IN VIVO FUNCTION OF RARE *G6pd* VARIANTS FROM NATURAL POPULATIONS OF *DROSOPHILA MELANOGASTER*

WALTER F. EANES AND JODY HEY

Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794

Manuscript received December 28, 1985
Accepted March 21, 1986

ABSTRACT

From 1981 to 1983, 15,097 *X*-chromosomes were genetically extracted from a number of North American populations of *D. melanogaster* and were electrophoretically screened for rare mobility and activity variants of glucose-6-phosphate dehydrogenase (G6PD). Overall, 13 rare variants were recovered for a frequency of about 10^{-3}. Eleven variants affect electrophoretic mobility and are apparently structural, and two variants exhibit low G6PD activity. One low activity variant is closely associated with a *P*-element insertion at 18D12-13—all of the variants were subjected to the previously described genetic scheme used to identify relative *in vivo* activity differences between the two common electrophoretic variants associated with the global polymorphism. Most of the rare variants exhibit apparent *in vivo* activities that are similar to one or the other of the common variants, and these specific rare variants appear to be geographically widespread. Several variants have significantly reduced function. All of the variants were measured for larval specific activity for G6PD as a first measure of *in vitro* activity. It appears that specific activity alone is not a sufficient predictor for *G6PD in vivo* function.

EXPERIMENTAL investigations of naturally occurring enzyme polymorphism have followed a progression from almost exclusive interest in the study of fitness to studies of proximal phenotypes, such as *in vivo* flux or immediate physiological differences associated with enzyme function. This progression is understandable because the measurement of fitness differences has remained an elusive problem, complicated by the need to uncouple the fitness variation associated with closely linked loci (see EANES *et al.* 1985; KIMURA and OHTA 1971), as well as by the recognition that most biologically significant selection differentials will be impossible to detect with any statistical power (LEWONTIN 1974). The possible exceptions to this limitation are those involving chemostat studies in *E. coli* (DYKHUIZEN and HARTL 1980; HARTL and DYKHUIZEN 1981; DYKHUIZEN, DE FRAMOND and HARTL 1984), where selective environments can be varied and exceptionally small selective differentials may be measured with sufficient statistical power.

In contrast, the recent interest in proximal phenotypes associated with spe-

cific enzyme polymorphisms in eukaryotes has been informative in a number of cases (see LEIGH-BROWN 1977; HILBISH, DEATON and KOEHN 1982; HILBISH and KOEHN 1985; BURTON and FELDMAN 1983; DIMICHELE and POWERS 1982a,b; CAVENER and CLEGG 1981). The experimental strength of this approach arises from the increased ability to partition single-locus variance contributions associated with variation at more proximal levels. If individual polymorphisms contribute to fitness variation, the problem can be viewed as a chain or path of causal effects leading from proximal to distal phenotypes. Specific polymorphisms will explain less variation at each level as more loci contribute to more distal levels. Hence, a specific polymorphism could explain all of the variation in flux associated with a step in its pathway, a significant, but smaller, amount of the physiological variation associated with the pathway, yet an experimentally immeasurable, albeit biologically significant, contribution to total fitness variation. For this reason the study of proximal phenotypes will remain a constructive approach to understanding the adaptive nature of enzyme polymorphism.

The natural polymorphism for glucose-6-phosphate dehydrogenase (G6PD; EC 1.1.1.49) has received increasing attention in recent years as a model polymorphism in experimental investigations. This sex-linked polymorphism at the *G6PD* (*Zwishenferment*) locus (YOUNG, PORTER and CHILDS 1964) consists of dimeric and tetrameric variants, hence it is a polymorphism for quaternary structure, at least as is determined *in vitro* (STEELE, YOUNG and CHILDS 1968; LEE, LANGLEY and BURKHART 1978; HORI and TANDA 1980; WILLIAMSON and BENTLEY 1983). We have shown that the dimeric or fast electrophoretic allele, $G6pd^A$, varies in frequency from nearly 100% in some European populations to only rare frequencies in southern and equatorial populations, and in a reciprocating fashion in Northern and Southern hemispheres (OAKESHOTT *et al.* 1983). Sequential electrophoresis indicates that most probably only these two molecular variants are present in any appreciable frequencies on a global basis and that little hidden common variation at the protein level exists (EANES 1983). Biochemical characterization of both variants has shown repeatable differences in specific activities, thermal stabilities and K_m values for glucose-6-phosphate (STEELE, YOUNG and CHILDS 1968; BIJLSMA and VAN DER MEULEN-BRUIJNS 1979; HORI and TANDA 1980, 1981; WILLIAMSON and BENTLEY 1983; W. F. EANES, unpublished observations). The fast allele is generally described as having lower activity than the slow or tetrameric form of the enzyme, as being less stable and as possessing lower apparent K_m values for glucose-6-phosphate. There is also considerable evidence that autosomal loci modify activity levels at the G6PD and 6PGD loci (LAURIE-AHLBERG *et al.* 1980; MIYASHITA and LAURIE-AHLBERG 1984).

In EANES (1984) the *in vivo* function of the two common *G6pd* genotypes was investigated using a simple genetic scheme that exploits the fact that the viability of flies bearing low activity 6-phosphogluconate dehydrogenase (6PGD) mutations critically depends on the activity of *G6pd* variants preceding it in the pentose shunt (HUGHES and LUCCHESI 1977; GVOZDEV *et al.* 1976, 1977). By using controlled chromosomal constructions it was possible to probe

the *in vivo* activity of the two common *G6pd* variants associated with the global polymorphism using a "leaky" *6Pgd* allele. In a full activity *6Pgd* background, both *G6pd* genotypes have effectively the same or very similar viabilities within the limits of the statistical power to detect such differences (EANES 1984; EANES *et al.* 1985). However, in the presence of the low activity $6Pgd^{lo1}$ allele, the putative high activity $G6pd^{B}$ allele is semilethal, whereas the lower activity $G6pd^{A}$ variant is almost normal in viability. The existence of a causal mechanism (6-phosphogluconate accumulation) and the direction of the outcome make a compelling case for *in vivo* functional differences between the common genotypes at the *G6pd* locus.

