Epistasis in Measured Genotypes: *Drosophila* P-Element Insertions

Andrew G. Clark and Lei Wang

*Institute of Molecular Evolutionary Genetics, Department of Biology, Pennsylvania State University,
University Park, Pennsylvania 16802*

Manuscript received September 16, 1996

Accepted for publication May 15, 1997

ABSTRACT

Transposon tagging provides an opportunity to construct large numbers of strains of organisms that differ by single insertional mutations. By scoring the phenotypes of these "measured genotypes," powerful tests of effects of mutations on phenotypic expression have been performed. Here we extend this approach by constructing with simple crosses all possible two-locus genotypes for each of eight pairs of P-element insertions. Analysis of metabolic phenotypes (fat and glycogen contents, enzyme activities, total protein, and body weight) of the resulting nine genotypes provides direct estimates of additive, dominance, and epistatic effects of the mutations. Nested two-way analysis of variance identified significant epistatic effects in 27% of the tests (35/128 of the trait $\times$ P-element combinations). Posterior contrasts were performed to partition the epistatic variance into the four orthogonal components of COCKERHAM, and the data exhibit a tendency toward additive $\times$ dominance and dominance $\times$ dominance epistasis. Mutations in this study have epistatic effects on metabolic traits that are on the same order of magnitude as main (additive and dominance) effects. Measured genotypes have been used in other contexts to quantify epistatic effects on phenotypic expression, and these results are also briefly reviewed.

EPISTASIS is an important topic in evolutionary genetics, particularly because of the fact that epistasis in fitness may lead to multiple fitness peaks (WHITLOCK *et al.* 1995). Population genetic theory has shown that multiple fitness peaks may give rise to multiple stable equilibria, a situation that may lead to a population failing to achieve the most adaptive (maximum fitness) state, or it may even lead to diversification and ultimately speciation. Epistasis in fitness, defined as any departure from additivity of effects across loci, is important in other evolutionary contexts. CHARLESWORTH *et al.* (1993) showed conditions under which epistasis could slow the rate at which MULLER's ratchet fixes deleterious mutations in finite populations. In general, epistasis can affect the dynamics of allele frequency change (KARLIN 1975) and the equilibrium distribution of allelic effects (GAVRILETS and DEJONG 1993).

Epistasis may also be of critical importance to populations that have fallen to very small size. The inbreeding that occurs at population bottlenecks results in a change in the variance components in quantitative genetic models. Empirical evidence for an epistatic trait realizing an increase in additive variance following a bottleneck (BRYANT and MEFFERT 1993) has been supported by theory that shows how additive variance can increase if there is epistasis (GOODNIGHT 1988). However, increased additive variance after a bottleneck can also occur if the founding population had high levels of dominance variance (WILLIS and ORR 1993), so in-

creased additive variance following a bottleneck is not evidence for epistasis. On the other hand, whereas variance components are useful in predicting the changes that occur because of natural selection, departure from simple additivity and dominance of measured genotypes is also of importance. With increasing application of anonymous molecular markers, there will be a huge opportunity to quantify this "physiological epistasis" (CHEVERUD and ROUTMAN 1995). Epistasis that is identified from measured genotypes is different from the variance components one might estimate from a North Carolina design III cross, because, with models using measured genotypes, one can make predictions about the phenotype of an individual given the genotype. The problem of predicting phenotypes arises in many genetic contexts, from agricultural to medical, so careful assessment of the adequacy of additive models is particularly important.

Epistasis may also play a large role in the maintenance of quantitative genetic variation. The fate of a new mutation depends on the pleiotropic effects of the mutation and the pattern of epistasis exhibited between the new mutation and the background genetic variation. Models of mutation-selection balance that depend on extensive pleiotropic effects and resulting "apparent" selection mediated through pleiotropic traits (TURELLI 1985) still imply a high mutation load. GAVRILETS and DEJONG (1993) show that epistasis among pleiotropically related genes can result in much stronger apparent selection and, hence, maintenance at equilibrium of greater genetic variation. Moreover, epistasis can maintain genetic variance with a lower vari-

Corresponding author: Andrew G. Clark, Institute of Molecular Evolutionary Genetics, Department of Biology, Pennsylvania State University, University Park, PA 16802. E-mail: c92@psu.edu

ance in fitness (genetic load) than either models with direct stabilizing selection or pleiotropic apparent selection without epistasis. Our understanding of the importance of epistasis in the maintenance of quantitative genetic variation hinges on empirical efforts to quantify both epistatic variance and physiological epistasis.

