Allozyme-Associated Heterosis in *Drosophila melanogaster*

David Houle

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

Manuscript received February 23, 1989
Accepted for publication September 7, 1989

ABSTRACT

Two large experiments designed to detect allozyme-associated heterosis for growth rate in *Drosophila melanogaster* were performed. Heterosis associated with allozyme genotypes may be explained either by functional overdominance at the allozyme loci, or closely linked loci; or by genotypic correlations between allozyme loci and loci at which deleterious recessive alleles segregate. Such genotypic correlations would be favored by consanguineous mating, small effective population size, population mixing and strong natural or artificial selection. *D. melanogaster* is outbred, has large effective population size and there is little evidence for genotypic disequilibria. Therefore it would be unlikely to show allozyme heterosis due to genotypic correlations. In the first experiment I estimated the genotypic values of 97 replicated genotypes. In the second experiment, 500 individuals were raised in a fluctuating, stressful environment. In neither experiment was there any consistent evidence for allozyme heterosis in size or development rate, fluctuating asymmetry for size or in tendency to deviate from the population mean. In the first experiment, heterosis explained less than 5.6% of the genetic variance in growth characters. In the second, heterosis explained less than 0.1% of the phenotypic variance in growth characters. Outside of the molluscs, species which show allozyme heterosis have population structures or histories which tend to promote genotypic correlations. There is little evidence that functional overdominance is responsible for observations of allozyme-associated heterosis.

A number of recent studies have shown a positive correlation between heterozygosity at allozyme loci and potentially fitness-related traits (Mitchon and Grant 1984; Zouros and Foltz 1987). Allozyme-associated heterosis has been detected both for character means, *(e.g.,* more rapid growth) and for variances *(e.g.,* lower asymmetry, or lower deviation from the population mean). These phenomena have been found in a variety of organisms, although published studies which fail to find evidence for allozyme heterosis are also common (Zouros and Foltz 1987).

Heterosis was originally defined to describe the phenomenon of hybrid vigor "irrespective of mechanism" (Shull 1914), and I will employ it to describe the observation that heterozygous genotypes perform better than more homozygous genotypes. Heterosis is the reverse of inbreeding depression, poor performance of inbred individuals compared to outbred relatives (Charlesworth and Charlesworth 1987). There are two basic genetic mechanisms which can explain observations of heterosis, usually termed dominance and overdominance (Crow 1952; Wright 1977). The overdominance hypothesis holds that heterozygotes at single loci have higher values of desir-
corresponding homozygotes in any single situation. Another possible explanation for allozyme heterosis is associative overdominance due to gametic disequilibrium between a scored locus and overdominant loci (FRANKLIN and LEWONTIN 1970; KIMURA and OHTA 1971; THOMSON 1977). If such loci are closely linked, this would be very difficult to distinguish from functional overdominance at the scored locus.

Under the dominance hypothesis, allozyme heterosis would be the consequence of genotypic correlations between scored loci and loci segregating for deleterious recessive alleles. This may be due to gametic-phase disequilibrium (OHTA 1971), or to inbreeding, which generates identity disequilibrium (HALDANE 1949; WEIR and COCKERHAM 1969). Gametic disequilibria are promoted by small effective population size (HILL and ROBERTSON 1968; KIMURA and OHTA 1971); factors that retard recombination, such as chromosomal inversions (OHTA 1971); and strong directional selection (MAYNARD SMITH and HAIGH 1974; THOMSON 1977). FRYDENBERG (1965), originally defined “associative overdominance” as heterosis not due to functional overdominance at a studied locus, and subsequent authors have used the term in this way. This definition is unfortunate, as it reduces the term overdominance to a synonym for heterosis, rather than the traditional definition as heterozygote superiority at a particular locus. I will refer to heterosis due to genotypic correlations with deleterious alleles as “dominance-correlation heterosis” to distinguish it from hypotheses involving overdominance at some loci.

In spite of the consensus that dominance accounts for most heterosis in artificial crosses, the genetic basis of allozyme-associated heterosis is obscure. There is very little direct evidence in favor of functional overdominance, but it is unclear that genotypic correlations are sufficiently common to explain all cases of allozyme heterosis. Tests of the competing hypotheses are difficult because allozyme heterosis is most frequently observed in organisms where genetic manipulations are very difficult. One area in which the dominance and overdominance hypotheses make different predictions is in relation to population structure and history. Under the overdominance hypothesis, effective population size will have no effect on the expectation that heterozygotes should perform better than homozygotes. Under the dominance-correlation hypothesis, populations with little gametic disequilibrium or inbreeding should not be expected to show heterosis. Consequently, species with large panmictic population structures provide a partial test of the two hypotheses. Failure to find heterosis in such species would tend to implicate dominance-correlation as the genetic basis of the phenomenon. I report here two experiments designed to detect allozyme heterosis in Drosophila melanogaster, which seems to approach the ideal panmictic situation.

D. melanogaster is abundant, commensal and has been introduced worldwide by humans in historical times, making a high level of gene flow extremely likely. This is supported by the near absence of geographic differentiation for restriction fragment length polymorphisms within North America (KREITMAN and AGUADÉ 1986; KREITMAN 1987; SIMMONS et al. 1989). Assuming that worldwide dispersal is also common, nucleotide heterozygosity at silent sites in samples combined from different continents can be used to estimate effective population size. Such estimates are always a million or more (KREITMAN 1987; AQUADRO et al. 1986; LANGLEY et al. 1988; SCHAFFER, AQUADRO and LANGLEY 1988; MIYASHITA and LANGLEY 1988; AQUADRO, LATO and NOON 1988). There is some clinal and patchy allozyme differentiation worldwide (e.g. SINGH and RHOMBERG 1987; OAKESHOTT et al. 1982; VOELKER, COCKERHAM and JOHNSON 1978), but much of this is most readily explained by selection, rather than limitations on gene flow. Direct studies of gametic disequilibrium have found little between allozyme loci (LANGLEY 1977). Studies at the nucleotide level have found some disequilibrium within loci and their immediate flanking regions but beyond about 10 kb there is little evidence that sites with equitable frequencies are liable to be in gametic disequilibrium (AQUADRO et al. 1986; SCHAEFFER, AQUADRO and LANGLEY 1988; MIYASHITA and LANGLEY 1988). There is no evidence for consanguineous mating in the wild. D. melanogaster is unlikely to possess sufficient genotypic disequilibria to cause dominance-correlation heterosis.