In this report, we extend the approach and results of the earlier study to a collection of 13 rare *G6pd* variants recovered from a large-scale electrophoretic screening of over 15,000 *X*-chromosomes from natural populations. We are interested in examining the spectrum of *in vivo* functional associated with variants of *G6pd* recovered from natural populations in considering the following question: Are rare variants in general defective, or do they often resemble the common alleles in function?

MATERIALS AND METHODS

Screening of natural populations: The 15,097 *X*-chromosomes recovered from natural populations were genetically extracted by crossing wild males or F_1 male progeny from isofemale lines (one per line) with *C(1)DX, yf* attached-*X* females. *G6pd* was screened electrophoretically on 12% starch gels (Sigma) designed to carry 80 samples per gel. The first 12,000 chromosomes were screened on a 0.05 M Tris-HCl, pH 8.9, buffer (YOUNG, PORTER and CHILDS 1964), and the final 3000 lines were screened with both the former system and a 0.01 M Tris-citrate, pH 7.0, buffer. All suggestive mobility or activity variation was repeated with controls. Male flies were used in all screening. After electrophoresis for 3 hr, the gels were sliced and stained using a 1% agar overlay containing 120 mg of glucose-6-phosphate (sodium salt), 45 mg of NADP, 40 mg of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, and 5 mg of phenazine methosulfate in 100 ml of 0.05 M Tris-HCl, pH 8.9, buffer. All electrophoretic variants of G6PD can be classified as dimeric or tetrametric in quaternary structure by their mobility relative to either common allele. Rare allele designations therefore reflect presumed quaternary structure (*A* or *B*), followed by mobility relative to either common allele (*F* (fast) or *S* (slow)) and the copy number.

All recovered rare variants possessing similar mobilities on the above systems (*AS1, AS2, AS3, AS4, AS5* and *AF1, AF2, AF3, AF4*) were examined in pairwise fashion in two additional buffer systems on 5 and 7.5% acrylamide slab gels as described in EANES (1983).

Specific activity measurements: Assays for G6PD were carried out on third instar larvae and adult flies separated by sex. In the former assays, ten third-instar larvae were taken from each of four replicate vials for each variant and weighed. Samples were ground in 1.5-ml Eppendorf tubes with 0.5 ml of 0.1 I Na-MOPS [3-(*N*-morpholino)propanesulfonic acid] buffer, pH 7.15, containing 0.15 mM NADP (determined enzymatically). Samples were centrifuged at 12,400 × *g* for 5 min, and the supernatant was assayed immediately. The assays were carried out at 25° in a Gilford 2500 recording spectrophotometer with temperature control.

The assay contained 50 μl of supernatant and 450 μl of assay solution with final concentrations of 2.5 mM glucose-6-phosphate and 0.10 mM NADP in the 0.1 I Na-MOPS, pH 7.15, buffer. Specific activity is expressed as units (micromoles of NADP reduced, min^{-1}) per milligram of wet weight.

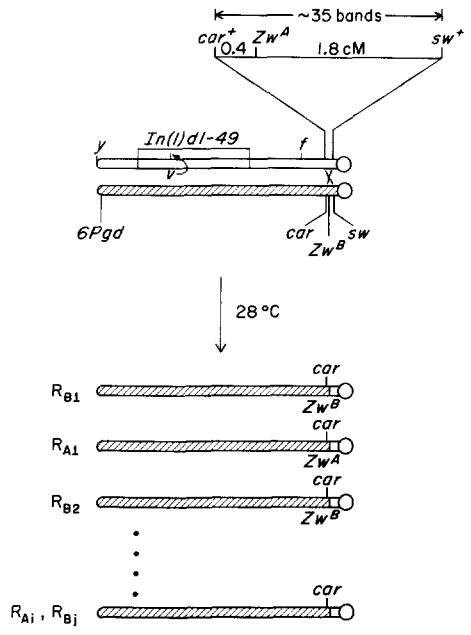

FIGURE 1.—The genetic scheme used to create an individual set of recombinant test chromosomes that are identical for the distal portion of chromosome *car sw* and the proximal portion of *In(1)dl-49, y v f*. The position of *G6pd* (=*Zw*) relative to the flanking markers is shown. Each rare allele set is established from this pair of stem chromosomes, in which the Zw^A allele has been replaced by specific rare variants. The total experiment necessitated the generation of 14 such sets (see text).

Adult assays were carried out in a similar design, except that the flies were aged for 8 days posteclosion and the sexes were assayed separately.

Recombinant test chromosome sets: For each rare variant, the indirect analysis of *in vivo* function involves two stages. First, a set of recombinant test chromosomes is generated bearing both the specific *G6pd* rare and $G6pd^B$ alleles. This step exploits recombination between visible markers closely flanking the *G6pd* locus using pairs of specifically constructed "stem" chromosomes. Using this approach we control for allelic viability variation potentially cosegregating with the *G6pd* allele, by homogenizing most of the *X*-linked background (the autosomes randomize). Second, these individual recombinant chromosomes are entered into a viability measurement scheme where *whole* chromosome viabilities are estimated relative to a reference genotype; in this case *FM6* males, where *FM6* is a balancer chromosome (LINDSLEY and GRELL 1968).