Considering epistasis in traits other than fitness, one finds a wide range of claims of its importance. To the extent that additivity of genic effects arises from the fact that effects have multiple independent causes, one might expect that whole-organism phenotypes such as body size would be largely additive. By contrast, traits closely related to expression of a single gene are dictated by fewer components that may exhibit strong interactions during normal functioning. GIBSON (1996) makes the case clearly that the determinants of variation in transcription rate of a single gene are likely to be highly epistatic, because of the intermolecular interactions that are necessary during assembly of the transcription complex. It is in some cases possible to determine the molecular mechanisms for epistasis in metabolic traits whose components are discernible (FRERIKSEN *et al.* 1994). On the other hand, there is little hope in understanding why a trait like body weight exhibits epistasis. Measured genotype approaches offer the possibility for considerably greater resolution of epistatic interactions, and they are accumulating evidence that epistasis is not uncommon (CABOT *et al.* 1994; PALOPOLI and WU 1994; DOEBLEY *et al.* 1995; LONG *et al.* 1995). Here we quantify the epistasis in the effects of pairs of *P*-element insertions on 16 quantitative characters in order to determine relative magnitudes of additive, dominance, and epistatic effects in many independent tests. Performance of so many tests at once also allows observation of patterns of significant epistasis across genotypes and across traits.

MATERIALS AND METHODS

Drosophila stocks: Single *P*-element-bearing lines in a homogeneous genetic background were constructed by standard methods, mobilizing the *CaSpeR* transposable element with the jumpstarter stock containing the engineered $\Delta 2-3$ *P*-element as a transposase source (COOLEY *et al.* 1988; ROBERTSON *et al.* 1988). Details of the initial stock construction appear in CLARK *et al.* (1995). Sixteen lines were selected from the original set of 263 *P*-element-tagged lines based on their stable low-level expression of *white*. This selection was necessitated by the need to be able to see differences in eye color among flies with zero, one, two, three, or four *P*-element insertions. Although the lines with low levels of *white* expression are not a random sample and probably have elements in regions of the genome whose chromatin structure dictates low levels of expression, it seems unlikely that this criterion biases the tendency of elements to exhibit epistatic interactions.

For each of eight pairs of *P*-element-insertion lines, a second-chromosome and a third-chromosome insertion line were crossed at random to produce F_1 flies, which were then crossed to give an F_2 generation, which exhibited a wide range in eye color phenotypes. The line with the darkest pigmentation was putatively the line doubly homozygous for the *P*-

element insertions. Virgins were crossed to produce a stable double-insert strain (*aabb*), and its genotype was verified by crossing to *Oregon-R*, and intercrossing the F_1 to recover expected ratios of single- and double-insertion lines in the F_2 . Crosses between the single-insert strains (*aaBB* and *AAbb*), the double-insert strain, and the no-insert *white* eye strain (*AABB*) used to found the lines produced all nine two-locus genotypes. The line that is doubly heterozygous was made both by crossing the two single-insert homozygotes and by crossing the double-insert line with the zero-insert line. These two crosses should produce identical offspring genotypes, and both were scored for all quantitative phenotypes to test this. At the time the double-insertion lines were constructed, the single *P*-element-tagged lines were 13 months old. Mutations in the genetic background could have accumulated in the lines, resulting in differences that should not be spuriously ascribed to the *P*-elements. To test for the magnitude of such background effects, each genotype was constructed by two independent sets of crosses in order to obtain random (and potentially different) samples of the background. This test really only asks whether the two constructs are the same, and it does not rule out the possibility that other mutations arose and went to fixation during the *P*-element transposition or subsequent 13 months before the lines in this study were constructed.