Previous studies of the relationship between allozyme genotypes and viability, fecundity, and development rate in D. melanogaster have given conflicting results (MUKAI et al. 1974; SERRADILLA and AYALA 1983). I chose to study growth-related characters, following studies in which differences in metabolism during rapid growth or stress have been implicated in allozyme heterosis (KOHEN and SHUMWAY 1982; DIEHL et al. 1985; DIEHL, GAFFNEY and KOHEN 1986; DANZMANN, FERGUSON and ALLENDORF 1987; MITTON and GRANT 1984; ALLENDORF and LEARY 1986). In D. melanogaster such an energetic advantage would most likely appear during larval growth.

MATERIALS AND METHODS

Experiment I: Wild Drosophila melanogaster were obtained from Davis Peach Farm, Mt. Sinai, New York, in a single collection in September 1984. I extracted haploid autosomal genomes from these flies using the second and third chromosome balancer stock A1, Cy L Ubx/B1, bwp (WALLACE, ZOUROS and KRIMBAS 1966). A1/B1, bwp carries a pair of reciprocal translocations between the second and third chromosomes, so viable gametes must carry either A1,
or B18. I substituted a single X chromosome, carrying the recessive marker yellow body (y) into A1/B18 as shown in Figure 1a. The y/+ stock had previously been converted to P cytotype (W. Eanes, personal communication), so the balanced stock was probably converted as well. Detailed descriptions of the mutations may be found in Lindsley and Grell (1968). The y, A1/B18 stock was used to extract haploid sets of wild second and third chromosomes, as shown in Figure 1b. Each line carried a wild Y chromosome; the y marked X chromosome, shared with all other lines; a wild second and third chromosome, balanced over A1y, and was probably segregating for a wild and a lab fourth chromosome. The fourth chromosome is small enough, about 1% of the genome of D. melanogaster, that its effect can be ignored.

Forty-nine of the 200 lines extracted were numbered at random, and reciprocally crossed with four other lines in a regular pattern, where line i was crossed with lines i + 23, through i + 26. This yielded 98 crosses, plus reciprocals. Virgin +/+ flies were collected over a 7-day period, and matings were set up on day 9. Eight females and six to ten males were placed in the same vials for each reciprocal cross. A replicate pair of reciprocal matings was set up for 11 of the crosses on day 10.

Starting on day 5, all flies were placed on a 20 hr light:4 hr dark regime to take advantage of pulses of egg laying at incubator "sunset" and "dawn" (Ashburner and Thompson 1978). On days 10 through 13, each set of mated flies was transferred to fresh food for 4–5 hr, to yield the offspring used for the experiment, and then transferred back to holding vials until the next day. This first transfer was made about 0.5 hr before "sunset," and the second about "dawn." This concentrated egg laying during the 5-hr period when the experimental offspring were sought, without preventing egg-laying in the previous period. Each day of oviposition was treated as an experimental block for analysis.

Vials containing offspring were placed in continuous light about 0.5 hr before "sunset," and the second about "dawn." This concentrated egg laying during the 5-hr period when the experimental offspring were sought, without preventing egg-laying in the previous period. Each day of oviposition was treated as an experimental block for analysis.

The food used was a standard cornmeal, sucrose, brewer's yeast, and agar medium, with antibiotics and propionic acid to inhibit the growth of microorganisms, seeded with live yeast. Abundant live yeast was provided in vials in which mated sets of flies were held. During the period of virgin collection, flies were held at 17°C, while the rest of the experiment was carried out at 25°C. Experiment II: For this experiment, I made a single collection of wild flies from Davis Peach Farm in September 1985. The Davis Peach Farm population reaches very high density in late summer, and is mostly D. melanogaster. Flies were held for 2 days in half-pint milk bottles on the cornmeal food described above, which was heavily inoculated with live brewer's yeast. Flies were maintained 150–200 to a bottle on their normal 12:12 September light-dark cycle. One thousand females were then placed individually in vials on standard food with no live yeast. These flies were allowed to oviposit overnight, in the dark, for 12 hr, transferred back to fresh vials, where they were allowed to lay eggs another 12 hr, until just after incubator "sunset." The parents were then discarded. The extraordinarily high density of the wild population makes it unlikely that this procedure greatly alters mating structure. The amount of common paternity among sibships was probably negligible, as females remate about once a day in the continuous presence of males, and males are unlikely to mate more than twice a day (Gromko, Gilbert and Richmond 1984). Half of the vials from each oviposition period were used for this experiment. Vials were placed in boxes of 100, five from the night oviposition and five from the day oviposition period. Data were analyzed with these boxes as blocks. The parents were maintained at 25°C.

The developing flies were maintained in a temperature regime of approximately 10 hr at 28°C, and 14 hr at 18°C, mimicking a temperature range routinely experienced by larvae in the field. The higher temperature is mildly stressful for flies (Ashburner and Thompson 1978), with, for example, many temperature dependent mutations becoming penetrant at this temperature (Lindsley and Grell 1968). As eclosion began, vials were checked twice daily, until the first females emerged. One of these females was then frozen and held for electrophoresis and size measurement. The total number of flies to emerge from each vial was counted after all flies had emerged.

