A Comparison of the Genetic Basis of Wing Size Divergence in Three Parallel Body Size Clines of *Drosophila melanogaster*

A. Stuart Gilchrist and Linda Partridge

Department of Biology, Galton Laboratory, University College London, London NW1 2HE, United Kingdom

Manuscript received November 30, 1998
Accepted for publication August 16, 1999

ABSTRACT

Body size clines in *Drosophila melanogaster* have been documented in both Australia and South America, and may exist in Southern Africa. We crossed flies from the northern and southern ends of each of these clines to produce F_1, F_2, and first backcross generations. Our analysis of generation means for wing area and wing length produced estimates of the additive, dominance, epistatic, and maternal effects underlying divergence within each cline. For both females and males of all three clines, the generation means were adequately described by these parameters, indicating that linkage and higher order interactions did not contribute significantly to wing size divergence. Marked differences were apparent between the clines in the occurrence and magnitude of the significant genetic parameters. No cline was adequately described by a simple additive-dominance model, and significant epistatic and maternal effects occurred in most, but not all, of the clines. Generation variances were also analyzed. Only one cline was described sufficiently by a simple additive variance model, indicating significant epistatic, maternal, or linkage effects in the remaining two clines. The diversity in genetic architecture of the clines suggests that natural selection has produced similar phenotypic divergence by different combinations of gene action and interaction.

THE genetics underlying the phenotypic evolution and divergence of populations of the same species has been a long-studied topic in evolutionary biology. The models developed by Fisher and Wright have provided the conceptual framework for most investigations of the topic. The Fisherian view is characterized as stressing the role of additive variance in evolution. Under his fundamental theorem, increase in fitness is proportional to the additive variance present in a population. Because large undivided populations have maximum additive variance, evolution is expected to proceed faster in large undivided populations. Wright provided an alternative view. His three-phase shifting balance theory envisaged evolution as occurring via processes of isolation and drift, intrademe and interdeme selection (Wright 1977). Under this theory, interaction between loci, or epistasis, plays a major role in producing the different fitness peaks reached by different populations.

To gauge the likely generality of the shifting balance theory, the occurrence of each of its component parts has been investigated. The roles of population structure and drift in evolution have been the subject of many studies over a long period. However, the occurrence and magnitude of epistatic variance has, until recently, received less attention for two main reasons (Whitlock *et al.* 1995; Fenster *et al.* 1997). First, especially in relation to fitness, most relevant interacting loci will be near fixation in a given interbreeding population, making their effects difficult to detect. Second, there are statistical difficulties associated with partitioning epistatic variance.

Nevertheless, a variety of experimental approaches are available that infer or measure epistatic interactions (summarized in Fenster *et al.* 1997). These methodologies measure either epistatic variances or the epistatic contribution to phenotypic means. The measurement of epistatic variances is valuable as it gives an indication of the importance of epistasis in any future short-term response to selection. However, the available methods suffer from low statistical power. The alternative is to measure the contribution of epistasis to current phenotypic means, using methods such as outbreeding depression and line-cross analysis. These methods are widely applicable and statistically powerful. However, although the finding of epistatic effects on the current population mean suggests that epistatic interactions were important in the evolution of populations, it is not direct evidence for past selection for those same coadapted gene complexes (Fenster *et al.* 1997).

In our experiments, we chose to use the second methodology, *i.e.*, the measurement of epistatic contributions to current phenotypic means. The two common experimental approaches employed to infer epistatic effects on mean phenotypes are outbreeding depression (or F_2 breakdown) and line-cross analysis. The common factor in both methods is that divergent populations are

Corresponding author: A. Stuart Gilchrist, Department of Biology, Galton Laboratory, Wolfson House, University College London, 4 Stephenson Way, London NW1 2HE, United Kingdom.
E-mail: a.gilchrist@ucl.ac.uk

crossed and the resulting offspring generation means measured. Outbreeding depression is expected where coadapted gene complexes present in the parental populations are disrupted by recombination in the F_2 generation. Line-cross analysis takes this a step further by comparing the observed means of a variety of hybrid generations with expected means (Mather and Jinks 1982). The expected means are derived from genetic models containing additive, dominance, digenic interaction, and maternal effects. More recently, similar expectations for additional generations, which incorporate inbreeding effects, have been derived (Lynch 1991). In the language of the shifting balance theory, outbreeding depression and line-cross analysis both attempt to infer epistatically generated peaks (the parental populations) by examining the genetic behavior of hybrid populations that are cast in the presumptive valleys between those peaks.

A number of instances of both outbreeding depression and epistatic effects on means have been reported. For example, outbreeding depression in the F_2 generation has been shown to affect fecundity in hybrids of *Drosophila pseudoobscura* from the western United States (Vetukhiv 1956), development time in a marine copepod, *Tigriopus californicus* (Burton 1990), and capacity for increase in the pitcher-plant mosquito, *Wyeomyia smithii* (Armbruster *et al.* 1997). Line-cross analysis in a variety of species has shown that epistatic interactions are the rule rather than the exception in crosses between divergent parental lines. Examples based on divergence between artificially selected parental lines include growth time in maize (Mohamed 1959) and body weight in mice (Chai 1956). Examples from populations where divergence has occurred under natural selection include photoperiod in *W. smithii* (Hard *et al.* 1992) and ovariole number in *D. hibisci* (Starmer *et al.* 1998; for further examples see Lynch and Walsh 1998).

Studies such as these that have actually measured epistatic parameters (*i.e.*, employed line-cross analysis) have usually found them. However, such studies are relatively rare, so we still do not know how common or how necessary epistatic interactions are in the process of population divergence. Furthermore, none of the studies of populations that have diverged under natural selection have had the opportunity to measure independent examples of parallel divergence. Without independent replicates, we can only guess at the relative importance of factors such as epistasis, mutational order, founder effects, and chance in the divergence of populations. We know of no measurements on natural populations with the necessary independent replication to answer this question. A potential source of such replication are the continental clines observed in various Drosophilid species. Clines have been identified in a number of characters, from complex traits such as body size (see below) and ovariole number (Starmer *et al.* 1998) to individual molecular markers such as inversions [*e.g.*, *In(2L)t*; Stalker 1976] and enzymes (*e.g.*, *Adh*; Oakeshott *et al.* 1982). A genetic cline can result from adaptation, drift, or particular historical circumstances (Endler 1977). However, the finding of similar clines for the same character on different continents is strong evidence that they are maintained by natural selection.

Previous studies of genetic divergence of clinal populations have been limited to single continents because of range limitations of the species under study, *e.g.*, photoperiodism in the North American pitcher-plant mosquito, *W. smithii* (Hard *et al.* 1992), fitness in *W. smithii* (Armbruster *et al.* 1997), and ovariole number in the Australian *D. hibisci* (Starmer *et al.* 1998). In each of these studies, it was found that two-locus epistatic interactions, in addition to additive and dominance effects, played a significant role in the genetic divergence of the populations. The most relevant study is that of Armbruster *et al.* (1997), who studied divergence in fitness in *W. smithii* in the eastern United States. They found different genetic architectures in crosses between two northern-derived populations and two southern parental populations.

In this article we investigated the genetic parameters of wing size for three parallel body size clines in *D. melanogaster* from the southern hemisphere. Wing length was measured because it is highly correlated with thorax length and has been used in many earlier studies of body size ($r = 0.7$; Reeve and Robertson 1952; Robertson 1959). However, because overall wing size is the product of complex interactions between different compartments (Guerra *et al.* 1997), wing area was also measured, as not all variation may be accounted for by wing length. Large and small populations from the extremes of each cline were crossed to produce F_1, F_2, and first backcross generations. Together with the parents and reciprocals of each generation, these crosses provide sufficient data (14 generation means) to test the adequacy of genetic models of expected generation means containing additive, dominance, maternal, cytoplasmic, and digenic epistatic effects (Mather and Jinks 1982). For the generation variances, a simpler model containing only additive and segregational parameters was tested. Assuming that the clines represent independent examples of similar natural selective forces, we were then able to compare the genetic basis of wing size divergence on the three continents, with particular emphasis on the occurrence of epistatic interactions. Our measurements of maternal effects also allowed us to assess the validity of using F_2 breakdown as an indicator of the presence of epistasis. Maternal effects can contribute significantly to F_2 breakdown, yet are often assumed to be absent. Because we were able to quantify both the epistatic and maternal effects present in each cline, we were able to test the validity of this assumption and its implications for the detection of epistasis by F_2 breakdown.