Assays of enzyme activities: *Drosophila* stocks were reared under uniform conditions, anesthetized, weighed, homogenized, centrifuged, and distributed into microtiter plates as described previously (CLARK *et al.* 1995). For each sample of flies, 16 characters were quantified, including live weight, total soluble protein, lipid and glycogen contents, and the activities of 12 enzymes in intermediary metabolism. All characters other than live weight were assayed with a microtiter plate reader (VMAX, Molecular Devices Inc.). Further details of the procedures appeared in CLARK and WANG (1994). Each microtiter plate contained homogenate samples as well as standards and controls. The 15 identical plates were used to assay triacylglycerol content, glycogen content, total protein content, and the activities of alcohol dehydrogenase (ADH, at 2-50.1), fatty acid synthase (FAS), glucose-6-phosphate dehydrogenase (G6PD, map location 1-63.0), alpha-glycerol-3-phosphate dehydrogenase (GPDH, map location 3-55.4), glycogen phosphorylase (GP), glycogen synthase (GS), hexokinase (HEX, Hex-A at 1-29.3 and Hex-C at 2-75.0), malic enzyme (ME, at 3-53.1), 6-phosphogluconate dehydrogenase (6PGD, at 1-0.64), phosphoglucose isomerase (PGI, at 2-58.0), phosphoglucumutase (PGM, at 3-43.4), and trehalase (TRE, at 2-92.0). The units of TRI and GLY are μg per fly, and the units of the enzyme activities are nm NADP (or NAD) reduced per fly per minute. The microtiter plate reader was programmed to record 6–12 optical densities of each well at prescribed time intervals depending on the assay. The design included eight *P*-element pairs $\times$ 10 genotypes per pair $\times$ two replicate constructions $\times$ four replicate assays.

Statistical analysis: The data were fitted to linear models that had the following classifications: second chromosome genotype (A_i , where $i = 0, 1$, or 2 copies of the *P*-element), third chromosome genotype (B_j , where $j = 0, 1, 2$), replicate construct (C_{ijk}), and an error term (e_{ijkl}) representing differences between replicate samples. Null hypotheses of interest included the equivalence of the two doubly heterozygous lines, additivity of effects of the two *P*-element inserts, and fits to models that parameterize epistasis. These null hypotheses were tested separately for all 16 characters assayed and for each of the eight *P*-element pairs. The following model for analysis of variance (ANOVA) was used for the simple test of additivity of the two *P*-elements' effects:

$$Y_{ijkl} = \mu + A_i + B_j + (A*B)_{ij} + C_{ijk} + e_{ijkl}$$

The 14 traits other than live weight and total protein content are expected to be correlated with body size, so the covariates weight and protein were removed in the following model of analysis of covariance:

$$Y_{ijkl} = \mu + A_i + B_j + (A*B)_{ij} + C_{ijk} \\ + \beta_w(w_{ijkl} - \bar{w}) + \beta_p(p_{ijkl} - \bar{p}) + e_{ijkl}$$

where μ is the grand mean, and C_{ijk} is the effect of the k^{th} construct nested in the i^{th} second chromosome and j^{th} third chromosome insert. The covariates in the analysis of covariance were weight (w_{ijkl}) and protein content (p_{ijkl}), whose means were $\bar{w}$ and $\bar{p}$ and whose regression coefficients were β_w and β_p , respectively. In applying this model there is an implicit assumption that variation in weight and protein is the cause of variation in activities. Models with and without weight and protein as covariates were fitted, because these linear models cannot reveal which character is causal. In the ANOVA, the interaction term ($A \times B$) was tested against the appropriate mean square (either error MS or the construct MS if the latter was significant). If the interaction was significant, four posterior contrasts were performed to assess significance of the four orthogonal components of epistasis as described by COCKERHAM (1954). Further details of subsequent statistical tests are presented along with the results.