Phenotypes: One or both wings of each fly were mounted on slides with transparent tape. The wings were projected on a 55 × 55 cm Hewlett-Packard 9864A digitizing board equipped with a cross-hair mouse. The coordinates of eight vein intersections, shown in Figure 2, were recorded, and the area enclosed by these calculated as a measure of size. Initial analyses also included linear wing dimensions, but the results obtained are qualitatively very similar to those for wing area (Houle 1988). In cases where two wings were digitized, the absolute value of the difference in area of the two was calculated as a measure of "fluctuating" asymmetry. The egg-to-adult development time of every fly was recorded, measured as the midpoint of the period between collections. Development rate is the inverse of development time, in days. A measure of the joint importance of size and development rate is the growth rate, the product of wing area and development rate, which is the amount of size added per unit development time. I also investigated the effect of heterozygosity on the absolute value of an individ-

![Figure 1](image-url)
null's deviation from the population mean for wing area, development rate, and growth rate.

**Analyses:** Tests for Hardy-Weinberg proportions and two locus gametic-phase disequilibria were performed with G tests of independence. Rarer allelic classes were pooled before testing to bring the expected frequencies in each cell above 5, if possible. Williams' correction was applied, except when there were empty cells, in which case Yates' correction for continuity was used (Sokal and Rohlf 1981).

Before further analysis of phenotype data, residuals were calculated from analyses of covariance on each variable, removing effects of experimental blocks, density (where significant), and sex and cross in experiment I. These residuals were tested for normality using the Box-Cox transformation. The data were then transformed until one was found which yielded log transformation. The data were then transformed, the ANCOVA repeated and the residuals again tested for normality. If the 95% confidence interval of X increased to 0.008, which is significantly less than the mendelian expectation of 0.5. Of the 850 vials set up for this experiment, 153 produced no wild-type flies at all. The 98 crosses, one failed to produce any wild-type flies, and two more crosses failed to yield any female wild-type flies. A total of 2,238 wild-type flies eclosed in this experiment.

The mean heterozygosity at each locus is shown in Table 1. Total heterozygosity, H, the number of heterozygous loci in each cross, ranged from 1 to 5 out of 7 loci which were polymorphic. The average H was 2.29. Five flies of each sex from each cross were selected at random, if available, and one wing mounted and its area measured. All the male flies from the crosses with 1, 4 or 5 heterozygous loci (273 flies) had both of their wings mounted, and the absolute value of the difference in wing area taken as a measurement of asymmetry. Undamaged wings were mounted for a total of 628 males, and 447 females. Means, variances, and Spearman rank correlations of the untransformed female data are given in Table 2.

Since heterozygosity varied only among crosses, to test for heterosis I first obtained least-squares means for each character in each cross. Least-squares means are an estimate of the mean that would have been obtained if sample sizes and covariates had been equal across days and sexes. Least-squares means were calculated with the SAS procedure GLM (SAS Institute, Inc. 1985b), after fitting models with main effects of sex, day, and cross. The transformations applied to each variable, the ANCOVA model used and the results of the analysis on each variable are shown in Table 3. I entered wing area into the model for development rate and development rate into the model for wing area to maximize the independence of the tests for heterosis in these characters. Tests for heterosis in growth rate, the product of development rate and wing area, would detect joint effects of heterozygosity on both characters. The resulting correlation structure for the cross least squares means is shown in the upper right portion of Table 2.
Allozyme Heterosis in Drosophila

TABLE 1
Enzymes run, number of alleles detected and heterozygosity (H)

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Locus</th>
<th>Chromosome map position</th>
<th>Alleles</th>
<th>H</th>
<th>Alleles</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid phosphatase-1</td>
<td>Acph</td>
<td>3-101.1</td>
<td>1</td>
<td>0.00</td>
<td>2</td>
<td>0.03</td>
</tr>
<tr>
<td>Alcohol dehydrogenase</td>
<td>Adh</td>
<td>2-50.1</td>
<td>2</td>
<td>0.42</td>
<td>2</td>
<td>0.44</td>
</tr>
<tr>
<td>Aldehyde oxidase</td>
<td>Ao</td>
<td>3-57.2</td>
<td>5</td>
<td>0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a-Amylase</td>
<td>Amy</td>
<td>2-77.7</td>
<td>3</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esterase-6</td>
<td>Est-6</td>
<td>3-36.8</td>
<td>2</td>
<td>0.42</td>
<td>3</td>
<td>0.46</td>
</tr>
<tr>
<td>Glucose-6-phosphate dehydrogenase</td>
<td>Zw</td>
<td>1-62.9</td>
<td>2</td>
<td>0.21</td>
<td>2</td>
<td>0.33</td>
</tr>
<tr>
<td>m-Glycerol-3-phosphate dehydrogenase</td>
<td>Gpdh</td>
<td>2-20.5</td>
<td>2</td>
<td>0.50</td>
<td>2</td>
<td>0.15</td>
</tr>
<tr>
<td>Hexokinase-C</td>
<td>Hex-C</td>
<td>2-73.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-Phosphogluconate dehydrogenase</td>
<td>Pgd</td>
<td>1-0.63</td>
<td>2</td>
<td>0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphoglucomutase</td>
<td>Pgm</td>
<td>3-43.4</td>
<td>3</td>
<td>0.13</td>
<td>4</td>
<td>0.27</td>
</tr>
<tr>
<td>Superoxide dismutase</td>
<td>Sod</td>
<td>3-54.6</td>
<td>2</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Genetic map positions are from TREAT-CLEMONS and DOANE (1984).

