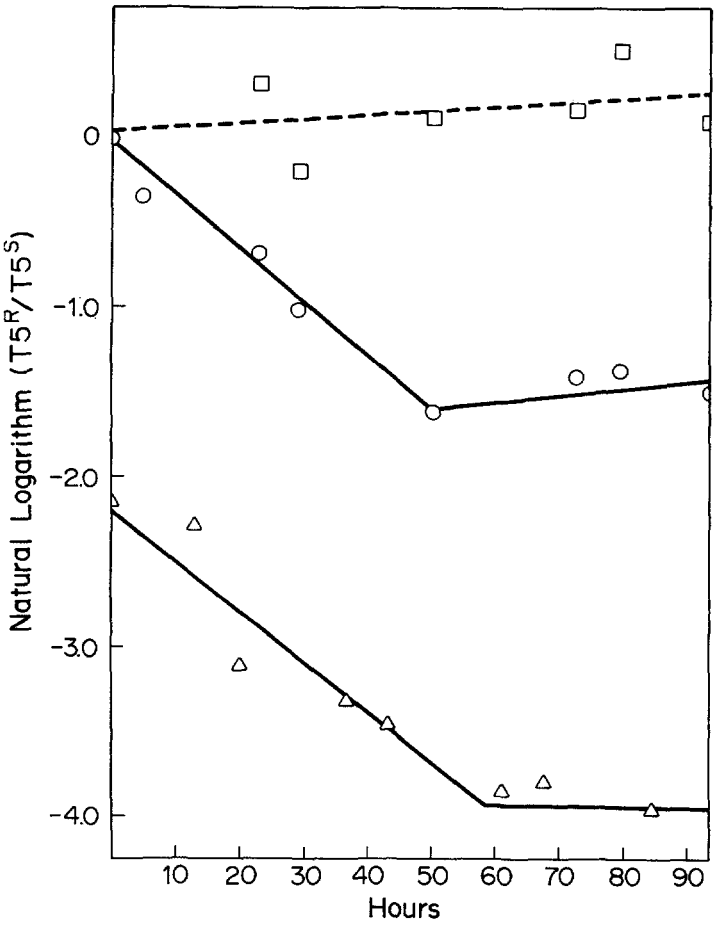

FIGURE 2.—Selection involving *gnd* alleles in an *edd*⁻ genetic background, strains *DD808 (gndF2 T5^R edd1) vs. DD811 (gndS4 T5^S edd1)*. Squares: ribose-limited control chemostat (0.4 gm/1 ribose) after strains had undergone adaptation in separate gluconate-limited chemostats for 3 days. Circles: gluconate-limited experimental chemostat involving pre-adapted strains as above. Triangles: gluconate-limited experimental chemostate involving *DD808* and *DD811*, but without prior adaptation. Straight-line segments are based on linear regressions. Doubling times in the chemostats ranged from 2.0 to 2.5 hr.

It is perhaps worthwhile to note that the selection in Figure 2 is in the direction that would be expected from the apparent K_m values for the *S4* and *F2* allozymes. In particular, the apparent K_m value for 6-phosphogluconate of the *S4* enzyme is slightly (but consistently) smaller than the apparent K_m value of the *F2* enzyme. When both K_m values are measured at the same time, using the same reagents, the apparent K_m value for the *S4* enzume is 2.7×10^{-2} mM and that of the *F2* enzyme is 3.7×10^{-2} mM (see Figure 3 for data and methods).

DISCUSSION

The results in Figure 2 support the hypothesis of metabolic compensation. In normal *edd*$^+$ cells, small differences in the functional efficiency of allozymes are not reflected as differences in fitness because the small differences are automatically compensated by means of alternative metabolic pathways. In *edd*$^-$ cells, the alternative pathway for gluconate is blocked, metabolic compensation becomes impossible and functional differences between allozymes do become important for fitness. Biochemical differences between allozymes are not limited to those between *F2* and *S4* discussed above. The apparent K_m value for 6-phosphogluconate of the *S8* enzyme is 8.7×10^{-2} mM, for example, although that for the W^+ enzyme is not distinguishable from that for the *F2* enzyme (see Figure 3). Yet, as shown in Table 2, *F2*, *S4*, W^+ and *S8* $T5^S$ are all selectively equivalent in an *edd*$^+$ genetic background. Of course, the interaction effects that we observed between *S8* and $T5^R$ (*i.e.*, *tonA*), and those observed between *gnd* allozymes and *edd*$^-$ may not be due to the *gnd* structural loci themselves, but rather to differences in closely linked regulatory sequences. The agreement between the observed selection in an *edd*$^-$ genetic background and the K_m values of *F2* and *S4* suggest to us that the allozymes themselves are involved, but the agreement could be coincidental. From the standpoint of the evolutionary implications of our results, it makes very little difference whether the discussion is phrased in terms of structural gene or regulatory polymorphisms.