MATERIALS AND METHODS

Fly populations: The six populations of *D. melanogaster* used in this study were chosen to represent the ends of three parallel body size clines, all in the southern hemisphere. Populations with genetically larger body size are found at more southerly latitudes. The first of these clines is found along the east coast of Australia and has been previously described in James *et al.* (1995). The two Australian populations used were a larger body size southern population Cygnet (Cyg, 43.08°S) and a smaller northern population Innisfail (Inn, 17.30°S). These two populations were collected in January 1997 and were a gift from Dr. Ary Hoffmann, La Trobe University, Melbourne.

The second cline is found along the west coast of South America and has also been previously described in Van't Land *et al.* (1995). The smaller body size population is from Guayaquil, Ecuador (Gu, 2°13'S) and the larger is from Porto Montt, Chile (Pm, 41°30'S). These two populations were collected in early 1995. Samples of these populations were provided by Dr. Jan Van't Land.

The third pair of populations were chosen in the expectation that a body size cline would also exist between equatorial Africa and southern Africa. Populations from Kenya and South Africa, originally collected as isofemale lines, were from the stock collection of Professor Chip Aquadro, Cornell University. The wing areas of these populations were measured and the southerly Capetown population (Cape, ~34°S) was found to have significantly larger wing area than the northern Kenya population (Kenya, approximately equatorial). This difference is consistent with the existence of a possible third body size cline.

All stocks were maintained as expanded bottle stocks at 25° on a 12:12-hr light:dark cycle on a standard cornmeal/yeast/sugar medium. While the Australian and South American flies were collected en masse, the African flies were descended from isofemale lines. Despite their earlier disparate culture histories and possible consequent variation in lab adaptation, these populations had maintained their original differences in body size under lab culture (see James *et al.* 1997).

Crosses: To exclude the possibility that hybrid dysgenesis affected the means and variances of the crosses, a simple test was carried out before the crosses were performed (Roberts 1986). Flies from cline ends were mated in reciprocal crosses, progeny raised at 25°, and 20 individuals of each sex dissected. The complete absence of unilateral or bilateral hypertropy of gonads indicated that all stocks were of identical cytotype with respect to the P-M system.

For each cline, the northern and southern populations were crossed to produce the six basic generations, *i.e.*, the two parental populations P_1 and P_2, their F_1 and F_2 generations, and the two backcross generations B_1 and B_2 (*i.e.*, $F_1 \times P_1$ and $F_1 \times P_2$). Reciprocals of all crosses were also established, denoted as F_1R, F_2R, B_1R, and B_2R. For each backcross generation, separate generations were raised in which the F_1 parent differed reciprocally [*e.g.*, $(P_1 \times P_2) \times P_1$ and $(P_2 \times P_1) \times P_1$]. These additional backcross generations were denoted as B_1a, B_1b, B_1Ra, B_1Rb, etc. In total, 14 distinct generations were raised from each cline (Table 1). All crosses were set up with at least 40 virgin individuals of each sex as parents. Crosses that provided progeny for measurement were allowed to oviposit on grape juice/agar medium. This allowed progeny to be picked as first instar larvae, which were then transferred to new vials of unyeasted standard medium at a constant density of 50 larvae per vial, conditions under which competition is minimal and body size maximized. Six replicate vials were established from each cross. After 14 days, all flies that had emerged in each vial were frozen for later measurement. All flies that were measured were reared simultaneously on the same batch of food medium.

Wing measurement: The right wings of all flies from each experimental vial were removed and mounted on microscope slides in Aquamount (left wings were used if the right was damaged). Wing images were captured using a compound microscope, with low power objective (2.5×) and attached video camera, connected to a Macintosh computer. The area of each wing was measured using the Object-Image program (by Norbert Vischer, based on the public domain NIH Image and available at http://simon.bio.uva.nl/object-image.html) to calculate areas and record coordinates of all landmarks. The area measured consisted of a polygon whose vertices were the humeral-costal break, the distal ends of longitudinal veins L2–5, and the base of the alula, as shown in Figure 1. This polygon area measurement is highly correlated with wing area (WA) as measured by tracing an outline on a graphics tablet ($r = 0.95$, data not shown). The polygon method is considerably faster and more reproducible than outline tracing. Using the landmark coordinates, wing length (WL) was also calculated. We report results for WL, in addition to WA, because

TABLE 1

The parameter coefficients used in the model for generation means

	Mather and Jinks notation: / Edinburgh/Iowa notation:	*m* / *m*	*[d]* / *[a]*	*[h]* / *[d]*	*[i]* / *[aa]*	*[j]* / *[ad]*	*[l]* / *[dd]*	*[dm]* / *[am]*	*[hm]* / *[dm]*	*[c]*	*[Y]*
P_1	*e.g.*, Cygnet	1	1	0	1	0	0	1	0	1	1
P_2	*e.g.*, Innisfail	1	−1	0	1	0	0	−1	0	−1	−1
F_1	$P_1 \times P_2$	1	0	1	0	0	1	1	0	1	−1
F_1R	$P_2 \times P_1$	1	0	1	0	0	1	−1	0	−1	1
F_2	$(P_1 \times P_2) \times (P_1 \times P_2)$	1	0	0.5	0	0	0.25	0	1	1	−1
F_2R	$(P_2 \times P_1) \times (P_2 \times P_1)$	1	0	0.5	0	0	0.25	0	1	−1	1
B_1a	$P_1 \times (P_1 \times P_2)$	1	0.5	0.5	0.25	0.25	0.25	1	0	1	−1
B_1b	$P_1 \times (P_2 \times P_1)$	1	0.5	0.5	0.25	0.25	0.25	1	0	1	1
B_1Ra	$(P_1 \times P_2) \times P_1$	1	0.5	0.5	0.25	0.25	0.25	0	1	1	1
B_1Rb	$(P_2 \times P_1) \times P_1$	1	0.5	0.5	0.25	0.25	0.25	0	1	−1	1
B_2a	$P_2 \times (P_1 \times P_2)$	1	−0.5	0.5	0.25	−0.25	0.25	−1	0	−1	−1
B_2b	$P_2 \times (P_2 \times P_1)$	1	−0.5	0.5	0.25	−0.25	0.25	−1	0	−1	1
B_2Ra	$(P_1 \times P_2) \times P_2$	1	−0.5	0.5	0.25	−0.25	0.25	0	1	1	−1
B_2Rb	$(P_2 \times P_1) \times P_2$	1	−0.5	0.5	0.25	−0.25	0.25	0	1	−1	−1

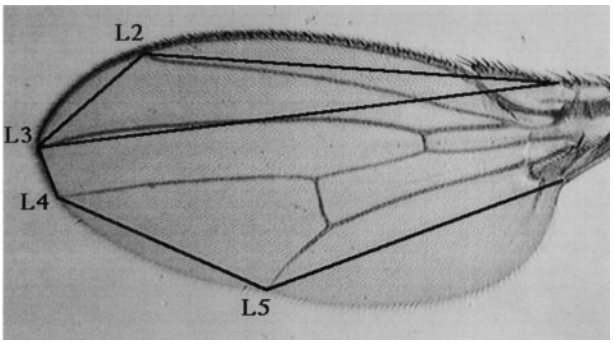

Figure 1.—Wing measurements used in these experiments. L2–L5 indicate the ends of the second to fifth longitudinal veins. The outline superimposed on the wing joins the six points used to determine wing area, while the line joining the humeral-costal break to the end of L3 shows the measurement used to calculate wing length.

it provides an alternate, linear, index of wing size. Also it is not clear whether WA or other traits such as aspect ratio (WL^2/wing width) or wing:thorax size ratios are the principal targets of natural selection (Azevedo *et al.* 1998).