TABLE 2
Means, variances, and correlations of variables in experiment I

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
<th>Area</th>
<th>D-rate</th>
<th>G-rate</th>
<th>Area</th>
<th>D-rate</th>
<th>G-rate</th>
<th>Asymmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing area (mm²)</td>
<td>0.9902</td>
<td>616 × 10⁻³</td>
<td>—</td>
<td>-0.044</td>
<td>0.804</td>
<td>-0.553</td>
<td>-0.021</td>
<td>-0.604</td>
<td>0.173</td>
</tr>
<tr>
<td>Development rate</td>
<td>0.1005</td>
<td>120 × 10⁻⁵</td>
<td>0.355</td>
<td>-0.055</td>
<td>-0.377</td>
<td>-0.519</td>
<td>0.605</td>
<td>-0.321</td>
<td></td>
</tr>
<tr>
<td>Growth rate (mm¹/day)</td>
<td>0.0998</td>
<td>238 × 10⁻⁴</td>
<td>0.688</td>
<td>0.900</td>
<td>-0.531</td>
<td>0.573</td>
<td>-0.649</td>
<td>0.428</td>
<td></td>
</tr>
<tr>
<td>Deviations-Area</td>
<td>0.0572</td>
<td>289 × 10⁻³</td>
<td>0.138</td>
<td>-0.069</td>
<td>-0.182</td>
<td>0.765</td>
<td>-0.705</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-D-rate</td>
<td>0.0096</td>
<td>284 × 10⁻⁷</td>
<td>0.028</td>
<td>0.173</td>
<td>0.106</td>
<td>-0.330</td>
<td>0.498</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-G-rate</td>
<td>0.0124</td>
<td>841 × 10⁻⁷</td>
<td>-0.101</td>
<td>0.152</td>
<td>0.353</td>
<td>0.731</td>
<td>-0.196</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymmetry (mm²)</td>
<td>0.0090</td>
<td>690 × 10⁻⁷</td>
<td>0.080</td>
<td>0.028</td>
<td>0.070</td>
<td>-0.002</td>
<td>-0.057</td>
<td>-0.025</td>
<td></td>
</tr>
</tbody>
</table>

Means and variances are for the untransformed data on individual female flies (N = 447). Deviations are the absolute value of the individual score minus the sample mean. Spearman rank correlations between the untransformed female data are below the diagonal, and Pearson correlations among the 97 cross least squares means are above the diagonal. Mean and variance of asymmetry, and correlations of asymmetry with other variables are for the male flies (N = 272).

TABLE 3
Transformations and analyses of variance for experiment I

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Model¹</th>
<th>λ²</th>
<th>d.f.</th>
<th>F</th>
<th>P</th>
<th>Day</th>
<th>Sex</th>
<th>Development rate</th>
<th>Wing area</th>
<th>Error d.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing area</td>
<td>α₀ + C_i + B_i + β_ΔR_i + ε_i</td>
<td>1</td>
<td>96</td>
<td>12.64</td>
<td>****</td>
<td>NS</td>
<td>****</td>
<td>****</td>
<td>968</td>
<td>246</td>
</tr>
<tr>
<td>Development rate</td>
<td>α₀ + C_i + B_i + β_ΔR_i + ε_i</td>
<td>1</td>
<td>96</td>
<td>4.90</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>968</td>
<td>246</td>
</tr>
<tr>
<td>Growth rate</td>
<td>α₀ + C_i + B_i + ε_i</td>
<td>0.8</td>
<td>96</td>
<td>10.06</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>969</td>
<td>246</td>
</tr>
<tr>
<td>Deviations-w wing area</td>
<td>α₀ + C_i + B_i + β_ΔR_i + ε_i</td>
<td>0.5</td>
<td>96</td>
<td>5.18</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>969</td>
<td>246</td>
</tr>
<tr>
<td>-Development rate</td>
<td>α₀ + C_i + B_i + β_ΔR_i + ε_i</td>
<td>0.7</td>
<td>96</td>
<td>2.26</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>968</td>
<td>246</td>
</tr>
<tr>
<td>-Growth rate</td>
<td>α₀ + C_i + B_i + ε_i</td>
<td>0.8</td>
<td>96</td>
<td>4.86</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td>969</td>
<td>246</td>
</tr>
<tr>
<td>Asymmetry</td>
<td>C_i + B_i + β_ΔR_i + β_ΔA_i + ε_i</td>
<td>-7</td>
<td>21</td>
<td>0.79</td>
<td>NS</td>
<td>NS</td>
<td></td>
<td>NS</td>
<td>45</td>
<td>45</td>
</tr>
</tbody>
</table>

¹ ANCOVA model applied. In addition to the mean, which is not shown, effects are: α_i for the ith sex, C_i for cross, B_i for day (block), β_ are regression coefficients for A_wing area, or R, development rate, and ε_i is the residual for the ith individual.

² Box-Cox power transformation applied to the data. 1 = no transformation.

Significance level for each variable entered into the ANCOVA. Degrees of freedom are always 3 for days, and 1 for sex, development rate, and wing area. NS P > 0.05, * P < 0.05, *** P < 0.001, **** P < 0.0001.

wings were measured only on male flies, so asymmetry has no sex factor in the model. I originally fitted models including density of flies in each vial as a covariate, but these terms were never significant. With
FIGURE 3.—Correlations between cross total heterozygosity and cross least-squares means in experiment I. The horizontal bars represent the 95% confidence interval for r. For correlations with the primary characters, heterosis would be indicated by positive correlations. For deviations from the character means, negative correlations would be heterotic.

the exception of asymmetry, all characters had a highly significant cross component of variance. This term tests for the presence of genetic variance in each trait.