To generate the individual recombinant chromosome sets, each rare variant was introduced into the genetic scheme outlined in detail in EANES (1984) and EANES *et al.* (1985) and illustrated in Figure 1. Initially, this involves placing each rare *G6pd* variant into a stem chromosome bearing the *In(1)dl-49, y v f* inversion (to restrict crossing over in the distal region), then permitting it to recombine with the $6Pgd^{lo1}$ *car* $G6pd^B$ *sw* stem chromosome and recovering all *car* males at 28° (*sw* is a temperature-sensitive lethal) using a w^+Y translocation to mask the viability effects of the $6Pgd^{lo1}$ mutation. This "leaky" variant possesses about 10% normal activity (BEWLEY and LUCCHESI 1975). Individual recombinant males (with w^+Y) were mated with $C(1)DX/w^+Y$ attached-*X* females, and the F_1 males were screened by electrophoresis (see above) for their *G6pd*

genotype. A minimum of 50 recombinant chromosomes were recovered for each set. Chromosomes with the appropriate genotypes were then balanced over the *FM6* balancer chromosome.

For each set of such chromosomes, recombinant between *car* and *sw*, the recovered *X* chromosomes are at a minimum identical-by-descent for 97% of the chromosome (average = 98.3%), including both the large distal and small proximal regions. Within each set, recovered chromosomes bear either the defined rare *G6PD* allele or the *B* allele, depending on whether crossing over occurred to the left or right of the *G6PD* locus, yet between the flanking markers. All chromosomes bear the $6Pgd^{lo1}$ allele. There are, therefore, 14 recombinant sets, one for each rare variant and the *A* allele. It must be emphasized that *all* sets contain recombinant chromosomes bearing the same *B* allele. These serve as internal controls. This is important because the variance in "unstandardized" viability (relative to *FM6* males) among *different* sets reflects not only the effects of the *G6PD* locus in its interaction with the $6Pgd^{lo1}$ allele but also a potential contribution due to viability variation segregating with the small proximal region (from *G6pd* to the centromere) containing about 9% of the total genetic map. However, because each set also contains *B*-bearing chromosomes that differ for the same proximal regions, the between-set variance associated with rare-bearing chromosomes when contrasted with the variance of the *B*-bearing control sets allows a partitioning of the among-set viability variance into a contribution from the *G6pd* locus and a contribution by the small proximal region.

Viability measurement: For each individual recombinant set, the experimental design involves a hierarchical nesting of levels. Viability variation is measured among *G6pd* variant genotypes, and the "error" variance includes a contribution due to different recoveries of recombinant chromosomes bearing the same *G6PD* allele, as well as environmental and binomial error (EANES 1984). If we designate individual recombinant chromosomes in the *i*th set (where $i = 1,2,3\ldots14$) as R_{ij} and B_{ik} (j and k stand for the *i*th and *k*th recombinant chromosomes in the *i*th set) for the rare and *B*-bearing chromosomes, respectively, the standard design can be outlined as follows. For the *i*th set, five crosses involving individual rare recombinant chromosomes ($j = 1,2,3\ldots5$) and five crosses involving individual *B*-bearing chromosomes were set up ($k = 1,2\ldots5$). Each individual cross involved a mass mating of eight virgin $FM6/R_{ij}$ or $FM6/B_{ik}$ females with the same number of *FM6/Y* males. Each such mass cross was then replicated five times, with replicates placed in eight drawn vials. Replicates were allowed to oviposit for 4 days with transfers to fresh food on days 4, 8 and 12, after which the adults were discarded. Hence, the viability measurement for the *i*th set involved a total of $2 \times 5 \times 5 \times 3 = 150$ vials. Adults emerging from individual vials were counted on days 11, 14 and 20, and the genotype counts [*FM6/Y*; R_{ij}/Y (or B_{ik}/Y); $FM6/R_{ij}$ (or $FM6/B_{ik}$); and *FM6/FM6*] were pooled across the vial transfers for each replicate cross. Each replicate cross served as an individual viability observation in the estimation of average viability.

In this study we used a viability index formulated to further standardize the hemizygous viability depression associated with each *G6pd* rare allele-bearing chromosome by the viability depression associated with the paired *B* allele-bearing chromosomes generated as controls. The viability of each chromosome was first measured relative to the *FM6* balancer chromosome. This was computed as the frequency of tested genotype males emerging among the total number of males in a vial (including *FM6* males). This was divided by 0.58 (the average relative viability of wild-type chromosomes) and was subtracted from one to characterize each rare-bearing chromosome as the proportion of viability depression from wild-type viability. Finally, this value was standardized by dividing by the same estimated depression for the *B* control chromosomes generated with in each recombinant rare allele set. This is expected to standardize results from experiment to experiment. The resulting index extends from a value of $v_i = 0$ for *G6pd* alleles that fully suppress $6Pgd^{lo1}$ lethality, and therefore possess wild-type viability, to values near $v_i = 1.0$ for semilethal alleles that are, as is the *B* allele, poor suppressors. We expect *in vivo* activity to be positively correlated with this measure.

TABLE 1

List of collections, common *G6pd* allele frequencies and rare variants

Locality	Date	$G6pd^A$	$G6pd^B$	Total	p^A	Rares[a]
Blacksburg, VA	7/81	64	497	561	0.11	
Maybrook, VA	7/81	18	252	270	0.07	
Stony Brook, NY	8–9/81	191	822	1013	0.19	
Setauket, NY	8–9/81	82	484	566	0.14	
St. James, NY	8/81	206	987	1193	0.17	*AF3*
Mt. Sinai, NY	9–10/81	526	2764	3290	0.16	*AF2*
Mt. Sinai, NY	9–10/82	223	1152	1375	0.16	
Freeport, ME	8/82	13	188	201	0.06	
Lincoln, MA	10/82	629	1999	2623	0.24	*AF4*
Nashville, TN	11/81	58	370	428	0.14	$Nash^-$
Death Valley, CA	3/82	357	634	991	0.36	
Watsonville, CA	1/83	401	462	863	0.46	*BF1,Bl01*
Alexander Springs, FL	4/83	35	306	342	0.10	*AS5*
Venus, FL	4/83	32	340	372	0.09	
Orchid, FL	4/83	31	357	390	0.08	*AS3,AS4*
Oklawaha, FL	4/83	43	337	380	0.11	
Ft. Pierce, FL	4/83	5	72	79	0.06	*AS1,AS2*
Fruit and Spice, FL	4/82	16	144	160	0.10	
Total				15,097		

[a] Variants *AF1* from Villeurbanne, France, and *BS1* from Somerville, Massachusetts, were recovered earlier (EANES 1983) in smaller surveys of allele frequency variation.