For each generation, wings of flies from up to six replicate vials were mounted and measured (mean of 5.5 replicate vials/generation with a mean of 22–23 flies of each sex per vial). Standard errors of generation means were calculated to take into account variation both within and between the vials contributing to each generation mean. For each cline and sex, variance components among and within vials were calculated by a nested ANOVA (*i.e.*, with generation and vial nested within generation as main effects). The sampling variance for each generation was then calculated as

$$V_{\text{between}} / n_{\text{vial}} + V_{\text{within}} / n_{\text{individuals}},$$

where V_{between} and V_{within} are the between- and within-vial variance components, respectively, and n_{vial} and $n_{\text{individuals}}$ are the numbers of vials and individuals in each generation (typically $n_{\text{vial}} = 5$ or 6, while $n_{\text{individuals}}$ averaged 128). By contrast, mean standard errors calculated by simply pooling flies from all vials within each generation underestimate the sampling variance (by anything up to 50% in our data).

Analysis of generation means: All analyses were performed separately for each sex on untransformed data (measured in square millimeters for WA and millimeters for WL). Square root and log transformations had little effect on the resulting models for WA and are not reported.

The generation means were analyzed by the methods of Mather and Jinks (1982). First, the observed generation means were used to estimate the parameters of a model consisting only of an overall mean *m* and composite additive *[a]* and dominance *[d]* genetic effects (Hayman 1958; Mather and Jinks 1982; notation follows Falconer 1989; as adopted by Kearsey and Pooni 1996). The estimated parameters were used in turn to calculate the expected generation means. The goodness of fit between observed and expected generation means was then examined, a significant χ^2 indicating a significant difference between the observed and expected generation means, which implied that a simple additive-dominance model was insufficient to explain the data. This method equates to the joint-scaling test of Cavalli (Mather and Jinks 1982).

If the additive-dominance model was found to be insufficient, then further parameters were added following Tables 11.4 and 13.2 in Kearsey and Pooni (1996). These extra parameters were, first, composite digenic epistatic effects (additive × additive *[aa]*, additive × dominance *[ad]* and dominance × dominance *[dd]*) and, second, parameters to account for maternal effects, and finally, a parameter to account for Y-linked effects *[Y]*. The maternal effect parameters were additive maternal and dominance maternal effects ($[a]_m$ and $[d]_m$) and cytoplasmic effects *[c]*, which account for mitochondrial genetic effects, symbionts, or infectious agents (*e.g.*, viruses; Thomas-Orillard 1984). The model parameters are shown in Table 1. From the model parameters, it can be seen that $[a]_m$ and $[d]_m$ account for genetic effects of the mothers on the means of their progeny. $[a]_m$ effects are only apparent in generations where the mother is from the P_1 or P_2 population. Conversely, $[d]_m$ effects are only apparent in F_1 hybrid mothers, reflecting the dominance interaction between different alleles from the parental populations that affect progeny means. From Table 1, it can also be seen that the presence of cytoplasmic effects *[c]* depends simply on the line of descent of the cytoplasm through the maternal line. In contrast to $[a]_m$ and $[d]_m$, cytoplasmic effects affect all generations and could persist for many generations (*e.g.*, as found by Cavicchi *et al.* 1989). Parameters accounting for interactions between additive and dominance effects in the progeny and maternal effects ($a.a_m$, $a.d_m$, $d.a_m$, and $d.d_m$) were also tested, but were not found to significantly improve the fit of any of the models. A further two parameters could have affected the generation means. The first accounts for X chromosome effects and the second could be included to account for the absence of recombination in males. From Table 1, it can be seen that the lack of recombination in males would affect the difference between generation pairings such as B_1a and B_1Ra. Neither of these parameters were found to improve the fit of any of the models and are not reported.

The coefficients used in our analysis followed Mather and Jinks (1982) in being based on an $F_{-\infty}$ metric, as opposed to the F_2 metric used in Lynch and Walsh (1998). In models based on an $F_{-\infty}$ metric, the mean in any model corresponds to the mean that would be expected after a large number of generations of inbreeding, whereas for an F_2 metric the expected mean corresponds to the mean observed in the F_2 generation. Although the coefficients for the parameters vary between the two systems, they produce similar parameter estimates.

Using an expanded model, expected generation means were then recalculated from the new parameter estimates. Because there were 14 generations, models could, in theory, contain up to 14 parameters. However, model fitting proceeded by adding only the digenic, maternal, and cytoplasmic effects. The significance of each of the extra parameters with respect to their standard errors gave an indication of which parameters could be omitted to simplify the model. After the least significant parameters were removed, the expected generation means were recalculated, and the new goodness of fit tested. In this way, models were constructed that contained the minimum number of parameters necessary to explain the observed generation means. To assess the importance of digenic epistatic parameters, their contribution was tested by comparing models before and after the removal of each digenic epistatic parameter. In all cases, the estimated digenic parameters were found to significantly improve the fit of the model.

The model parameters, errors, and χ^2 values were estimated using weighted least-squares methods (Mather and Jinks 1982; Kearsey and Pooni 1996; Lynch and Walsh 1998). A vector of parameters, $\hat{\mathbf{y}}$, was calculated using

$$\hat{\mathbf{y}} = (\mathbf{C}^T\mathbf{V}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{V}^{-1}\mathbf{x},$$

where **C** is the matrix of coefficients of the parameters of

the expected generation means, **V** is the diagonal matrix of sampling variances of each line mean (calculated as described above), and **x** is the vector of observed line means. The standard error of each parameter is obtained from the square root of the corresponding diagonal element of the sampling covariance matrix **S**

$$\mathbf{S} = (\mathbf{C}^T\mathbf{V}^{-1}\mathbf{C})^{-1}.$$

The χ^2 used to determine the goodness of fit of each model was calculated as

$$(\mathbf{x} - \hat{\mathbf{x}})^T\mathbf{V}^{-1}(\mathbf{x} - \hat{\mathbf{x}})$$

(Lynch and Walsh 1998), where $\hat{\mathbf{x}}$ is the vector of expected generation means calculated as $\hat{\mathbf{x}} = \mathbf{C}\hat{\mathbf{y}}$, with the degrees of freedom equal to the number of generation means minus number of parameters in the model.

A significant improvement in the goodness of fit was measured using a likelihood-ratio test

$$\Lambda = \chi^2_{\text{initial}} - \chi^2_{\text{enlarged}},$$

which, if significant, indicates an improved fit (Lynch and Walsh 1998). This equation is sufficient for the large sample sizes used here. The degrees of freedom are equal to the difference in the number of parameters between the models (*i.e.*, one where the effect of adding individual epistatic parameters was tested).

Analysis of generation variances: The analysis of generation variances was mathematically similar to that of the generation means. A weighted least-squares procedure was again used to provide parameter estimates, which, in turn, allowed the goodness of fit of the model to be tested.

The principal differences from the analysis of generation means were, first, that the sampling variance of the generation variances cannot be estimated from the data as was possible for the generation means. Instead, an iterative process is required, with the diagonal elements of **V** estimated initially as $2v^2/n$, where v is the observed generation variance and n the number of individuals (Hayman 1960; Mather and Jinks 1982). In subsequent iterations, the new estimates of the generation variances were used to calculate new weights, *i.e.*, diagonal elements of **V**. Although parameter estimates usually stabilized with 4 or 5 iterations, 10 iterations were performed for each model.

The second difference from the generation means analysis was in the formulas used in the model. Because the parental lines were not completely homozygous, expectations of a model for the generation variances in terms of V_a, V_d, etc. are not straightforward because the parental variances cannot simply be assumed to be due entirely to environmental variance. To circumvent this problem, the method of Lynch and Walsh (1998) was used, which uses the parental additive genetic variances σ^2_{A1} and σ^2_{A2}, and the segregational variance σ^2_s, as parameters. The model predicts the generation variances based on a simple additive model for the trait. As with the generation means, these predictions can be tested for goodness of fit against the observed variances. A significant χ^2 indicates only that the additive model is insufficient to explain the observed data. The statistical power of the estimation of variance parameters is relatively poor (in comparison to the generation means analysis) because the sampling errors associated with variance estimates are relatively large. Even with testable models, it would be difficult to distinguish more complex predictive models (Kearsey and Pooni 1996; Lynch and Walsh 1998). Furthermore, some models of epistatic interactions produce mainly additive genetic variance (Keightley 1989), decreasing the likelihood of detecting nonzero epistatic variance parameters. Consequently, further analysis of variance was not pursued in our experiments.