The product-moment correlations (r) between cross least-squares means and total heterozygosity are shown in Figure 3, along with approximate 95% confidence limits of each correlation coefficient, calculated as in Sokal and Rohlf (1981, pg. 585). For the primary characters in the upper part of Figure 3 heterosis would be indicated by significant positive correlations. None of the correlations approach significance, and the average r is only 0.039. The upper 95% confidence limits average 0.24, which sets an upper limit on the proportion of the genetic variance explained by H (r^2) at 5.6%. With the asymmetry and deviations from the mean in the lower part of Figure 3, heterosis would be indicated by negative correlations. Again, there is no hint of heterosis. Leaving out asymmetry, which is based on a smaller sample size, the lower 95% confidence limits rule out heterotic r^2 values of 4.1% or greater. The qualitative results remained the same when alternative ANCOVA models were applied, and when more or less stringent outlier removal schemes were used.

In testing for correlations with total heterozygosity, it is possible that I am obscuring heterosis at single loci. To test this possibility I calculated single classification analyses of variance for each locus-character combination separately. For each analysis, crosses were scored as either heterozygous or homozygous for the locus in question, without further regard for genotype. In Figure 4 I graph the probability value (P) associated with a one-tailed test of the hypothesis that heterozygotes are superior (higher mean or lower deviation from the mean) to homozygotes. Significant heterosis might be indicated if many tests cluster in the rejection region on the left hand side of the figure, or by isolated, highly significant tests. Overall, four of 49 tests are significant. Esterase-6 seems to be signifi-

FIGURE 4.—Single-locus tests for heterosis for experiment I. The graph shows the probability that heterozygotes do not have higher means than homozygotes for primary characters, indicated by diamonds; and the probability that heterozygotes do not have lower means for the remaining characters, indicated by circles. The vertical dotted line is drawn at P = 0.05, so values falling to the left of this line call for rejection of the null hypothesis.

FIGURE 5.—Plot of cross least-squares means for wing area versus Esterase-6 genotype in experiment I. The bars to the right of the distribution for each genotype represents the mean, ± 2 SE.

The possibility of heterosis at Esterase-6 is further investigated in Figure 5. Here I graph the cross least-squares means for wing area for the three genotypes found in this experiment, along with the genotype means and their 95% confidence intervals. A multiple comparisons test following an analysis of variance of the genotypes shows that the genotype 22 is significantly less than both 23 and 33. When the three crosses with very low means are deleted, there is still significant genetic variance among genotypes, although 22 is no longer significantly different from 33. Thus, while there is fairly strong evidence of
genotype specific differences associated with Est-6, the means suggest a dominance pattern rather than heterosis, although the power to distinguish the two is low. This pattern also holds for growth rate and deviation from mean wing area.

**Experiment II:** In this experiment, only one of the first female flies to eclose from each vial was measured and subjected to electrophoresis. Of the 1000 wild females allowed to oviposit, 622 produced viable female offspring, and 14 produced only male offspring. Sixty-five of these were identified as Drosophila simulans, a sibling species of D. melanogaster, by examining puparia (Kidwell 1981) and male siblings, or by the electrophoretic phenotype at Acph. The average number of flies eclosing from the remaining 557 vials that produced females was 23.5. A number of wings were damaged in mounting, so the sample size for the full analysis is 535. Both wings were mounted and measured from 200 flies selected at random. Means, variances and Spearman correlations of the raw data are shown in Table 4.

Comparison of means from the two experiments shows that the fluctuating temperature scheme in which flies were raised lengthened the developmental period substantially. The first females to emerge from each vial in experiment II had an average development time of 12.77 days, while the average development time of all female flies in experiment I was 9.95 days. The mean wing area for females in the two experiments was essentially identical; however, this includes late emerging flies in experiment I, which tended to be small, suggesting that female size was reduced in experiment II when adjusted for development rate. Both the longer development time and smaller size suggest that the temperature fluctuations were stressful for the developing flies. Complete genotype data for all 8 loci were obtained for 489 flies, all but 5 of these missing Hex-C. For individuals scored at all 8 loci, $H$ ranged from 0 to 6, and averaged 2.35.

Tests for heterosis were made on residuals from analyses of variance, removing the effects of extraneous variables. The transformations to normality used, and the results of the analyses of variance are presented in Table 5. I again removed the effect of wing area on development rate, and development rate on wing area, and depend on growth rate to detect heterosis due to joint effects. Unlike experiment I, density had strong effects on many characters. Density was uncorrelated with $H$ (analysis not shown), so I removed as much of its effect as possible by fitting higher order regression coefficients. The correlations among the resulting character values are shown in the upper right portion of Table 4.

The correlation of each character with $H$ is shown in Figure 6. Confidence intervals were calculated as in experiment I. Because each correlation is based on a larger sample size than in experiment I, the scale is reduced from Figure 3. As before, there is no evidence of positive correlations with size or development rate, and no evidence for negative correlations with asymmetry of deviations. Several individual correlations approach significance, but none in a heterotic direction. The average upper confidence limit for heterosis in size and rate is only 0.029, so heterotic $r^2$ values must be less than 0.1%. For deviations from the mean, heterotic $r^2$ values are less than 0.3%. The $P$ values associated with single-locus tests of heterosis are presented in Figure 7. No single locus shows a concentration of significant tests. Analyses of residuals from alternative ANCOVAs and of the raw data gave very similar results.