It might be worthwhile here to summarize the evolutionary implications of our results. With respect to allozymes of *gnd* in a normal *E. coli* K12 genetic background in the admittedly artificial, but highly stressful, environment of a bacterial chemostat, the allozymes are selectively neutral or nearly neutral (*i.e.*, selection cofficients less than 0.005). The allozymes, or perhaps their associated regulatory sequences, do have the potential for being selected, as evidenced by the interaction effects observed between *S8* and *tonA* and between the allozymes and *edd*$^-$. Thus, we believe, the selective potential of allozymes can be realized in certain genetic backgrounds or perhaps under particular combinations of environmental circumstances. Although, at the moment, we can speak only about the allozymes of 6-phosphogluconate dehydrogenase, the method of analysis is readily applicable to many other polymorphic loci.

Does the inference of near neutrality of *gnd* allozymes carry over to natural populations of *E. coli*? ROBERT K. SELANDER and BRUCE R. LEVIN have kindly provided us with data from their analysis of 109 naturally occurring *E. coli* strains, and DAVID HAYMER of our laboratory has performed a statistical test of selective neutrality based on the observed distribution of allele frequencies. In the SELANDER-LEVIN sample, there were 15 distinguishable alleles of *gnd*, making this locus one of the most polymorphic. Using the computer-simulation methods of WATTERSON (1978), HAYMER generated 1,000 simulated samples of size 109 containing exactly 15 alleles and confroming to EWENS' (1980) formula for selective neutrality. Compared to the expectations based on selective neutrality,

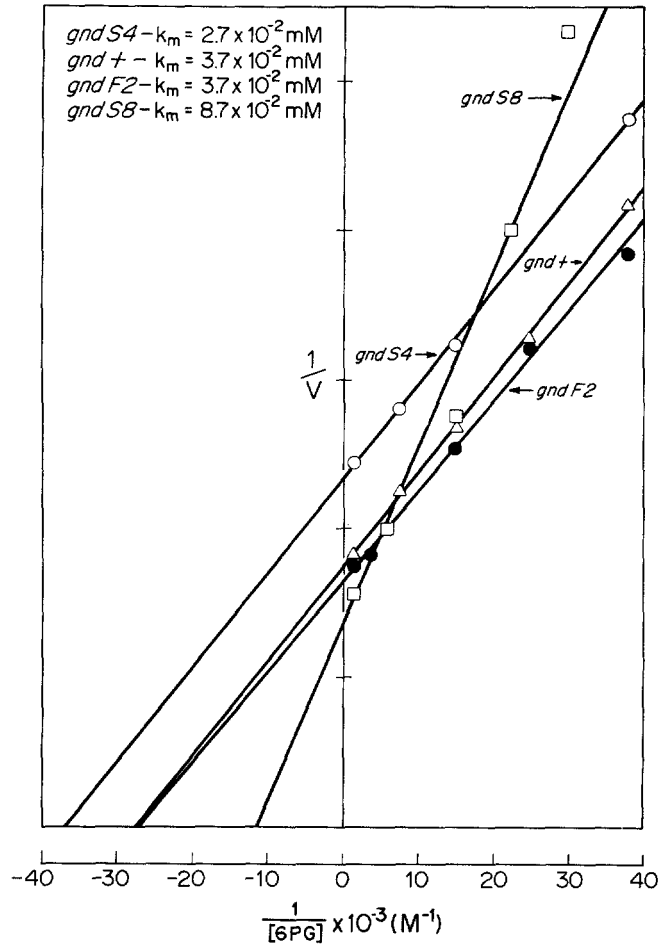

FIGURE 3.—Double reciprocal plots of measurements of four 6PGD allozymes. Assays were carried out simultaneously using the same reaction mixtures and were monitored spectrophotometrically according to the procedures of EISENBERG and DOBROGOSZ (1967) except that 2 μmoles of 6-phosphogluconate (trisodium) and 1.2 μmoles of NADP were used and the reactions were carried out at 37°. The apparent K_m values are calculated as the negative reciprocal of the x-intercept, which is independent of the scale of the y-axis. The apparent K_m values are within the range reported for purified enzyme (DE SILVA and FRANKEL 1979).

the observed distribution of allele frequencies has an excess of the most frequent allele, but the excess is not statistically significant. In short, the observed distribution of allele frequencies is not incompatible with the hypothesis of selective neutrality. On the other hand, statistical tests of neutrality are rather low in their power to detect selection. Thus, while *gnd* allozymes do seem to be selectively nearly neutral, in both laboratory and natural populations, the selective potential of allozymes based on genetic interactions may sometimes be realized in natural populations, though not with sufficient intensity or duration to alter the overall distribution of allele frequencies from neutrality.