RESULTS

Generation means analysis: In analyses excluding maternal effects, it is common to pool reciprocal generation means. The only generations that potentially could be pooled in our analysis were the B_1a/B_1b and B_2a/B_2b pairs, because neither pair differs in their maternal parameters (*i.e.*, $[a]_m$, $[d]_m$, and $[c]$). However, because significant differences were observed between approximately half of the reciprocal pairs of generation means, no generation means were pooled and maternal effects were included in the analysis.

The mean wing areas of both sexes of all 14 generations from each of the three clines are shown in Figure 2. The overall difference in wing area between the large and small parental populations of each cline is similar (note that in Figure 2 the coordinate axes are drawn on the same scale for each sex). The regression line shown on each graph is the weighted least-squares estimate of a simple genetic model containing only the overall mean and additive effects, *i.e.*, area = mean + additive effects. In no case is this simple model sufficient to describe the observed means (minimum $\chi^2 = 57.48$, $P < 0.001$ for Australian male WA).

Models of generation means were fitted for both WA and WL. For each character, the model with the fewest parameters producing the best fit to the observed generation means (determined by nonsignificant χ^2 values) is shown in Tables 2 and 3. Because WA and WL are genetically correlated characters, similarity between the models is expected within clines and sexes, allowing more confidence to be attached to general conclusions regarding the presence or absence of particular genetic effects in each cline. A number of general observations can be made about the models presented.

First, none of the clines was adequately described by a simple additive-dominance model. The closest fit between the model and observed generation means was for the African female cline (minimum $\chi^2 = 27.03$, $P = 0.0045$, WA). In this case only one additional parameter ($[aa]$) was necessary for the model to sufficiently describe the observed means (*i.e.*, mean, additive, dominance, and dominance maternal effects).

When digenic epistatic and maternal effects were included in the models, all 12 sets of generation means (3 clines × 2 traits × 2 sexes) were adequately described; *i.e.*, the χ^2 for each model was nonsignificant. Significant χ^2 values would have indicated that linkage and/or higher order interactions (*e.g.*, trigenic interactions) were required to explain the observed generation means. Of all the commonly used experimental organisms, *D. melanogaster* has one of the lowest recombination indices (Lynch and Walsh 1998, p. 211), and it is therefore perhaps surprising that all the observed gener-

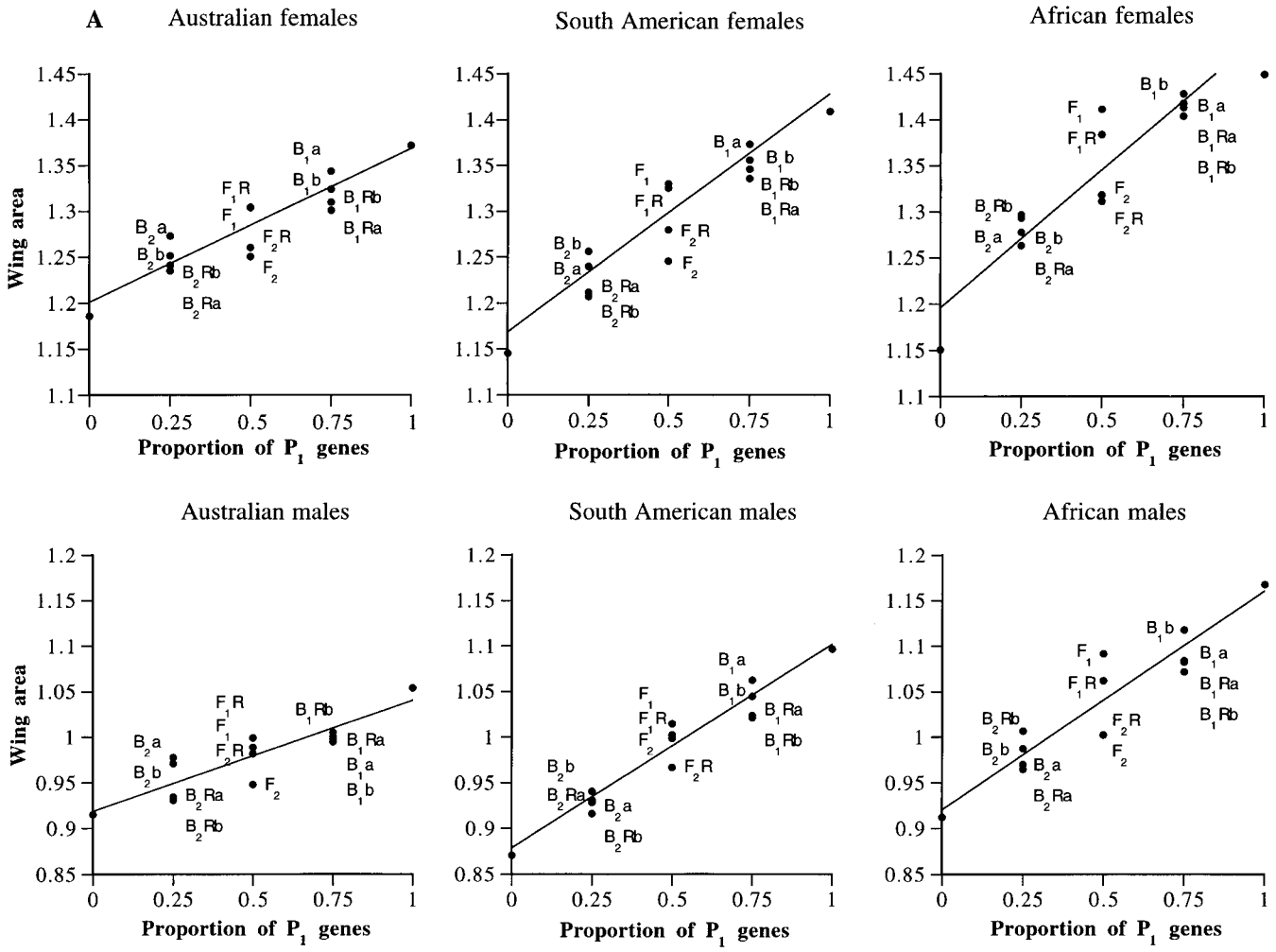

Figure 2.—Mean wing area in square millimeters (A) and wing length in millimeters (B) as a function of the proportion of genes derived from P_1, the larger (southern) parent, in each cross. The solid line is the expectation of a maximum-likelihood additive model. Means for the different reciprocal generations are identified using the notations (B_1a, B_1b, etc.) shown in Table 1. Note that for each sex, the coordinate axes are shown at the same scale, showing the broad overlap of the phenotypic ranges of each character between the clines. For clarity, standard errors are not shown, but the width of the symbols used approximates one standard error in all cases.

ation means could be adequately explained without linkage or higher order interactions.

Second, digenic epistatic interactions were present in both sexes in the Australian and African clines, but notably absent in the South American cline. In the eight models of the Australian and African clines (2 clines × 2 traits × 2 sexes), significant *[aa]* effects were present in six of the models and all positive in sign. Significant *[ad]* and *[dd]* effects were each present only once in the models and were both negative. As noted by Kearsey and Pooni (1996), *[d]* and *[dd]* are always of opposite sign.

Third, additive and dominance maternal effects occurred commonly in the models, although never together. Additive maternal effects were confined to the African cline, while dominance maternal effects were common in the South American and Australian clines. All dominance maternal effects were negative, indicating unfavorable combinations of parental genes in hybrid mothers.