All viability experiments were carried out at 25° in 8-dram vials in 10 ml of media containing 1000 ml of water, 19 g of agar, 32 g of cornmeal, 54 g of D-glucose, 32 g of dead yeast and 7 ml of propionic acid.

RESULTS

Rare *G6pd* variants: From 1981 to 1983 over 15,000 *X*-chromosomes recovered from natural populations in North America were screened for rare electrophoretic and activity variants. In all, 11 individual variants were recovered, nine as mobility variants and two as activity variants. The collections are summarized in Table 1, which includes the common allele frequencies as well as the list of rare variants. The frequency of individual variant copies was 9.3×10^{-4}. It would appear that, other than the two globally polymorphic variants, additional allelic types are exceptionally rare in North America. Unfortunately, there are no other systematic studies of this order of magnitude for single gene-enzymes in Drosophila; therefore, it is difficult to make contrasts. In the treatments that follow, two other variants, *AF1* and *BS1*, recovered earlier from Villeurbanne, France, and Somerville, Massachusetts (EANES 1983), are included in the comparisons.

The molecular identity-by-descent of these rare variants is impossible to establish unequivocally at this time. We have compared alleles *AF1* through *AF4*, and *AS1* through *AS5*, under different electrophoretic conditions of pH,

TABLE 2

Mean-specific activities (units per milligram of wet $wt^{-1} \times 10^{-3}$) and standard errors of larvae and adults (male and female) with the 15 alleles described in the text

Allele	Larvae	Adults	
		Males	Females
A	3.70 ± 0.02	6.58 ± 0.49	3.95 ± 0.69
B	3.72 ± 0.08	7.54 ± 0.71	5.32 ± 0.08
AF1	2.51 ± 0.32	5.70 ± 0.42	3.72 ± 0.19
AF2	4.07 ± 0.30	7.79 ± 0.24	4.35 ± 0.44
AF3	3.52 ± 0.10	6.73 ± 0.26	5.02 ± 0.62
AF4	3.78 ± 0.24	7.84 ± 0.16	3.08 ± 0.09
AS1	3.58 ± 0.18	8.21 ± 0.31	3.98 ± 0.29
AS2	4.22 ± 0.09	6.45 ± 0.16	3.68 ± 0.23
AS3	3.22 ± 0.20	7.33 ± 0.62	3.62 ± 0.16
AS4	4.56 ± 0.12	9.18 ± 0.38	5.90 ± 0.07
AS5	3.91 ± 0.30	7.48 ± 0.33	4.24 ± 0.18
BF1	3.98 ± 0.22	6.42 ± 0.23	3.73 ± 0.30
BS1	2.13 ± 0.18	3.00 ± 0.07	2.31 ± 0.29
Blo1	0.80 ± 0.04	2.75 ± 0.06	1.21 ± 0.16
Nash$^-$	0.09 ± 0.01[a]	0.08 ± 0.02[a]	0.08 ± 0.01[a]

[a] Not significantly different from background control.

ionic strength and concentrations of acrylamide and are able to remove *AF1* and *AS2* from those sets of variants. Sequential electrophoresis of human rare hemoglobin variants has proved very effective in separating human variants that have previously been assigned to general charge classes (RAMSHAW, COYNE and LEWONTIN 1979). Therefore, we tentatively believe that these two homogeneous mobility classes represent multiple recoveries of the same two variants, one rare variant possibly restricted to Long Island and Massachusetts collections and the other to Florida. If so, copies of these rare variants have still been recovered hundreds of kilometers apart. The question of molecular identity is not technically untractable. The *G6pd* locus has been cloned (GANGULY, GANGULY and MANNING 1985), and we are in the processes of restriction mapping and hope also to sequence these alleles, as has been done with *Adh* and its flanking regions (KRIETMAN 1983; C. F. AQUADRO *et al.*, unpublished results).

The 13 rare and two common *G6PD* variants span a range of specific activities (Table 2), which are highly correlated in the third instar larvae and adult (sexes pooled) stages ($r = 0.95$, $P < 0.01$). Both common variants appear to have the same specific activity. One variant, *Nash*$^-$, has zero activity over background and thus appears completely null in activity. The *Blo1* variant has identical mobility to *B* under all conditions, but has only 20% of "normal" *B* activity.

These two low activity variants were examined cytologically and were also screened with a *P*-element probe ($p\pi 25.1$; O'HARE and RUBIN 1983) by *in situ* hybridization to the polytene chromosomes on the chance that these mutations might be dysgenically derived as proposed for a number of autosomal nulls

(BURKHART *et al.* 1984). The *Blo1* chromosome shows a distinct *P*-element insertion at 18D12-13 that, given the cytogenetic mapping of the locus to the 18D-F region (STEWART and MERRIAM 1974; WILLIAMSON and BENTLEY 1983), the genetic mapping using *car* at 18D1-2 and *sw* at 19B (EANES 1983) and the *in situ* hybridization of the cloned locus (GANGULY, GANGULY and MANNING 1985), is in the site anticipated for the locus. This insertion segregated without exception with the *Blo1* phenotype in the nine chromosomes recombinant between the *car* and *sw* closely flanking visible markers. We have also direct evidence, from the restriction mapping of this variant, for a 2.2-kb insertion within 1.5 kb of the 5′ end of the *G6pd* coding region. This variant, if a *P*-element insertion, did not arise spontaneously during chromosome extraction, because the iso-*X* line so derived would have been heterogeneous.