**DISCUSSION**

I have performed two very different experiments designed to detect allozyme heterosis. In the first experiment, I characterized a modest number of genotypes well. This allowed a direct test for a correlation between heterozygosity and genotypic value. This ex-

---

**TABLE 4**

Means, variances and correlations of variables in experiment II

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Variance</th>
<th>Area</th>
<th>D-rate</th>
<th>G-rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing area (mm$^2$)</td>
<td>0.9800</td>
<td>244 $\times$ 10$^{-8}$</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Development rate (day$^{-1}$)</td>
<td>0.0783</td>
<td>795 $\times$ 10$^{-8}$</td>
<td>0.154</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Growth rate (mm$^2$/day)</td>
<td>0.0769</td>
<td>268 $\times$ 10$^{-7}$</td>
<td>0.838</td>
<td>0.081</td>
<td>—</td>
</tr>
<tr>
<td>Deviations--Area</td>
<td>0.0396</td>
<td>961 $\times$ 10$^{-6}$</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>-D-rate</td>
<td>0.0023</td>
<td>286 $\times$ 10$^{-8}$</td>
<td>0.133</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>-G-rate</td>
<td>0.0040</td>
<td>107 $\times$ 10$^{-7}$</td>
<td>—</td>
<td>0.043</td>
<td>0.146</td>
</tr>
<tr>
<td>Asymmetry (mm$^2$)</td>
<td>0.0100</td>
<td>916 $\times$ 10$^{-7}$</td>
<td>0.019</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Means and variances are for the untransformed variables. Spearman rank correlations between the untransformed variables are below the diagonal, and Pearson correlations of the residuals from ANCOVAs on transformed variables above the diagonal. Sample sizes are given in Table 5.
TABLE 5
Transformations and analyses of variance for experiment II

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Model</th>
<th>$\lambda$</th>
<th>$N$</th>
<th>Box</th>
<th>$D$</th>
<th>$D^2$</th>
<th>$D^3$</th>
<th>Development rate</th>
<th>Wing area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing area</td>
<td>$B_1 + \beta_1 D_{m} + \beta_2 D_{m} + \beta_3 A_{m} + \beta_4 R_{m} + \epsilon_{m}$</td>
<td>1</td>
<td>535</td>
<td>**</td>
<td>*</td>
<td>NS</td>
<td>NS</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>Development rate</td>
<td>$B_1 + \beta_1 D_{m} + \beta_2 D_{m} + \beta_3 A_{m} + \epsilon_{m}$</td>
<td>1</td>
<td>557</td>
<td>****</td>
<td>***</td>
<td>**</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Growth rate</td>
<td>$B_1 + \beta_1 D_{m} + \beta_2 D_{m} + \beta_3 A_{m} + \epsilon_{m}$</td>
<td>1</td>
<td>535</td>
<td>****</td>
<td>***</td>
<td>****</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Deviations-wing area</td>
<td>$B_1 + \beta_1 D_{m} + \beta_2 D_{m} + \beta_3 A_{m} + \epsilon_{m}$</td>
<td>0.3</td>
<td>535</td>
<td>NS</td>
<td>NS</td>
<td>*</td>
<td>*</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>-development rate</td>
<td>$B_1 + \beta_1 D_{m} + \beta_2 D_{m} + \beta_3 A_{m} + \epsilon_{m}$</td>
<td>0.3</td>
<td>535</td>
<td>NS</td>
<td>NS</td>
<td>*</td>
<td>*</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>-growth rate</td>
<td>$B_1 + \beta_1 D_{m} + \beta_2 D_{m} + \beta_3 A_{m} + \epsilon_{m}$</td>
<td>0.4</td>
<td>535</td>
<td>NS</td>
<td>NS</td>
<td>*</td>
<td>*</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Asymmetry</td>
<td>$B_1 + \beta_1 D_{m} + \beta_2 D_{m} + \beta_3 A_{m} + \beta_4 R_{m} + \beta_5 A_{m} + \epsilon_{m}$</td>
<td>0.27</td>
<td>201</td>
<td>NS</td>
<td>*</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
</tbody>
</table>

* ANCOVA model applied. In addition to the mean, which is not shown, effects are: $B_k$ for the $k$th box (block); $\beta_*$ are regression coefficients for density, $D$, wing area, $A$, development rate, $R$, and $\epsilon_*$ is the residual for the individual from the $m$th vial.

* Power transformation applied to the data. $1$ = no transformation, $0$ = log transformation.

Sample sizes and significance levels for each variable entered into each ANCOVA. Degrees of freedom are always 9 for boxes, and 1 for each covariate: density, development rate, and wing area.

A persistent problem in the detection of weak phenomena, such as allozyme heterosis, is provided by departures from normality and the resulting inflation of type I error rates. Growth character distributions in $D. melanogaster$ are very commonly skewed, with a long tail of poorly performing individuals. A substantial portion of this is likely to be due to environmental causes in outbred populations. This is supported by the fact that in experiment I such individuals occurred in a large proportion of crosses. In experiment I, the influence of such individuals was minimized because many individuals of the same genotype were obtained. In experiment II, I precluded the inclusion of such individuals by sampling only the fastest developing female from each sibship. This sampling scheme resulted in lower phenotypic variance in every trait except asymmetry, and therefore probably obscured some of the genetic variance in all traits, especially development rate. The optimal tradeoff between inclusion of all individuals and the necessarily somewhat arbitrary truncation of the distribution sampled depends on the degree to which extreme poor performance is genetically based. This is unknown. The scheme chosen for experiment II successfully normalized the phenotypic distributions, so no individuals were sampled with unusually high influence on the results.

Experiment II also included temperature variation during development of the flies. The temperature range chosen, 28–18 degrees, is less than the Davis Peach Farm population often experiences. Although 28$^\circ$ is a stressful temperature for $D. melanogaster$ in the laboratory (ASHBURNER and THOMPSON 1978), peaches with living larvae at Davis Peach Farm often reach higher temperatures (personal observation).