Significant Y effects were found in two of the three male clines. The *[Y]* parameter was not found to be significant in any of the female models. The magnitudes of both the *[c]* and *[Y]* effects were small, and close to the limits of detection for the data. In the case of the South American male clines for both WA and WL, the addition of either a *[c]* or *[Y]* parameter produced a statistically nonsignificant model, but a smaller χ^2 value resulted from the inclusion of the *[Y]* parameter. However, because the effects are small, the biological distinction between cytoplasmic and Y-linked effects should not be stressed. Significant cytoplasmic effects were observed only in the clines for Australian males, where the *[Y]* parameter did not produce a nonsignificant model. No significant cytoplasmic effects were observed in any of the models of female means.

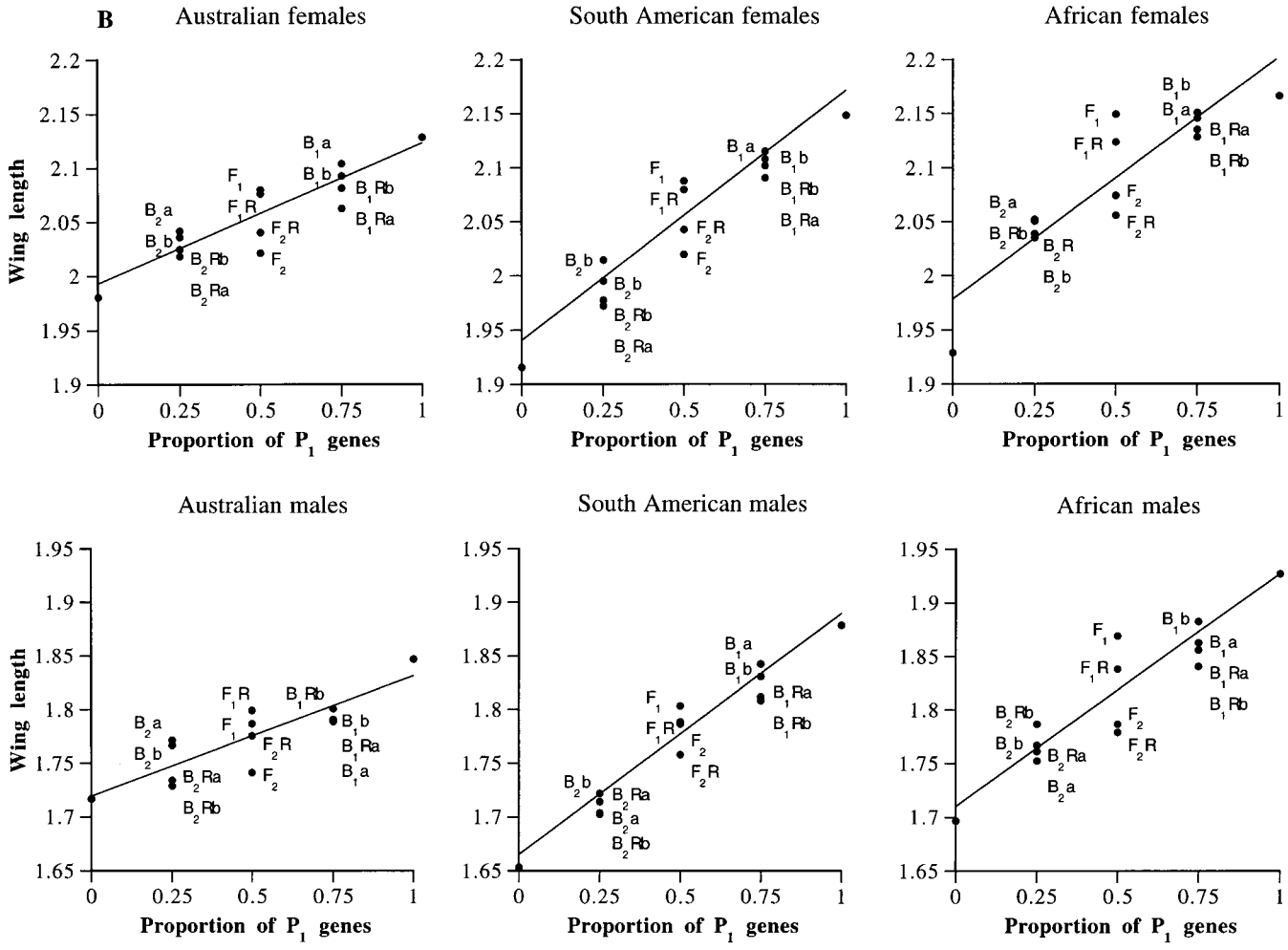

Figure 2.—(*Continued*)

Finally, although all the data could be described adequately using the models shown in Tables 2 and 3, the effect of adding further epistatic parameters was also tested. The results are shown in Tables 2 and 3, where unaccompanied asterisks indicate that the addition of that parameter to the model shown significantly improved the fit of the model (*i.e.*, significantly decreased the χ^2 value). The effect of taking these extra parameters into account is to make the models for WA and WL appear even more similar, indicating that the genetic control of WA and WL is highly integrated.

Generation variance analysis: Figure 3 shows the relation between observed variances for WA and the maximum-likelihood expectations based on an additive model. An additive model adequately described variances for both sexes in the Australian cline. Therefore it appears that there is no need to invoke further variance parameters (dominance, epistatic, and maternal variances) to explain the Australian WA cline. However, there is a strong caveat on this conclusion because the power of these tests is generally low (due to relatively large standard errors: Kearsey and Pooni 1996; Lynch and Walsh 1998). In contrast, in both sexes in the South American and African clines, the additive model was strongly rejected. From Figure 3, it appears that the South American and African clines each failed to fit the additive model for different reasons. In the South American cline, the parental and F_1 variances were generally lower than expected under the additive model, while the reciprocal F_2 variances deviated in both directions from the additive expectation. In the African cline, reciprocal differences are also apparent in the F_2 as well as significant deviations from the additive expectation in the F_1 and backcross variances. Because relationships between generation variances and the underlying variance components (especially epistatic and maternal variances) in the segregating generations were complex, no conclusions can be drawn from the present data regarding likely reasons for the lack of fit between observed variances and the additive expectation (Mather and Jinks 1982; Kearsey and Pooni 1996). For WL, the same pattern was observed, with only the Australian cline being adequately described by the additive model. Graphs corresponding to Figure 3 for WL are not shown,

TABLE 2

Estimates of composite genetic effects underlying divergence in wing area between the southern and northern populations of the three body size clines

		Cline		
Sex	Parameter	Australia	South America	Africa
Females	*m*	1.2819 ± 0.0045***	1.2789 ± 0.0052***	1.1254 ± 0.0146***
	[a]	0.0820 ± 0.0037***	0.1277 ± 0.0042***	0.1368 ± 0.0047***
	[d]	0.0253 ± 0.0072***	0.0505 ± 0.0082***	0.1553 ± 0.0208***
	[aa]			0.0478 ± 0.0165**
	[ad]	*		
	[dd]			***
	$[a_m]$			
	$[d_m]$	−0.0278 ± 0.0040***	−0.0333 ± 0.0043***	
	c			
	χ^2	13.98 NS	16.99 NS	17.42 NS
Males	*m*	0.9352 ± 0.0087***	0.9826 ± 0.0039***	0.9390 ± 0.0118***
	[a]	0.0662 ± 0.0037***	0.1164 ± 0.0037***	0.0808 ± 0.0082***
	[d]	0.0583 ± 0.0120***	0.0255 ± 0.0062***	0.1412 ± 0.0166***
	[aa]	0.0503 ± 0.0102***		0.1020 ± 0.0135***
	[ad]	**		
	[dd]			
	$[a_m]$			0.0254 ± 0.0050***
	$[d_m]$		−0.0196 ± 0.0032***	
	c	−0.0094 ± 0.0020***		
	Y		−0.0065 ± 0.0019**	0.0099 ± 0.0030**
	χ^2	12.88 NS	8.19 NS	11.06 NS

The χ^2 values were calculated using only those parameters whose values are shown. Asterisks without numbers indicate parameters that were not necessary to produce a satisfactory model, but which significantly improved the fit between model and data. The number of asterisks indicates the significance of the improved fit when the parameter was added. NS, not significant. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

but the χ^2 values for each of the models are shown in Table 4.