***In vivo* function:** In order to compare the *in vivo* function of the rare *G6pd* variants, each was introduced into the same genetic protocol outlined in EANES (1984) and was assessed for the ability to suppress the male hemizygous lethal effects of the "leaky" $6Pgd^{lo1}$ allele. We must emphasize that the viability measured here should not be construed as an indication of the viability or fitness (partial) of these variants in natural populations. It is used in this artificial construct to infer *in vivo* function and should be interpreted in no other way. The hemizygous viability index values for 15 *G6pd* alleles are summarized on the ordinant of Figure 2. The index ranges from zero viability depression for full suppression (wild-type viability) to values greater than one for high activity variants that are even poorer suppressors than their *B* controls, which are generally semilethal. The best suppressor, as anticipated, is the null variant, $Nash^{-}$, and the worst is the nearly lethal *AS3*. It should be noted immediately that the basic observation from EANES (1984) has been repeated; the common polymorphic *A* and *B* alleles suppress very differently. The second important observation is that the majority of rare variants show a pattern of essentially "wild-type" *in vivo* function with a number of variants associated with each common variant. One variant, *AS2*, has intermediate function, whereas another, *AS3*, appears superactive. Four alleles have activity significantly below wild-type function (*A* or *B*). It should be pointed out that the two homogeneous electrophoretic mobility sets observed above are not identical with respect to this measure of function. This may either indicate their real nonidentity or represent regulatory divergence accrued since the unique mutational origin of each set. It is unlikely that different autosomal backgrounds contribute to these differences, because in the construction of the final lines, each allele has gone through six crosses to the same laboratory stocks.

In this study, we have assumed that contributions by the proximal *X*-chromosome regions as well as by autosomal modifiers segregating among lines contribute negligibly to the large among-rare-set variance in viability. However, we may examine this assumption and quantify the contribution of the proximal regions and autosomes by computing the among-set variance components for the rare-bearing chromosomes and their *B*-bearing control chromosome sets separately. Although the variance in viability among rare sets contains variance contributions by the *G6pd* locus, as well as their associated proximal regions

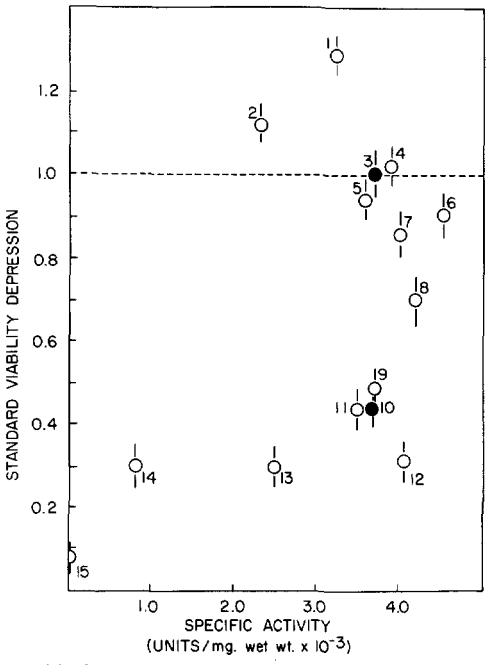

FIGURE 2.—The relationship between the standardized viability depression (see text) for males carrying the $6Pgd^{lo1}$ allele and each *G6pd* variant, and the measured larval-specific activities. The common alleles (●) are $A = 10$, $B = 3$. The rare variants (○) are $AS3 = 1$, $BS1 = 2$, $AS5 = 4$, $AS4 = 5$, $AS1 = 6$, $BF1 = 7$, $AS2 = 8$, $AF4 = 9$, $AF3 = 11$, $AF2 = 12$, $AF1 = 13$, $Blo1 = 14$, and $Nash^{-} = 15$.

and the segregation of potential autosomal modifiers, the variance among sets for the *B*-control chromosomes contains only contributions due to the latter two components; hence, no contribution of *G6pd* to the among-set variance. The difference in estimated among-set variance components represents that additional contribution by the rare variants at the *G6pd* locus and their individual interactions with the $6Pgd^{lo1}$ allele. The among-set variance (on arcsine transformed relative viabilities) contributions to viability are $V_{rare} = 191.97$ and $V_B = 21.90$ for the rare and *B* chromosome subsets, respectively. This nearly ninefold difference in variances is, statistically, highly significant ($F_s = 8.77$; $P < 0.001$). The greatly increased variance associated with the rare subset indicates that the majority of the variance in viability observed between rare variants in this system is associated with the variants *per se*, not the heterogeneity in proximal regions or autosomal backgrounds.

In this analysis, we have not examined the viability variation of these rare variants in a "normal" *6Pgd* background. In EANES (1984) the viabilities of the *A* and *B* alleles were also examined in a normal $6Pgd^{S}$ background because there was no *a priori* reason to suspect there were no large *intrinsic* viability differences between genotypes. This was another level of control. However, the genotype viabilities were virtually identical compared to the relatively large viability differences observed in the $6Pgd^{lo1}$ background. We have assumed

here that there are relatively insignificant intrinsic viability differences among rare variants and that the very large variation in viabilities observed is due to their interaction in each case with the $6Pgd^{lo1}$ allele. This is not unreasonable. Superficially, there were no obvious differences in fitness associated with these alleles before placement in the $6Pgd^{lo1}$ background. In fact, the $Nash^{-}$ null allele, the most likely candidate for a variant with intrinsically lowered viability, exhibits completely wild-type hemizygous viability relative to *FM6* males.

DISCUSSION

This is the first study designed to investigate the functional activities of rare variants from natural populations of *Drosophila melanogaster*. We expect spontaneous mutation at the *G6pd* locus to generate new variation with a range of *in vivo* function, and while most new mutations are lost stochastically, natural selection will also screen this functional array, eventually determining the number of generations that many mutations persist. Because of hemizygosity, defective alleles at *X*-linked loci are rapidly eliminated, whereas alleles with function equivalent to either common allele should be effectively neutral with respect to the fitness of that common counterpart. These quasineutral variants, although globally rare, could assume widespread distributions, depending on migration (see SLATKIN 1981). Understanding the particular allele frequency spectrum and geographic distribution of rare variants at the *G6pd* locus requires knowledge of the mutation rates to alleles of different functional (fitness) classes, as well as elusive information on the history and population structure of *D. melanogaster* in North America. As observations, we possess only the collection of rare variants, each with an unknown history. Several of these variants appear geographically widespread, yet others must have ambiguous distributions because of the vagaries of sampling rare alleles *per se*. Recognizing these constraints, what can we say about the functional nature of rare *G6pd* alleles recovered from natural populations?