RESULTS

Heterogeneity among constructs: If the independent constructs of lines were heterogeneous it would imply that background variation had accumulated by mutation and that particular *P*-element genotypes include a heterogeneous array of genotypes. This would not be fatal to the experiment, but it would be necessary to perform subsequent tests taking the construct heterogeneity into account. The most powerful test of a construct effect is to combine the results into a large ANOVA with three main crossed factors: *P*-element set, element *A* presence/absence, and element *B* presence/absence. The construct effect was nested within each combination of *P*-element pairs $\times A \times B$, and a separate test was done for each trait. Of the 16 tests, 10 were found to be significant at the 5% level. In subsequent analyses of variance, the construct mean square was used as the error term in assessing significance.

Homogeneity of double heterozygotes: Lines that were doubly heterozygous for the two *P*-element insertions were constructed by the crosses $AABB \times aabb$ and $AAbb \times aaBB$. Provided all the lines were constructed without introduction of errors, the progeny from these two crosses should be identical. We tested this null hypothesis for all eight *P*-element pairs and all 16 phenotypes by simple ANOVA. Of the 128 tests, four were found to be significant at the 5% level. Because one expects 5% of the tests to be significant by chance alone, we applied the Bonferroni correction for multiple comparisons and found that none of the tests was significant at an experiment-wide 5% level. (Analysis of covariance and factoring out weight and protein covariates gave the same results.) Failing to reject the null hypothesis of equivalence of the two double heterozy-

FIGURE 1.—Graphic representation of the phenotypic measures of the nine genotypes that were constructed for each *P*-element pair. Parallel lines imply additivity across loci, and departures from parallel indicate interaction on the given scale.

gotes, we pool them and end up with nine genotypes for each pair of *P*-element insertions.

ANOVA: The original definition of epistasis arose in the context of an interaction term in ANOVA. Graphically, interaction variance arises when there is a departure from parallel lines in a plot like Figure 1. Analogous plots of corresponding phenotypes for all eight pairs of *P*-elements and all 16 traits appear to show frequent departures from parallel (Figure 2). Among the plots in Figure 2, one can find examples of apparently pure additive genes, dominance, overdominance, and, in cases where a single homozygous genotype has a strongly aberrant phenotype, clear cases of epistasis. Test of the significance of epistatic variance amounts to a test for significance of the interaction term in ANOVA (Table 1). Of the 128 tests for an interaction between the two *P*-elements (eight *P*-element pairs $\times$ 16 traits), 35 were found to be significant at the 5% level. Even after adjusting significance levels for multiple comparisons, there were 23 cases of significant interaction. Repeating the analyses with analysis of covariance, adjusting out the effects of live weight and total protein, the number of cases of significant interaction (without correction for multiple comparisons) was 33. These figures suggested that a deeper look into the structure of the epistatic effects was warranted.

Posterior contrasts for component of epistasis: COCKERHAM (1954) noticed that the nine phenotypes for two diallelic loci could be tested in eight orthogonal contrasts, corresponding to the two additive effects parameters, two dominance parameters, and four epistatic terms, additive $\times$ additive ($A \times A$), additive $\times$ dominance ($A \times D$), dominance $\times$ additive ($D \times A$), and dominance $\times$ dominance ($D \times D$) (Table 2). These components were tested with the CONTRAST statement in the SAS procedure GLM. The magnitudes of the *F* statistics for tests of additive, dominance, $A \times A$, $A \times D$, $D \times A$, and $D \times D$ are plotted in Figure 3. The

FIGURE 2.—Thumbnail plots like that in Figure 1 for all 16 traits in all eight *Pelement* pairs. The vertical scale is linear in all cases.

striking pattern is that, of the epistatic parameters, $A \times A$ is much less important than the other two, $A \times D$ effects are roughly equal in magnitude, and epistatic effects seem to be even stronger than additive or dominance effects. Results of the significance tests appear

TABLE 1

Example analysis of variance and covariance for triacylglycerol levels of *Pelement* pair 2