**Figure 6.**—Correlations, $\pm 95\%$ confidence limits, between total individual heterozygosity and characters in experiment II. Note that the scale of the $x$-axis is larger than in Figure 3.

**Figure 7.**—Single-locus tests for heterosis for experiment II. Presentation like that for Figure 4.

Experiment was carried out at low density at a constant temperature. In the second experiment I sampled one of the first female flies to emerge in sibships obtained from wild females, so each fly was genetically independent. These flies were raised in a stressful, fluctuating temperature regime, and at a higher density. In neither experiment was there evidence of heterosis.
Temperature fluctuations might be expected to favor the expression of heterosis for two reasons. Heterozygotes at allozyme loci are usually intermediate to corresponding homozygotes in their biochemical phenotypes (Gillespie and Langley 1974; Cacser and Burns 1981), so if functional overdominance occurs it is likely to be due to marginal overdominance. While it is possible to conceive of some source of variance that would affect flux through any enzymatic step, temperature is the only variable which is likely to affect all enzymes. Studies of enzyme kinetics also make it clear that allozymes very often differ in their response to temperature (Zera, Koehn and Hall 1985). In addition, the reduction in size and lengthening of development brought about by the temperature fluctuations indicate that culture conditions were less than optimal. There is some evidence that physiological stress enhances heterosis (Barlow 1981; Scott and Koehn 1989).

These experiments share two sources of uncertainty about the identity of genotypes scored. The first is inversion polymorphism. Inversions are found on each autosomal chromosome arm of D. melanogaster, and are often in strong gametic disequilibrium with allozyme loci which map within or near their breakpoints, including Adh, Gpdh and Est-6 (Voelker, Cockerham and Johnson 1978). Flies were not scored for inversions in my experiments. Populations at low latitudes tend to be moderately polymorphic for these inversions, but in populations above 40°N latitude in North America, such as Davis Peach Farm, the frequency of the common gene arrangement is usually much greater than 90%. Inversion polymorphisms in gametic disequilibrium with the allozymes I scored would lead allozymes to mark large blocks of chromosome, which should increase my chances of observing heterosis. The significance of my failure to observe heterosis is not affected by inversion polymorphism.

Another important source of error in scoring genotypes is cryptic electrophoretic variation, which has been relatively well characterized in D. melanogaster. Of the loci used in this study, the apparently biallelic polymorphisms at Adh, Gpdh, Zw, and Pgd remain biallelic when examined more closely, although rare variants are sometimes found (Kreitman 1980; Coyne et al. 1979; Eanes 1983; Eanes and Hey 1986; W. F. Eanes, unpublished results). Pgm (Trippa, Loverre and Catamo 1976) and Ao (Coyne 1982) both may have some cryptic variation. Est-6, at which I found two common alleles, is in fact much more polymorphic (Cochrane and Richmond 1979; Cooke, Richmond and Oakeshott 1987; Labate et al. 1989), with "allele" 3 (Figure 5) being particularly likely to consist of many rare variants. No characterization of this variation has been published for Acph, Amy, Hex-C, or Sod. Alleles at each of these loci separate very well on the electrophoretic systems I used, and have low heterozygosity. Such loci do not generally possess large amounts of cryptic variation (Coyne 1982), the notable exception being Est-6. There is some potential for underestimating heterozygosity in these experiments, as there is in all studies of allozyme heterosis.

In spite of the differences in design of the two experiments, the results are consistent: there is no evidence for allozyme-associated heterosis in size or development rate in the outbred D. melanogaster studied. Previous studies testing for allozyme-associated heterosis in D. melanogaster have given conflicting results. Very large studies of egg-to-adult viability and fertility by Mukai et al. (1974) turned up no evidence for genotype specific differences at six polymorphic allozyme loci, including Adh, Gpdh, and Est-6, as well as for the common gene arrangements on both chromosomes. Serradilla and Ayala (1983) found exceptionally large amounts of heterosis for fertility at Adh, Gpdh, and Acph; but little evidence of genotypic differences for viability or development rate. The results of these studies are diametrically opposed for fertility, and difficult to reconcile. A potential explanation is that Serradilla and Ayala did not control for inversion polymorphism, which is likely to be high in the population they sampled, although the failure of Mukai et al. (1974) to find inversion specific phenotypic differences in their populations argues against this explanation. Taken together, the experiments reported here and those of Mukai et al. provide strong evidence that allozyme-associated heterosis does not occur in outbred D. melanogaster.

However, D. melanogaster is capable of expressing allozyme heterosis under conditions favoring a correlation between genotypes. Biembourg (1983) assayed the viability of offspring of brother-sister matings and the heterozygosity of the parents, and found a consistent relationship between \( H \) and whether offspring viability was greater or less than 0.90. Crosses with offspring viability less than 0.90 probably segregated for recessive lethals, and allozymes would have marked the identity by descent of chromosomes in these inbred families. Sheerif and Skibinski (1988) found that heterozygotes for Adh and Gpdh possessed sternopleural bristle counts significantly nearer the mean than those of homozygotes in a cage population of D. melanogaster where there was significant gametic disequilibrium between these two second chromosome loci. This could have been due to inversion In(2L)\( t \), which is sometimes in gametic disequilibrium with both loci (Voelker, Cockerham and Johnson 1978), or to selection or founder events in the cage. Girard (1986) found a negative correlation between development time and heterozygosity in the F\(_2\) from
crosses between a single pair of lines differentiated for six allozyme loci.

As outlined in the Introduction, allozyme and molecular studies have shown that there is very little gametic or identity disequilibrium in natural populations of *D. melanogaster*, and hence little opportunity for dominance-correlation heterosis. Other large studies have also failed to find evidence for allozyme heterosis in mobile populations of large size (WARD et al. 1985; McANDREW, WARD and BEARDMORE 1986).