We have recovered only one *bona fide* null variant at the *G6pd* locus in the more than 15,000 gene copies screened for a frequency of 6.67×10^{-5}. This result is consistent with two other studies of *X*-linked null alleles (LANGLEY *et al.* 1981; VOELKER *et al.* 1980) which found no *X*-linked nulls at four loci in 1600 chromosomes, whereas autosomal nulls were regularly recovered (average frequency, 2.45×10^{-3}). Studies of activity and electrophoretic mobility mutation have suggested that the mutation rate to null alleles is equal to or as much as an order of magnitude greater than that for electrophoretic mobility (MUKAI and COCKERHAM 1977; VOELKER, SCHAEFFER and MUKAI 1980). It would therefore appear that selection is rapidly eliminating null alleles at this locus.

Most of the *G6pd* variants examined here appear to have *in vivo* function similar to one or the other common variant. We speculate that seven of these copies belong to two variant sets (*AF2, AF3, AF4,* and *AS1, AS3, AS4, AS5*), with individual copies collected often hundreds of kilometers apart. This implies that the regional population contains thousands of copies of these variants and that they represent examples of quasineutral variants. Of course, we can-

not rule out that these alleles possess different, perhaps detrimental, function under other conditions and may be selectively maintained at low equilibrium frequencies over large areas. The analysis of DNA sequence diversity associated with these variants will shed considerable light on their identity and history.

There are aspects of *in vivo* function still needing refinement. The measure used here is an indirect reflection of G6PD flux. The mechanism presumedly involves the accumulation of 6-phosphogluconate preceding 6PGD, which, in turn, inhibits phosphoglucose isomerase (PGI) activity (SMIRNOVA, KOGAN and GVOZDEV 1982), leading to lowered viability (PGI nulls are apparently lethals; BURKHART *et al.* 1984). The functional shape of this response is unknown; hence, flux is not directly inferable except as a decreasing monotonic function of viability. However, we may speculate on several points. If *Blo1* is a *P*-element-mediated transcriptional mutation as proposed, then we may say that an 80% reduction in *in vivo* enzyme (V_{max}) concentration and flux results in a 70% suppression relative to the full activity *B* allele (Figure 2). For a relationship that is approximately linear from *B* to *Blo1*, we may expect the 60% viability suppression associated with the *A* variant to reflect a reduction in flux of about 50%. We are currently planning *in vivo* flux studies using radiolabeled D-[6^{14}C] and D-[1-^{14}C]glucose, axenic culture and controlled autosomal backgrounds. By examining several genotypes (*Nash*$^{-}$, *A*, *B*) representing the range of activities seen here, we shall be able to map *6Pgd*^{lo1} suppression with "real" flux and, thus, infer *in vivo* function for the entire set of *G6pd* variants.

We have examined specific activity here as an initial step in the full characterization of the *G6pd* common and rare variants. Several variants with lowered activities have clearly lowered *in vivo* activity (*Nash*$^{-}$, *Bl01*, *AF1*); however, a general observation is that specific activity alone is inadequate as a measure of *in vivo* function. This is not unexpected. First, due to *in vitro* stability differences, the measure simply may not reflect true *in vivo* V_{max}. Second, and probably the most important point, the intracellular substrate levels are not saturating but are near the K_m for each substrate (FERSHT 1977). For this reason, the examination of *in vivo* function should require a full kinetic characterization under physiologically relevant *in vitro* conditions. We found no difference in larval-specific activity between the two common variants when assayed under our buffer conditions. Most studies have reported lower activity for the *A* variant (STEELE, YOUNG and CHILDS 1968; BIJLSMA and VAN DER MEULEN-BRUIJNS 1979; HORI and TANDA 1981; WILLIAMSON and BENTLEY 1983), but all have used Tris-HCl buffers at pH values 7.5 to 8.6. CAVENER and CLEGG (1981) demonstrated that there were *in vitro* stability differences in this buffer and found no differences in larvae when assayed in 0.05 M phosphate, pH 7.5, buffer where stability was conferred. Our choice of a physiological buffer (SOMERO 1981) with pH in the range of reported intracellular pH values (ROOS and BORON 1981) has concluded that there are no specific activity differences between *A* and *B* in larvae, although there is the suggestion of differences in adults. However, as the viability measure reported here reflects prepupal mortality, we have concentrated on larval activity differences. Presumedly, the *in vivo* activity differences in larvae arise from dif-

ferences in other parameters, such as the K_m values for glucose-6-phosphate and NADP or the K_i for NADPH. We are currently purifying to homogeneity a set of these rare variants, which, along with the two common variants, will be subjected to a full kinetic characterization.

The final issue we wish to address concerns the relationship of this study to the KACSER and BURNS (1981) model for the evolution of genetic dominance. This model proposes that enzymatic flux is a systematic property of all the steps through a biochemical pathway and that overall flux will be insensitive to activity variation at any step unless that variation is drastic. It appears their hypothesis serves as a robust explanation for the recessive nature of most discrete Mendelian mutations, as well as for the additive nature of polygenic variation. As an extension of this logic, some investigators have questioned the ability of relatively subtle activity differences observed for most enzyme polymorphisms to lead to biologically significant flux differences on which selection could act (HARTL, DYKHUIZEN and DEAN 1985). Indeed, a recent study of G6PD allozymes in *E. coli* by DYKHUIZEN, DE FRAMOND and HARTL (1984) failed to detect significant growth rate differences between naturally occurring variants. This observation makes an excellent case for the selective neutrality of these alleles, and kinetic studies also failed to detect any differences in K_m or V_{max} values between enzyme variants. An analogous situation may emerge for the rare *G6pd* variants within the classes associated with each common allele in *D. melanogaster*. MIDDLETON and KACSER (1983) in an investigation of ethanol flux associated with different genotypes of the *Adh* polymorphism were unable to detect *in vivo* flux differences in spite of the presence of reported kinetic differences. Their results contrast with this study and those of CAVENER and CLEGG (1981) and EANES (1984), which all report *in vivo* differences associated with the *common* allele polymorphisms for the pentose shunt enzymes G6PD and 6PGD. Should their model remain a robust explanation for the properties associated with pathway flux, what can be learned from these apparent exceptions?