Source	d.f.	SS	MS	F	P
A insert	2	0.2449	0.1224	0.69	0.5250
B insert	2	0.1466	0.0733	4.15	0.0529
$A \times B$	4	0.4132	0.1033	5.84	0.0134
Construct ($A \times B$)	9	0.1591	0.0176	2.79	0.0098
Error	50	0.3163	0.0063		
Weight	1	0.0064	0.0064	2.17	0.1475
Protein	1	0.1505	0.1505	50.83	0.0001
A insert	2	0.0008	0.0004	0.08	0.9252
B insert	2	0.0715	0.0357	6.44	0.0184
$A \times B$	4	0.1410	0.0352	6.34	0.0104
Construct ($A \times B$)	9	0.0500	0.0055	1.88	0.0786
Error	48	0.1422	0.0029		

Note that the construct mean square is used as the denominator in tests of significance. The r^2 for the top model is 0.72 and that for the bottom model is 0.87.

TABLE 2

COCKERHAM's (1954) weights for posterior contrasts to test significance of variance components

	$\frac{AB}{AB}$	$\frac{AB}{Ab}$	$\frac{Ab}{Ab}$	$\frac{AB}{aB}$	$\frac{AB}{ab}$	$\frac{Ab}{ab}$	$\frac{aB}{aB}$	$\frac{ab}{aB}$	$\frac{ab}{ab}$
Additive-1	2	2	2	0	0	0	-2	-2	-2
Additive-2	2	0	-2	2	0	-2	2	0	-2
Dominance-1	1	1	1	-2	-2	-2	1	1	1
Dominance-2	1	-2	1	1	-2	1	1	-2	1
$A \times A$	1	0	-1	0	0	0	-1	0	-1
$A \times D$	1	-2	1	0	0	0	-1	2	-1
$D \times A$	1	0	-1	-2	0	2	1	0	-1
$D \times D$	1	-2	1	-2	4	-2	1	-2	1

Additive-1 and additive-2 correspond to the additive effect of insert 1 and insert 2, respectively. Components of epistasis are as follows: $A \times A$, additive $\times$ additive; $A \times D$, additive $\times$ dominance; $D \times A$, dominance $\times$ additive; and $D \times D$, dominance $\times$ dominance.

in Figure 4, which shows all 512 tests (eight *Pelement* pairs $\times$ 16 traits $\times$ four epistasis parameters). The heterogeneity across *Pelement* pairs is readily apparent: set 3 exhibited very little epistasis, while set 2 exhibited considerable interaction. Note also that if one trait exhibited particularly strong $A \times D$ epistasis, then other traits were more likely to show the same component as significant (see set 5 for example).

Single-genotype departures: Regression methods were used to fit the phenotype data to models having additive and dominance effects of both loci with strict additivity of effects across loci. These least-squares estimates gave predicted phenotypes and confidence intervals for all nine phenotypes, under the model of no epistasis. If observed phenotypes fell outside the 95% confidence interval for the predicted phenotype, then this was scored as significant evidence for epistasis. Note that this approach is independent of population allele frequency and is not a test of variance components. Rather it is simply a test of fit of observed and predicted phenotypes. For this reason, CHEVERUD and ROUTMAN (1995) call it "physiological epistasis." The results of all 1052 tests (eight *Pelement* pairs $\times$ 16 traits $\times$ nine phenotypes) appear in Figure 5. The figure shows 160 significant tests (15.2% of the tests were significant). As in the classical parameterization, there appears to be heterogeneity among crosses in the patterns of epistasis. It also appears significant epistatic effects are clustered in the sense that some genotypes exhibit significant epistasis for multiple traits. This is expected given the correlation across these metabolic traits.

Scaling: If genes act in a multiplicative manner, then they may appear to be spuriously epistatic on an additive scale. Log transformation will put the gene effects on an appropriate scale and produce the result of additive effects on a log scale. Other transformations besides logarithmic may successfully remove the apparent epistasis seen on the additive scale. Figure 2 shows, however,

FIGURE 3.—Distributions of F -statistics obtained from the eight P -element pairs $\times$ 16 traits = 128 analyses of variance. Symbols are for the Cockerham parameterization: additive (A), dominance (D), additive $\times$ additive (A $\times$ A), additive $\times$ dominance and dominance $\times$ additive pooled (A $\times$ D), and dominance $\times$ dominance (D $\times$ D).

that it is unlikely that any transformation will remove epistatic effects from many of the tests, because no transformation will eliminate the crossing of the lines seen in this figure. Crossing lines indicate departure from additivity on any scale. Nevertheless, we tried log transforming the data, to see how much it affected the ANOVA test of epistasis or the fits to COCKERHAM and CHEVERUD's models. Whereas 35 of 128 tests of the interaction term in the ANOVA were significant on an additive scale, 37 of 128 were significant on the log scale.