By way of comparison, a variance analysis was also carried out using a model consisting of four variance components (additive, dominance, additive × dominance, and environmental) as outlined in Kearsey and Pooni (1996). Underlying this analysis is the major assumption that the parental lines are homozygous. Obviously, this assumption does not hold for our clinal populations, yet the analysis produced very similar results to those described for the variance models detailed above. For both WA and WL, a simple additive model (*i.e.*, a model containing environmental variance and additive variance only) adequately described only the Australian cline. No satisfactory variance model could be found for the other clines (data not shown).

DISCUSSION

We examined the genetic basis of body size divergence in the parallel clines, using wing size as an indicator of body size. Analysis of both means and variances of hybrid generations indicated that the occurrence and magnitude of epistatic, maternal, and sex-linked effects vary greatly among the three clines investigated. This finding suggests that markedly different genetic architectures underlie the phenotypic divergence of each cline. If so, our results imply that similar natural selective forces can produce divergence via quite different types of gene action and interaction. In terms of Wright's (1977) shifting balance theory, our results support the notion that isolated populations can reach different adaptive peaks of similar "height," having evolved toward those different peaks by different combinations of gene interactions.

The major question concerning our results is whether the differences between the models reflect significant interclinal differences in the genetic basis of divergence or are simply due to sampling error. Ideally, intracline replicates would be used to estimate variation in models within clines, but such replicates were not included in our survey of wing size divergence.

However, we think it unlikely that the large differences were due mainly to sampling differences. Sampling error could potentially occur at both intra- and interdeme levels. First, at the intrademe level, it is unlikely that sampling bias occurred because size is a polygenic character and the crosses were made from stocks maintained as large outbred populations (with the exception of the African populations). Because large num-

TABLE 3

Estimates of composite genetic effects underlying divergence in wing length between the southern and northern populations of the three body size clines

		Cline		
Sex	Parameter	Australia	South America	Africa
Females	*m*	2.0555 ± 0.0036***	2.0321 ± 0.0041***	1.9292 ± 0.0246***
	[a]	0.0741 ± 0.0041***	0.1154 ± 0.0033***	0.1055 ± 0.0038***
	[d]	0.0233 ± 0.0057***	0.0519 ± 0.0065***	0.3380 ± 0.0565***
	[aa]			0.1143 ± 0.0240***
	[ad]	−0.0380 ± 0.0117**		
	[dd]			−0.1312 ± 0.0349***
	$[a_m]$			
	$[d_m]$	−0.0253 ± 0.0031***	−0.0237 ± 0.0034***	
	c			
	χ^2	14.78 NS	10.77 NS	12.18 NS
Males	*m*	1.7222 ± 0.0083***	1.7630 ± 0.0036***	1.7217 ± 0.0098***
	[a]	0.0626 ± 0.0035***	0.1176 ± 0.0034***	0.0858 ± 0.0046***
	[d]	0.0697 ± 0.0114***	0.0309 ± 0.0057***	0.1365 ± 0.0139***
	[aa]	0.0598 ± 0.0097***		0.0917 ± 0.0112***
	[ad]	**		*
	[dd]			
	$[a_m]$			0.0179 ± 0.0034***
	$[d_m]$		−0.0153 ± 0.0029***	
	c	−0.0098 ± 0.0019***		
	Y		−0.0055 ± 0.0018**	
	χ^2	14.29 NS	11.35 NS	15.16 NS

Explanation of entries as for Table 2. NS, not significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

bers of parents were used in the crosses, it is improbable that the crosses are unrepresentative of the deme sampled. Second, interdeme sampling bias may have occurred; *i.e.*, had collections been made from different northern and southern localities on each continent, the models might have been quite different. This suspicion is supported by the fact that population structuring in *D. melanogaster* seems relatively high (Powell 1997). However, with respect to the genetics of wing size, the only evidence we have suggests relatively little genetic differentiation between demes at the cline ends. This evidence comes from crosses between our southern Australian population (Cygnet) and another population collected ~20 km away (Flowerpot). We raised and measured wing area of parental, F_1, and F_2 generations and calculated genetic models exactly as for the clinal data. For males, there was no evidence of any genetic differentiation (*i.e.*, all generations had the same mean), while for females, only dominance effects were necessary to explain the data. The *C*-scaling tests for nonadditive effects (see below for further discussion) were all negative. Assuming that these two southern Australian populations represent different demes, it seems interdeme differences may be minimal, although only a large-scale survey could confirm this.

The results of other line-cross studies also suggest that models of composite genetic effects are minimally affected by sampling bias. First, Pooni *et al.* (1985) performed parallel crosses between two pairs of divergent inbred *Nicotiana rustica* lines ($V_1 \times V_5$ and $V_2 \times V_{12}$) and found their results adequately described by quite similar patterns of additive, dominance, and digenic interaction effects. For five traits measured, two (2- and 4-wk height) showed the same significant parameters. All were the same sign and the magnitude varied at most by a factor of three. For the remaining three traits (6-wk, flowering, and final height) models varied only by one parameter and the magnitudes of the parameters were again similar. The Nicotiana lines were collected from different countries before the Second World War and are quite unlikely to be related by pedigree (H. S. Pooni, personal communication). As each inbred line represents a single sampling from the species, and with only four independent samples, there is considerable scope for sampling error, yet the models describing the various characters were quite consistent. This suggests that composite genetic parameters for polygenic characters may vary little (although this would not necessarily be the case for individual loci). Second, Hard *et al.* (1992) performed a generation means analysis of critical photoperiod in the pitcher-plant mosquito, *W. smithii*. Crosses were performed between a common southern source population and each of two northerly, derived populations. The parameters (including digenic epistatic parameters) of the models describing the parallel divergence between southern (source) and northern