In considering this problem we may turn the issue around and, assuming their model, ask another question: Where and how might selection operate on a pathway if it becomes advantageous to modify flux (*i.e.*, change the fitness function; HARTL, DYKHUIZEN and DEAN 1985) in some environments? Obviously, selection will operate at those points best able to confer flux change with catalytic or regulatory mutation. Enzymatic steps that do not conform to the systematic properties implicit in the KACSER and BURNS (1981) model will become the targets of selective influence and polymorphism if the circumstances arise. The effective irreversibility of the G6PD step and the lack of 6-phosphogluconate inhibition of G6PD may contribute to an ability to influence flux at this juncture and the appearance of this polymorphism. On the other hand, to the extent that steps with large "sensitivity coefficients" do not exist in this or any pathway, the model also predicts that selected polymorphisms may require large differences in regulatory or catalytic function as measured *in vivo*. This may yet be the case for the G6PD and 6PGD polymorphisms.

We should like to thank D. HOULE, J. COYNE, S. ORZACK and D.MCCAULEY for assisting in providing flies from a number of sites. D. HOULE, B. BINGHAM and J. AJIOKA assisted in a variety of aspects. R. K. KOEHN critically read an earlier version of the manuscript. The work was supported by National Science Foundation grant DEB-8106930 to W. F. EANES. This is contribution 583 from the Program in Ecology and Evolution, State University of New York, Stony Brook.

LITERATURE CITED

BEWLEY, G. C. and J. C. LUCCHESI, 1975 Lethal effects of low and "null" activity alleles of 6-phosphogluconate dehydrogenase in *Drosophila melanogaster*. Genetics **79:** 451–457.

BIJLSMA, R. and C. VAN DER MEULEN-BRUIJNS, 1979 Polymorphism at the *G6PD* and *6PGD* loci in *Drosophila melanogaster*. III. Developmental and biochemical aspects. Biochem. Genet. **17:** 1131–1144.

BURKHART, B. D., E. MONTGOMERY, C. H. LANGLEY and R. A. VOELKER, 1984 Characterization of allozyme null and low activity alleles from two natural populations of *Drosophila melanogaster*. Genetics **107:** 295–306.

BURTON, R. S. and M. W. FELDMAN, 1983 Physiological effects of an allozyme polymorphism: glutamate-pyruvate transaminase and response to hyperosmotic stress in the copepod, *Tigripus californicus*. Biochem. Genet. **21:** 239–251.

CAVENER, D. R. and M. T. CLEGG, 1981 Evidence for biochemical and physiological differences between enzyme genotypes in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **78:** 4444–4447.

DIMICHELE, L. and D. A. POWERS, 1982a LDH-B genotype specific hatching times of *Fundulus heteroclitus* embryos. Nature **296:** 563–564.

DIMICHELE, L. and D. A. POWERS, 1982b Physiological basis for swimming endurance differences between LDH-B genotypes of *Fundulus heteroclitus*. Science **216:** 1014–1016.

DYKHUIZEN, D., J. DE FRAMOND and D. L. HARTL, 1984 Selective neutrality of glucose-6-phosphate dehydrogenase allozymes in *Escherichia coli*. Mol. Biol. Evol. **1:** 162–170.

DYKHUIZEN, D. and D. L. HARTL, 1980 Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. Genetics **96:** 801–817.

EANES, W. F., 1983 Genetic localization and sequential electrophoresis of G6pd in *Drosophila melanogaster*. Biochem. Genet. **21:** 703–711.

EANES, W. F., 1984 Viability interactions, *in vivo* activity and the *G6PD* polymorphism in *Drosophila melanogaster*. Genetics **106:** 95–107.

EANES, W. F., B. BINGHAM, J. HEY and D. HOULE, 1985 Targeted selection experiments and enzyme polymorphism: negative evidence for octanoate selection at the *G6PD* locus in *Drosophila melanogaster*. Genetics **109:** 379–391.

FERSHT, A., 1977 *Enzyme Structure and Mechanism*. W. H. Freeman, San Francisco.

GANGULY, R., N. GANGULY and J. E. MANNING, 1985 Isolation and characterization of the glucose-6-phosphate dehydrogenase gene of *Drosophila melanogaster*. Gene **35:** 91–101.

GVOZDEV, V. A., T. I. GERASIMOVA, T. I. KOGAN, G. L. BRASLAVSKAYA and O. YU, 1976 Role of the pentose phosphate pathway in metabolism of *Drosophila melanogaster* elucidated by mutations affecting glucose-6-phosphate and 6-phosphogluconate dehydrogenases. FEBS Lett. **64:** 85–88.

GVOZDEV, V. A., T. I. GERSIMOVA, G. L. KOGAN and J. M. ROSOVSKY, 1977 Investigations on the organization of genetic loci in *Drosophila melanogaster*: lethal mutations affecting 6-phosphoglucoante dehydrogenase and their suppression. Mol. Gen. Genet. **153:** 191–198.

HARTL, D. L. and D. E. DYKHUIZEN, 1981 Potential for selection among nearly neutral allozymes

of 6-phosphogluconate dehydrogenase in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78:** 6344–6348.

HARTL, D. L., D. E. DYKHUIZEN and A. M. DEAN, 1985 Limits of adaptation: the evolution of selective neutrality. Genetics **111:** 655–674.