FIGURE 4.—Parameters of COCKERHAM's model that deviated significantly from zero (at $P < 0.05$) are plotted as black rectangles. 1–8, sets of P -element pairs. Within each set the four rows of the table correspond to A $\times$ A, A $\times$ D, D $\times$ A, and D $\times$ D epistasis.

DISCUSSION

The following five conclusions can be drawn from the data presented here: (1) There was heterogeneity among P -element pairs in the extent of epistasis exhibited. (2) The A $\times$ D epistasis was directional in the sense that if A $\times$ D was significant for a given trait and P -element pair, other traits for that P -element pair tended to have A $\times$ D and not D $\times$ A epistasis. (3) The distribution of significant components of COCKERHAM's tests is not random. That is, if a particular component is significant for one trait, it is more likely to be significant for other traits as well. This is caused in part by the correlation across traits. (4) Similarly, if one genotype appears aberrant in one phenotype in the test of physiological epistasis, that same genotype is likely to be aberrant in other traits as well. (5) The COCKERHAM parameterization of epistasis and CHEVERUD's physiological epistasis give different perspectives on the pattern of epistasis.

The two ways to consider epistasis are not independent, because if there is no significant interaction variance, then that implies that none of the individual genotypes can deviate significantly from additivity across

FIGURE 5.—Phenotypes that deviated significantly from the nonepistatic model of CHEVERUD and ROUTMAN (1995) are plotted as black rectangles. Within each of the eight *P*-element pair blocks, the nine rows correspond to the nine genotypes *AABb*, *AABb*, *AAbb*, *AaBB*, *AaBb*, *Aabb*, *aaBB*, *aaBb*, and *aabb*.

loci. Formal statistical testing for epistasis should be done following COCKERHAM's approach, since it provides a full partitioning of the interaction variance whether the sample is from a HARDY-WEINBERG equilibrium population or a sample of measured genotypes. On the other hand, the idea of identifying which particular genotypes seem to be responsible for the departure from additivity has some biological merit. This analysis made it clear that the departures did not tend to occur in the doubly homozygous genotypes, but were distributed among the nine genotypes fairly evenly (Figure 5).

It is important to ask whether the epistasis uncovered in this study is caused by the particular *P*-element insertions, and there are two reasons to be cautious in making this claim. First, the independent constructs of particular *P*-element genotypes were found to be significantly heterogeneous in the majority of cases. Such heterogeneity may have arisen from additional mutations that accumulated in the *P*-element-bearing lines between the time of their original construction and the time the double-insert stocks were made (~13 months). It is possible that some of the observed epistatic variance was generated by this unidentified background varia-

tion. But the analysis used the among-construct variance in assessing the significance of epistasis, so the differences among *P*-element genotypes were significantly greater than this residual variation. In addition, the *P*-elements bear a copy of the *white* minigene, and some of the interaction may be caused by the differential expression of *white*. This possibility cannot be totally excluded, but it is not very plausible to argue that *white* expression could affect interactions in expression of more than a few of the diverse array of metabolic traits as was examined here.

It may be meaningful that epistasis is particularly prevalent in crosses among most widely divergent races and among interspecific crosses (WHITLOCK *et al.* 1995). Particularly good examples are seen in the genes that cause interspecific hybrid male sterility in the *Drosophila melanogaster* subgroup (CABOT *et al.* 1994; PALOPOLI and WU 1994). This may be because polymorphism for strong interactions are not retained in populations for some reason, perhaps because of their frequent lack of global stability. Or perhaps the wide crosses result in stronger effects at all levels, including pleiotropic effects which ultimately result in apparent interactions, even if the interactions are very indirect. In either case, the consequences for gene dynamics and response to selection are real and need to be accommodated in population genetics models.