One further comment: population genetics is split between those who think that natural selection must somehow be involved in the maintenance of allozyme polymorphisms and those who think that allozyme polymorphisms are selectively neutral or nearly neutral. Those committed to selection can argue that we have studied only a restricted range of environments with gluconate as the energy source, that the chemostat is not a suitable experimental model of the mammalian intestine, that the population size of *E. coli* is so large that even experimentally undetectable selection coefficients are evolutionarily significant, and that our finding of significant selective effects due to nonallelic interactions is proof positive of the potential for selection between allozymes in natural populations. Those committed to neutrality, on the other hand, can argue that gluconate was chosen as the energy source in order to maximize the opportunity for selection and to amplify such fitness differences as may exist, that competition between *E. coli* strains in chemostats is at least as intense as competition within the mammalian intestine, that the *effective* size of *E. coli* populations is many orders of magnitude less than the actual size because of clonal reproduction and periodic selection, and that the observed nonallelic interactions involve *tonA* and *edd*⁻ mutations, which have not been reported to occur in natural populations. Our own view is intermediate. The results suggest to us that the allozymes of 6PGD are selectively neutral or nearly neutral, but that they have a latent potential for selection; this potential may on occasion be realized under the appropriate conditions of environment or genetic background, and under these conditions allozymes can become the raw material for adaptive evolution.

Thanks to R. MILKMAN for bacterial strains, R. K. SELANDER and B. R. LEVIN for providing their data, D. HAYMER for carrying out the statistical test of neutrality and especially LUANNE WOLFGRAM for excellent technical assistance. Thanks also to R. MILKMAN, B. R. LEVIN, M. LEVY and J. MCNAIR for their helpful comments on the manuscript.

LITERATURE CITED

ALLARD, R. W., G. R. BABBEL, M. T. CLEGG and A. L. KAHLER, 1972 Evidence for coadaptation in *Avena barbata*. Proc. Natl. Acad. Sci. U.S. **69**: 3043–3048.

AYALA, F. J. (editor), 1976 *Molecular Evolution*. Sinauer, Sunderland, Mass.

BACHMANN, B. J. and K. B. LOW, 1980 Linkage map of *Escherichia coli* K-12, Edition 6. Microbiol. Rev. **44**: 1–56.

CAVENER, D. R. and M. T. CLEGG, 1978 Dynamics of correlated genetic systems. IV. Multilocus effects of ethanol stress environments. Genetics **90**: 629–644.

CHRISTIANSEN, F. B. and O. FRYDENBERG, 1974 Geographical patterns of four polymorphisms in *Zoarces viviparus* as evidence of selection. Genetics **77**: 765–770.

CLEGG, M. T. and R. W. ALLARD, 1972 Patterns of genetic differentiation in the slender wild oat species *Avena barbata*. Proc. Natl. Acad. Sci. U.S. **69**: 1820–1824.

DARWIN, C., 1872 *The Origin of Species*, 6th ed. p. 60, 88, 181, 182, 184. New American Library of World Literature, New York (1958).

DAY, T. H., P. C. HILLIER and B. CLARKE, 1974 The relative quantities and catalytic activities of enzymes produced by alleles at the alcohol dehydrogenase locus in *Drosophila melanogaster*. Biochem. Genet. **11**: 155–165.

VAN DELDEN, W., A. C. BOERMA and A. KAMPING, 1978 The alcohol dehydrogenase polymorphism in populations of *Drosophila melanogaster*. I. Selection in different environments. Genetics **90**: 161–191.

VAN DELDEN, W., A. KAMPING and H. VAN DIJK, 1975 Selection at the alcohol dehydrogenase locus in *Drosophila melanogaster*. Experientia **31**: 418–420.

DYKHUIZEN, D., 1978 Selection for tryptophan auxotrophs of *Escherichia coli* in glucose limited chemostats as a test of the energy conservation hypothesis of evolution. Evolution **32**: 125–150.

DYKHUIZEN, D. and M. DAVIES, 1980 An experimental model: Bacterial specialists and generalists competing in chemostates. Ecology **61**: 1213–1227.

DYKHUIZEN, D. and D. HARTL, 1981 Evolution of competitive ability in *Escherichia coli*. Evolution **61**: 1213–1227.