HILBISH, T. J., L. E. DEATON and R. K. KOEHN, 1982 Effect of an allozyme polymorphism on regulation of cell volume. Nature **3298:** 688–689.

HILBISH, T. J. and R. K. KOEHN, 1985 Dominance in physiological phenotypes and fitness at an enzyme locus. Science **229:** 52–54.

HORI, S. H. and S. TANDA, 1980 Purification and properties of wild-type and mutant glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase from *Drosophila melanogaster*. Jpn. J. Genet. **55:** 211–223.

HORI, S. H. and S. TANDA, 1981 Genetic variation in the activities of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in *Drosophila melanogaster*: evidence for an autosomal modifier system. Jpn. J. Genet. **56:** 257–277.

HUGHES, M. B. and J. C. LUCCHESI, 1977 Genetic rescue of a lethal "null" activity allele of 6-phoshogluconate dehydrogenase in *Drosophila melanogaster*. Science **196:** 1114–1115.

KACSER, H. and J. A. BURNS, 1981 The molecular basis of dominance. Genetics **97:** 639–666.

KIMURA, M. and T. OHTA, 1971 *Theoretical Aspects of Population Genetics. Monographs in Population Biology*. Princeton University Press, Princeton.

KREITMAN, M., 1983 Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature **304:** 412–417.

LANGLEY, C. H., R. A. VOELKER, A. J. LEIGH-BROWN, S. OHNISHI, B. DICKSON and E. MONTGOMERY, 1981 Null allele frequencies at allozyme loci in natural populations of *Drosophila melanogaster*. Genetics **99:** 151–156.

LAURIE-AHLBERG, C. C., G. MARONI, G. C. BEWLEY, J. C. LUCCHESI and B. S. WEIR, 1980 Quantitative genetic variation of enzyme activities in natural populations of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **77:** 1073–1077.

LEE, C. Y., C. H. LANGLEY and J. BURKHART, 1978 Purification and molecular weight determination of glucose-6-phosphate dehydrogenase and malic enzyme from mouse and *Drosophila*. Anal. Biochem. **86:** 697–706.

LEIGH-BROWN, A. J., 1977 Physiological correlates of an enzyme polymorphism. Nature **290:** 803–804.

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

LINDSLEY, D. L., and E. H. GRELL, 1968 Genetic variations of *Drosophila melanogaster*. Carnegie Inst. Wash. Publ. 627.

MIDDLETON, R. J. and H. KACSER, 1983 Enzyme variation, metabolic flux and fitness: alcohol dehydrogenase in *Drosophila melanogaster*. Genetics **105:** 633–650.

MIYASHITA, N. and C. C. LAURIE-AHLBERG, 1984 Genetical analysis of chromosomal interaction effects on the activities of the glucose-6-phosphate and 6-phosphogluconate dehydrogenases in *Drosophila melanogaster*. Genetics **106:** 655–668.

MUKAI, T. and C. C. COCKERHAM, 1977 Spontaneous mutation rates at enzyme loci in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **74:** 2514–2517.

OAKESHOTT, J. G., G. K. CHAMBERS, J. B. GIBSON, W. F. EANES and D. A. WILLCOCKS, 1983 Geographic variation in G6pd and Pgd allele frequencies in *Drosophila melanogaster*. Heredity **50:** 67–72.

O'HARE, K. and G. M. RUBIN, 1983 Structure of *P* transposable elements of *Drosophila melanogaster* and their sites of insertion and excision. Cell. **34:** 25–35.

RAMSHAW, J. A. M., J. A. COYNE and R. C. LEWONTIN, 1979 The sensitivity of gel electrophoresis as a detector of genetic variation. Genetics **93:** 1019–1037.

ROOS, A. and W. F. BORON, 1981 Intracellular pH. Physiol. Rev. **61:** 196–434.

SLATKIN, M., 1981 Estimating levels of gene flow in natural populations. Genetics **99:** 323–335.

SMIRNOVA, S. G., G. L. KOGAN and V. A. GVOZDEV, 1982 On the mode of lethal mutations affecting 6-phosphogluconate dehydrogenase in *Drosophila melanogaster.* Drosophila Inform. Serv. **58:** 136–138.

SOMERO, G. N., 1981 pH-temperature interactions on proteins: optimal pH and buffer system design. Marine Biol. Lett. **2:** 163–178.

STEELE, M. C., W. J. YOUNG and B. CHILDS, 1968 Glucose-6-phosphate dehydrogenase in *Drosophila melanogaster:* starch gel electrophoretic variation due to molecular instability. Biochem. Genet. **2:** 159–175.

STEWART, B. R. and J. R. MERRIAM, 1974 Segmental anenploidy and enzyme activity as a method for cytogenetic localization in *Drosophila melanogaster.* Genetics **76:** 301–309.

VOELKER, R. A., C. H. LANGLEY, A. J. LEIGH-BROWN, S. OHNISHI, B. DICKSON, E. MONTGOMERY and S. C. SMITH, 1980 Enzyme null alleles in natural populations of *Drosophila melanogaster*: frequencies in a North Carolina population. Proc. Natl. Acad. Sci. USA **77:** 1091–1095.

VOELKER, R. A., H. E. SCHAEFFER and T. MUKAI, 1980 Spontaneous allozyme mutations in *Drosophila melanogaster*: rate of occurrence and nature of the mutants. Genetics **94:** 961–968.

WILLIAMSON, J. H. and M. M. BENTLEY, 1983 Comparative properties of three forms of glucose-6-phosphate dehydrogenase in *Drosophila melanogaster.* Biochem. Genet. **21:** 1153–1166.

YOUNG, W. J., J. E. PORTER and B. CHILDS, 1964 Glucose-6-phosphate dehydrogenase in Drosophila: *X*-linked electrophoretic variants. Science **143:** 140–141.

Communicating editor: D. L. HARTL