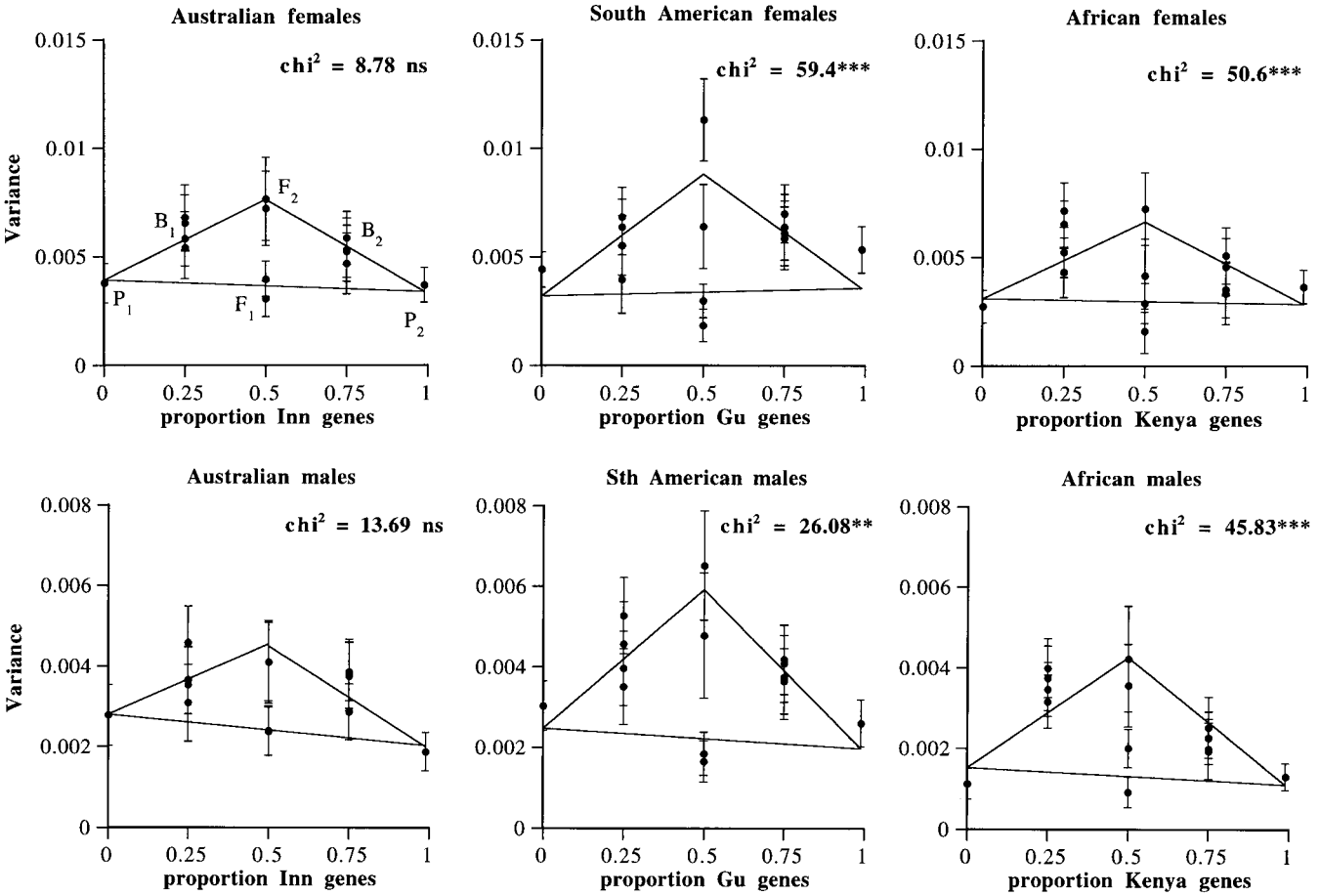

Figure 3.—Variance of wing area (in square millimeters) in each cross plotted as a function of the proportion of genes derived from the smaller (northern) parent. The triangle connects points (not shown) representing the expected variances under a purely additive model. The vertices correspond to the expected variances for the P_1, P_2, and F_2 generations and the midpoints of each side of the triangle correspond to the expected variances for the F_1 and backcross generations (the positions are indicated in the Australian females graph). χ^2 values indicate the significance of the fit of the observed variances to the simple additive model. Error bars are ±2 SE. Errors were estimated as described in the text.

populations were again remarkably similar in occurrence and magnitude. Because the northerly populations are thought to derive from the southern populations, the northern populations are, to some extent, samples of the source population. Their results suggest that sampling differences are a minor source of variation between models.

By contrast, the large differences observed in this study between the models for each cline suggest that the clines are based on radically different genetic archi-

TABLE 4

Significance tests for the fit between the observed generation variances and the maximum-likelihood additive model (11 d.f.)

		Wing area		Wing length	
Sex	Cline	χ^2	*P*	χ^2	*P*
Female	Australian	8.78	0.64	7.60	0.75
	South American	59.38	<0.001	52.28	<0.001
	African	50.57	<0.001	34.52	<0.001
Male	Australian	13.69	0.25	8.57	0.66
	South American	26.08	0.006	63.17	<0.001
	African	45.83	<0.001	41.18	<0.001

tectures, possibly different genes. As the sample sizes were similar, this cannot be due simply to variation in statistical power in the analysis of each cline. Supporting evidence comes from studies of the cellular basis of wing size divergence in the Australian and South American clines. Wing area in *D. melanogaster* is a product of cell area and cell size (Robertson 1959; Cavicchi *et al.* 1985; James *et al.* 1995). The body size cline observed in eastern Australia has been shown to be largely, though not exclusively, due to variation in cell number (James *et al.* 1995), while the South American cline is due to changes in both cell number and cell size (B. Zwaan, R. Azevedo, A. James and L. Partridge, unpublished data). Therefore, a given wing size may be produced by diverse genetic mechanisms. Another indicator that there are real genetic differences between the clines is that they also differ for other nonclinal characters. For example, there was no significant cline in aspect ratio (wing length^2/wing area) in our experiment (a similar result was found for different samples from the same cline by Azevedo *et al.* 1998), yet there are significant differences in aspect ratio between the clines. The ranking of aspect ratios of each cline (Australia > South America > Africa) is maintained in all the generations raised in this experiment.

How much of this diversity can be ascribed to different segregating loci can be fully answered only by a quantitative trait loci (QTL) analysis, but our results indicate that different clines may contain quite different segregating alleles and loci. It is possible that differences between parallel clines are as dependent on stochastic processes (*e.g.*, founder effects) as they are on deterministic processes, echoing the conclusions of Armbruster *et al.* (1997). Simulations have shown that the effects of mutation order can be important in divergence between populations (Clarke *et al.* 1988).

Recently, there has been considerable interest in the potential role of epistasis in population divergence (*e.g.*, Whitlock *et al.* 1995; Fenster *et al.* 1997), particularly with the advent of QTL analysis and the application of techniques such as line-cross analysis to animal populations (*e.g.*, Cohan *et al.* 1989; Hard *et al.* 1992; Starmer *et al.* 1998). As highlighted by Fenster *et al.* (1997), methods for investigating epistasis fall into two broad categories—those that measure epistatic contributions to current phenotypic means (*e.g.*, include F_2 breakdown, line-cross analysis, multilocus associations, and QTL mapping) and those that measure epistatic variances directly. Our results provide an illustration of the distinction between these two classes of results, showing that epistatic contributions to current generation means do not necessarily reflect epistatic variances for the same character. For instance, the generation means analysis of the Australian males indicated the presence of significant epistatic effects, yet a simple additive model adequately described the generation variances. The converse is true for the South American cline. Different explanations are possible for these discrepancies. First, when epistatically interacting loci are close to fixation, as expected near local fitness peaks, the ratio of epistatic to additive variance approaches a minimum. Consequently, little or no epistatic variance is observed despite potentially large amounts of underlying epistatic interaction (Whitlock *et al.* 1995). This may explain the situation observed in the Australian cline. In the South American cline, no epistasis is apparent in the generation means, yet the variances show a highly significant divergence from the expectations of an additive model. This suggests a possible role for dominance and maternal variances. Data from additional generations is required to test these possibilities (Mather and Jinks 1982).

A number of previous investigations of wing size in various Drosophila species have found evidence of maternal effects based on significant differences between reciprocal F_1 and F_2 generations (McFarquhar and Robertson 1963; Anderson 1968; Cavicchi *et al.* 1985). Our results strongly support the idea that maternal effects can influence wing size. Perhaps the most striking example was that of Cavicchi *et al.* (1989), who found cytoplasmic effects on female wing size in artificially selected lines that persisted over five generations (males were not measured). In contrast, we found evidence of cytoplasmic effects only in the Australian males. This effect was very small, of similar magnitude to the Y effects measured in the males from the other clines. It appears that cytoplasmic effects on wing size may be highly variable and population-specific.

An important implication of our results for other studies is that it could often be wrong to assume that outbreeding depression reflects the disruption of coadapted gene complexes, unless alternative reasons for the depression have been eliminated. Outbreeding depression or F_2 breakdown (as distinct from F_1 breakdown) may result when populations of the same species originating from different localities are crossed. Depression of trait values or performance in the F_2 generation is usually taken as evidence of the breakup of coadapted gene complexes by recombination (*e.g.*, Vetukhiv 1956; McFarquhar and Robertson 1963; Anderson 1968; Burton 1990; Blows 1993; Armbruster *et al.* 1997). Under a simple additive-dominance model, the deviation of the F_2 generation from the midparent is expected to be half that of the deviation of the F_1 generation from the midparent. This deviation is usually tested using the *C*-scaling test of Mather and Jinks (1982), in which the quantity $4F_2 - 2F_1 - P_1 - P_2$ is expected to be zero. Significant deviations from this expectation, *i.e.*, outbreeding depression, are often assumed to be evidence of coadapted gene complexes in the parental strains. However, the deviation of *C* from zero indicates not only the presence of epistatic effects but also maternal effects. The epistatic deviation is $-2[aa] - [dd]$, while the maternal effects deviation is $4[d]_m - 2[a]_m +$