The use of measured genotypes, identified either from anonymous markers in the genome or through controlled crosses, has greatly increased our ability to resolve and quantify epistatic interactions. Unlike classical quantitative genetic methods of inferring epistasis, measured genotypes allow tests of epistasis that nearly match the statistical power of tests of additivity and dominance. EDWARDS *et al.* (1987) examined two sets of F_2 maize populations, scoring 20 markers and 82 traits. They found QTLs for every trait examined, extensive overdominance, but relatively little digenic epistasis (~3% of interaction terms in a two-way ANOVA were significant). This result was consistent with classical estimates of genetic variance components in maize, which generally show little epistasis. Measured genotype methods in fact have revealed relatively little epistasis in any plant species (TANKSLEY 1993). One notable exception is found in the crosses between maize and teosinte (DOEBLEY and STEC 1991; DOEBLEY *et al.* 1995). Here quite strong epistasis is found in several characters. This is another example of wide crosses showing more epistatic interaction than crosses between similar lines. In *Drosophila*, it appears that some "near" crosses can yield pronounced levels of epistasis. LONG *et al.* (1995) were able to test pairwise interactions on bristle phenotypes among genomic regions marked by *roo* transposable element insertions. The design was an analysis of F_2 flies derived from crosses of high- and low-bristle-number selected lines. Of the 60 tests that they could perform, 13 exhibited significant epistasis. They also

found that the magnitude of epistatic effects was on par with the additive effects. It seems that the measured genotype approach provides a good tool for quantifying gene interactions for a variety of traits.

It may also be informative to consider the sorts of traits that tend to be more highly epistatic. At some level, it seems obvious that the determinants of an individual gene's expression must be highly nonadditive. The assembly of a transcription complex occurs only through delicate pairwise and multiway interactions among a large (and growing) set of transcription factors and related proteins. It follows that traits that are likely to be manifestations of expression of one or a few genes, are likely to be epistatic because genetic variation in the factors that result in that gene's expression are likely to interact. Our results, showing extensive epistasis among *P*-element inserts that affect enzyme activities, is further supported by the study of DAMERVAL *et al.* (1994). Using two-dimensional gel electrophoresis, they examined expression of 72 proteins in 85 F₂ maize plants from a distant cross and assembled a genetic map with 118 markers. Their finding that 14% of the proteins exhibited significant epistasis between genomic regions appears to be independent of problems of multicollinearity in this set of data. Recall that morphological traits in maize have repeatedly exhibited little epistasis, yet traits close to the expression of individual genes appear to exhibit significant levels of epistasis.

Finally, we emphasize that our study examined epistatic effects of new mutations caused by *P*-element insertion. Such effects are likely to be different from the distribution of epistatic effects in an equilibrium population. Large differences between levels of mutational epistasis and equilibrium epistasis may provide evidence for the action of natural selection removing a nonrandom set of alleles from the population. Of course, natural selection for a two-gene system may be complex, especially if one considers pleiotropic effects of mutations on different life-stage components (CLARK and FELDMAN 1981). Measured genotypes provide the opportunity to directly estimate both distributions of mutational and equilibrium effects, and efforts are underway to obtain parallel estimates of epistatic effects from natural populations.

We thank JAMES ROSINSKI for technical assistance. This work was supported by National Science Foundation grant DEB-9419631.