EISENBERG, R. C. and W. J. DOBROGOSZ, 1967 Gluconate metabolism in *Escherichia coli*. J. Bact. **93**: 941–949.

EWENS, W. J., 1980 *Mathematical Population Genetics*. Springer Verlag, New York.

EWENS, W. J. and M. W. FELDMAN, 1976 The theoretical assessment of selective neutrality. pp. 303–338. In: *Population Genetics and Ecology*. Edited by S. KARLIN and E. NEVO. Academic Press, New York.

GIBSON, J., 1970 Enzyme flexibility in *Drosophila melanogaster*. Nature **227**: 959–960.

HARPER, R. A. and F. B. ARMSTRONG, 1973 Alkaline phosphatase of *Drosophila melanogaster*. II. Biochemical comparison among four allelic forms. Biochem. Genet. **10**: 29–38.

HARRIS, H., 1966 Enzyme polymorphisms in man. Proc. Roy. Soc., Ser. B **164**: 298–310.

HARTL, D. L., 1980 *Principles of Population Genetics*. Sinauer, Sunderland, Mass.

HICKEY, D. A., 1978 Selection for amylase allozymes in *Drosophila melanogaster*. Evolution **31**: 800–804.

DE JONG, G. and W. SCHARLOO, 1976 Environmental determination of selective significance or neutrality of amylase variants in *Drosophila melanogaster*. Genetics **84**: 77–94.

KIMURA, M., 1968 Evolutionary rate at the molecular level. Nature **217**: 624–626. ——, 1979 The neutral theory of molecular evolution. Scientific American **241**: 98–126.

KING, J. L. and T. H. JUKES, 1969 Non-Darwinian evolution: Random fixation of selectively neutral mutations. Science **164**: 788–798.

KOEHN, R. K., 1978 Physiology and biochemistry of enzyme variation: The interface of ecology and population genetics pp. 51–72. In: *Ecological Genetics: The Interface*. Edited by P. BRUSSARD. Springer Verlag, New York.

KUBITSCHEK, H. E., 1970 *Introduction to Research with Continuous Cultures*. Prentice-Hall, Englewood Cliffs, New Jersey.

LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

LEWONTIN, R. C. and J. L. HUBBY, 1966 A molecular approach to the study of genetic heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics **54**: 595–609.

MILKMAN, R., 1973 Electrophoretic variation in *Escherichia coli* from natural sources. Science **182**: 1024–1026.

MILLER, J. H., 1972 *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

MOSER, H., 1958 *The Dynamics of Bacterial Populations Maintained in the Chemostat*. Carnegie Inst., Publ. No. 614. Washington, D. C.

NEI, M., 1975 *Molecular Population Genetics and Evolution*. American Elsevier, New York.

PROUT, T., 1965 The estimation of fitness from genetic frequencies. Evolution **19**: 546–551. ——, 1969 The estimation of fitness from population data. Genetics **63**: 949–967. ——, 1971a The relation between fitness components and population prediction in *Drosophila*. I. The estimation of fitness components. Genetics **68**: 127–149. ——, 1971b The relation between fitness components and population prediction in *Drosophila*. II. Population prediction. Genetics **68**: 151–167.

ROGERSON, A. C. and M. FREUNDLICH, 1970 Control of isoleucine, valine and leucine biosynthesis. VIII. Mechanism of growth inhibition by leucine in relaxed and stringent strains of *Escherichia coli K-12*. Biochim. Biophys. Acta **208**: 87–98.

SCHARLOO, W. and G. DE JONG, 1980 Selection for amylase allozymes in *Drosophila melanogaster*: A reply. Evolution **34**: 608–610.

SELANDER, R. K. and B. R. LEVIN, 1980 Genetic diversity and structure in populations of *Escherichia coli*. Science **210**: 545–547.

DE SILVA, A. O. and D. G. FRANKEL, 1979 The 6-phosphogluconate dehydrogenase reaction in *Escherichia coli*. J. Biol. Chem. **254**: 10237–10242.

SOMERO, G. N., 1978 Temperature adaptation of enzymes. Ann. Rev. Ecol. Syst. **9**: 1–30.

WATERSON, G. A., 1978 The homozygosity test of neutrality. Genetics **88**: 405–417.

YAMAZAKI, T., 1971 Measurement of fitness at the esterase-5 locus in *Drosophila pseudoobscura*. Genetics **67**: 579–603.

YARDLEY, D. G., 1978 Selection at the amylase locus of *Drosophila melanogaster*: A word of caution. Evolution **32**: 920–921.

Corresponding editor: M. NEI