On the other hand, bivalve molluscs, with an apparently panmictic population structure due to a pelagic larval phase, consistently show allozyme heterosis (reviewed by ZOUROS 1987; see also KOEHN, DIEHL and SCOTT 1988; ZOUROS, ROMERO-DOREY and MALLET 1988; GAFFNEY 1989). However, bivalves also show consistent heterozygote deficiencies (ZOUROS and FOLTZ 1984), which are difficult to reconcile with the overdominance hypothesis. The occurrence of a deficiency seems to be a necessary condition for heterosis: studies which do not find heterozygote deficiencies also do not find heterosis (ZOUROS 1987). Furthermore, within the two largest studies of allozyme heterosis in bivalves, the degree of heterosis at individual loci is correlated with the amount of heterozygote deficiency (ZOUROS, SINGH and MILES 1980; P. M. GAFFNEY, T. M. SCOTT, R. K. KOEHN and W. J. DIEHL, unpublished data). Population structure seems to be involved in the expression of heterosis, as progeny from mass spawnings of few individuals, or pair matings show little evidence of heterosis (GAFFNEY and SCOTT 1984; MALLET et al. 1986; ZOUROS 1987).

There are a number of possible explanations for these correlated phenomena, such as spatial or temporal population structure, segmental aneuploidy, or underdominance in larvae, coupled with overdominance in adults (ZOUROS and FOLTZ 1984; ZOUROS 1987; ZOUROS, ROMERO-DOREY and MALLET 1988; However, all of these hypotheses have their difficulties, and no consensus exists among workers in this area as to which are most significant. Given these peculiarities, it is possible that the allozyme heterosis observed in bivalves is a different phenomenon from heterosis in other organisms.

Outside of these studies, a review of the literature suggests that there is little evidence for allozyme heterosis in any population which is not likely to be partially inbred, have very small effective population size, or have recently undergone strong directional selection. All these factors enhance the probability that dominance-correlation heterosis will occur. Obvious examples are plant populations which primarily outbreed, but self at a low rate (LEDIG, GURIES and BONEFELD 1983; STRAUSS 1986). Inbreeding depression is common in such species (LEDIG 1986; CHARLESWORTH and CHARLESWORTH 1987) and homozygosity will be correlated with the probability an individual is the product of selfing (HALDANE 1949; WEIR and COCKERHAM 1973). Contrary to the claims of BUSH, SMOUSE and LEDIG (1987), allozyme-associated heterosis in such cases is entirely consistent with inbreeding (HOULE 1988). STRAUSS (1986) showed that there is a strong correlation of heterozygosity and growth or reproduction among selfed individuals, but equivocal evidence for heterosis in outbred individuals.

Many animals have population structures which may be expected to promote gametic disequilibrium. Examples include populations centering on small discrete patches of resource which are probably subject to relatively frequent local extinction, followed by recolonization. An excellent example of this are populations of amphibians in small ponds, such as *Ambystoma tigrinum* (PIERCE and MITTON 1982) and *Bufo boreas* (SAMOLLOW and SOULE 1983) in which some evidence of heterosis has been found. Limited dispersal will also promote local inbreeding and therefore identity disequilibrium. This may be enhanced by social structure, or habitat fragmentation. This situation applies to many vertebrates in which heterosis has been observed, such as grouse (REDFIELD 1974) and white-tailed deer (COTHRAN et al. 1983). In plants, neighbors are often relatives, as shown by allozyme distributions (LEDIG 1986; SCHAAI 1975). In such a population, homozygosity will again tend to be correlated with inbreeding. A good example of this may be the perennial herb *Liatris cylindracea*, in which there is substantial spatial structure to allozyme frequencies (SCHAAI 1975), and young cohorts have large heterozygote deficiencies which decrease with age (SCHAAI and LEVIN 1976).

In domesticated organisms, it is extremely likely that there has been substantial inbreeding associated with domestication. Breeders very frequently resort to crosses among more distantly related strains to counteract these effects. This will generate substantial gametic disequilibrium which may persist for many generations beyond such crosses, particularly as selection will tend to favor heterozygous individuals. In addition, strong directional artificial or natural selection in the novel artificial environments most domesticated populations are exposed to will tend to promote gametic disequilibrium as favorable alleles move towards fixation (MAYNARD SMITH and HAIGH 1974; THOMSON 1977). The observation that populations subject to artificial directional selection often reach a selection limit due to decreasing fecundity or viability before the genetic variance is exhausted, and that temporary relaxation of selection sometimes allows such limits to be extended provides an extreme example of this phenomenon (YOO 1980; FALCONER 1981). Other data consistent with this are the large
allozyme frequency changes in newly established cage populations of Drosophila, which cannot be repeated in older populations (Yamazaki 1971; Fontevila et al. 1975; Hedrick 1976; Barker 1977; Powell and Wistrand 1978; Yamazaki et al. 1985; Arnason 1988). Consequently, heterosis in domesticated plants and animals cannot be regarded as evidence of functional overdominance. The semi-domesticated hatchery populations including such selection in the novel hatchery environment.

In conclusion, there is little evidence favoring the hypothesis that functional overdominance plays a role in observations of allozyme heterosis. Outside of the bivalves, studies which have demonstrated allozyme heterosis were carried out in organisms whose population structure or history promotes genotypic correlations, allowing the expression of deleterious recessive alleles to become associated with allozyme homozygosity. Until there is evidence that allozyme heterosis occurs in large, panmictic, natural populations, the hypothesis of functional overdominance must be regarded with considerable skepticism.


Allozyme Heterosis in Drosophila