LITERATURE CITED

- BRYANT, E. H., and L. M. MEFFERT, 1993 The effect of serial founder-flush cycles on quantitative genetic variation in the housefly. *Heredity* **70**: 122–129.
- CABOT, E. L., A. W. DAVIS, N. A. JOHNSON and C.-I. WU, 1994 Genetics of reproductive isolation in the *D. simulans* clade: complex epistasis underlying hybrid male sterility. *Genetics* **137**: 175–189.
- CHARLESWORTH, D., M. T. MORGAN and B. CHARLESWORTH, 1993 Mutation accumulation in finite outbreeding and inbreeding populations. *Genet. Res.* **61**: 39–56.
- CHEVERUD, J. M., and E. J. ROUTMAN, 1995 Epistasis and its contribution to genetic variance components. *Genetics* **139**: 1455–1461.
- CLARK, A. G., and M. W. FELDMAN, 1981 The estimation of epistasis in components of fitness in experimental populations of *Drosophila melanogaster*. II. Assessment of meiotic drive, viability, fecundity and sexual selection. *Heredity* **46**: 347–377.
- CLARK, A. G., and L. WANG, 1994 Comparative analysis of metabolism in nine species of *Drosophila*. *Evolution* **48**: 1230–1243.
- CLARK, A. G., L. WANG and T. HULLEBERG, 1995 *P*-element induced variation in metabolic regulation in *Drosophila*. *Genetics* **139**: 337–348.
- COCKERHAM, C. C., 1954 An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* **39**: 859–882.
- COOLEY, L., R. KELLEY and A. SPRADLING, 1988 Insertional mutagenesis of the *Drosophila* genome with single *P* elements. *Science* **239**: 1121–1128.
- DAMERVAL, C., A. MAURICE, J. M. JOSSE and D. DE VIENNE, 1994 Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of genome expression. *Genetics* **137**: 289–301.
- DOEBLEY, J., and A. STEC, 1991 Genetic analysis of the morphological differences between maize and teosinte. *Genetics* **129**: 285–295.
- DOEBLEY, J., A. STEC and C. GUSTUS, 1995 *teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* **141**: 333–346.
- EDWARDS, M. D., C. W. STUBER and J. F. WENDEL, 1987 Molecular-marker facilitated investigations of quantitative trait loci in maize I. Numbers, genomic distribution and types of gene action. *Genetics* **116**: 113–125.
- FRERIKSEN, A. B. L. A. DE RUITER, W. SCHARLOO and P. W. H. HEINSTR, 1994 *Drosophila* alcohol dehydrogenase polymorphism and carbon-13 fluxes: opportunities for epistasis and natural selection. *Genetics* **137**: 1071–1078.
- GAVRILETS, S., and G. DEJONG, 1993 Pleiotropic models of polygenic variation, stabilizing selection, and epistasis. *Genetics* **134**: 609–625.
- GIBSON, G., 1996 Epistasis and pleiotropy as natural properties of transcriptional regulation. *Theor. Popul. Biol.* **49**: 58–89.
- GOODNIGHT, C. J., 1988 Epistasis and the effect of founder events on the additive genetic variance. *Evolution* **42**: 441–454.
- KARLIN, S., 1975 General two-locus selection models: some objectives, results and interpretations. *Theor. Popul. Biol.* **7**: 364–398.
- LONG, A. D., S. L. MULLANEY, L. A. REID, J. D. FRY, C. H. LANGLEY *et al.*, 1995 High resolution mapping of genetic factors affecting abdominal bristle number in *D. melanogaster*. *Genetics* **139**: 1273–1291.
- PALOPOLI, M. F., and C.-I. WU, 1994 Genetics of hybrid male sterility between *Drosophila* sibling species: a complex web of epistasis is revealed in interspecific studies. *Genetics* **138**: 329–341.
- ROBERTSON, H. M., C. R. PRESTON, R. W. PHILLIS, D. M. JOHNSON-SCHLITZ, W. K. BENZ *et al.*, 1988 A stable genomic source of *P*-element transposase in *Drosophila melanogaster*. *Genetics* **118**: 461–470.
- TANKSLEY, D. S., 1993 Mapping polygenes. *Annu. Rev. Genet.* **27**: 205–233.
- TURELLI, M., 1985 Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* **111**: 165–195.
- WHITLOCK, M. C., P. C. PHILLIPS, F. B. G. MOORE and S. J. TONSOR, 1995 Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.* **26**: 601–629.
- WILLIS, J. H., and H. ALLEN ORR, 1993 Increased heritable variation following population bottlenecks: the role of dominance. *Evolution* **47**: 949–956.

Communicating editor: D. CHARLESWORTH