$2c$ for the cross $P_1 \times P_2$ and $4[d]_m + 2[a]_m - 2c$ for the reciprocal cross (Kearsey and Pooni 1996). Therefore, the reciprocal crosses have the expectations

$$-2[aa] - [dd] + 4[d]_m - 2[a]_m + 2c$$

and

$$-2[aa] - [dd] + 4[d]_m + 2[a]_m - 2c.$$

Clearly, significant maternal effects, and dominance maternal effects in particular, may produce significant C values when, in fact, no significant composite epistasis is present. For wing area, we have shown that epistatic effects on size were present in three clines (Australian male and both African clines) and absent in the remaining clines. Therefore, using our data, we were able to test the common assumption that failure of the C-test indicates epistatic interactions. For the clines where epistatic effects were present, five out of six C-tests were significant as expected (after Bonferroni correction; the six tests consisted of 3 clines $\times$ 2 reciprocals). However, in the remaining clines, where no significant epistatic effects were measured, five out of six C-tests were again significant (Australian females, $C = -0.165$, $P < 0.001$; $C_{\text{reciprocal}} = -0.125$, $P < 0.001$; South American females, $C = -0.232$, $P < 0.001$; $C_{\text{reciprocal}} = -0.087$, $P = 0.015$; South American males, $C = -0.001$, not significant; $C_{\text{reciprocal}} = -0.105$, $P < 0.001$). The significant C-test for the Australian female WA cline may be explicable by the $[ad]$ parameter, shown in Table 2, which was shown to significantly improve the fit of that model. However, no similar suggestion of near-significant epistatic effects could be found in the South American data. The only remaining explanation is that the South American cline failed the C-test due to the presence of maternal effects. Therefore, assuming that significant maternal effects were absent in the South American clines would have led to a mistaken conclusion that the outbreeding depression indicated epistasis in the parental lines. The common assumption that maternal effects are absent can be highly misleading.

Our survey of the genetic basis of wing size variation has shown that apparently similar phenotypic divergence can result from quite different underlying genetic architectures. Epistasis, maternal, and cytoplasmic effects can each make significant but highly variable contributions to the ultimate phenotype. Given the close association between body size and fitness in *D. melanogaster*, it will be interesting to see if the multiple peaks of the phenotypic landscape are reflected in a similar fitness landscape.

We thank Giselle Geddes and Lillian Tsang for invaluable technical assistance, Ricardo Azevedo for bringing Object-Image to our attention, and Mauro Santos and John Sved for helpful comments. We thank the anonymous reviewers whose comments greatly improved this article, especially with respect to the statistical analysis. This work was supported by a grant from the Natural Environment Research Council (United Kingdom).

LITERATURE CITED

Anderson, W. W., 1968 Further evidence for coadaptation in crosses among geographic populations of *Drosophila pseudoobscura*. Genet. Res. **12:** 317–330.

Armbruster, P., W. E. Bradshaw and C. M. Holzapfel, 1997 Evolution of the genetic architecture underlying fitness in the pitcher-plant mosquito, *Wyeomyia smithii*. Evolution **51:** 451–458.

Azevedo, R. B. R., A. C. James, J. McCabe and L. Partridge, 1998 Latitudinal variation of wing:thorax size ratio and wing aspect ratio in *Drosophila melanogaster*. Evolution **52:** 1353–1362.

Blows, M. W., 1993 The genetics of central and marginal populations of *Drosophila serrata*. 2. Hybrid breakdown in fitness components as a correlated response to selection for desiccation resistance. Evolution **47:** 1271–1285.

Burton, R. S., 1990 Hybrid breakdown in developmental time in the copepod *Tigriopus californicus*. Evolution **44:** 1814–1822.

Cavicchi, S., D. Guerra, G. Giorgi and C. Pezzoli, 1985 Temperature-related divergence in experimental populations of *Drosophila melanogaster*. 1. Genetic and developmental basis of wing size and shape variation. Genetics **109:** 665–689.

Cavicchi, S., D. Guerra, V. Natal, C. Pezzoli and G. Giorgi, 1989 Temperature-related divergence in experimental populations of *Drosophila melanogaster*. II. Correlation between fitness and body dimensions. J. Evol. Biol. **2:** 235–251.

Chai, C. K., 1956 Analysis of quantitative inheritance of body size in mice. II. Gene action and segregation. Genetics **41:** 165–178.

Clarke, B. C., P. R. Shelton and G. S. Mani, 1988 Frequency-dependent selection, metrical characters and molecular evolution. Philos. Trans. R. Soc. Lond. B Biol. Sci. **319:** 631–640.

Cohan, F. M., A. A. Hoffmann and T. W. Gayley, 1989 A test of the role of epistasis in divergence under uniform selection. Evolution **43:** 766–774.

Endler, J. A., 1977 *Geographic Variation, Speciation and Clines*. Princeton University Press, Princeton, NJ.

Falconer, D. S., 1989 *Introduction to Quantitative Genetics*. Longman, London.

Fenster, C. B., L. F. Galloway and L. Chao, 1997 Epistasis and its consequences for the evolution of natural populations. Trends Ecol. Evol. **12:** 282–286.

Guerra, D., M. C. Pezzoli, G. Giorgi, F. Garoia and S. Cavicchi, 1997 Developmental constraints in the Drosophila wing. Heredity **79:** 564–571.

Hard, J. J., W. E. Bradshaw and C. M. Holzapfel, 1992 Epistasis and genetic divergence of photoperiodism between populations of the pitcher-plant mosquito, *Wyeomyia smithii*. Genetics **131:** 389–396.

Hayman, B. I., 1958 The separation of epistatic from additive and dominance variation in generation means. Heredity **12:** 371–390.

Hayman, B. I., 1960 Maximum likelihood estimation of genetic components of variation. Biometrics **16:** 369–390.

James, A. C., R. B. R. Azevedo and L. Partridge, 1995 Cellular basis and developmental timing in a size cline of *Drosophila melanogaster*. Genetics **140:** 659–666.

James, A. C., R. B. R. Azevedo and L. Partridge, 1997 Genetic and environmental responses to temperature of field adapted populations of *Drosophila melanogaster*. Genetics **146:** 881–890.

Kearsey, M. J., and H. S. Pooni, 1996 *The Genetical Analysis of Quantitative Traits*. Chapman & Hall, London.

Keightley, P. D., 1989 Models of quantitative variation of flux in metabolic pathways. Genetics **121:** 869–876.

Lynch, M., 1991 The genetic interpretation of inbreeding depression and outbreeding depression. Evolution **45:** 622–629.

Lynch, M., and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.

Mather, K., and J. L. Jinks, 1982 *Biometrical Genetics: The Study of Continuous Variation*. Cambridge University Press, Cambridge.

McFarquhar, A. M., and F. W. Robertson, 1963 The lack of evidence for co-adaptation in crosses between geographical races of *Drosophila subobscura* Coll. Genet. Res. **4:** 104–131.

Mohamed, A. H., 1959 Inheritance of quantitative characters in Zea mays. I. The estimation of the number of genes controlling the time of maturity. Genetics **44:** 713–724.

Oakeshott, J. G., J. B. Gibson, P. R. Anderson, W. R. Knibb, D. G. Anderson *et al.*, 1982 Alcohol dehydrogenase and glycerol-3-

phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. Evolution **36:** 86–96.

Pooni, H. S., J. L. Jinks and J. F. F. Toledo, 1985 Predicting and observing the properties of second cycle hybrids using basic generations and inbred line × *F1* crosses. Heredity **54:** 121–129.

Powell, J. R., 1997 *Progress and Prospects in Evolutionary Biology: The* Drosophila *Model.* Oxford University Press, New York.

Reeve, E. C. R., and F. W. Robertson, 1952 Studies in quantitative inheritance. II. Analysis of a strain of *Drosophila melanogaster* selected for long wings. J. Genet. **51:** 276–316.

Roberts, D. B., 1986 *Drosophila. A Practical Approach.* IRL Press, Oxford.

Robertson, F. W., 1959 Studies in quantitative inheritance. XII. Cell size and number in relation to genetic and environmental variation of body size in Drosophila. Genetics **44:** 869–896.

Stalker, H. D., 1976 Chromosome studies in wild populations of *D. melanogaster*. Genetics **82:** 323–347.

Starmer, W. T., M. Polak, L. L. Wolf and J. S. F. Barker, 1998 Reproductive characteristics of the flower breeding *Drosophila hibisci* Bock (Drosophilidae) in eastern Australia: genetic and environmental determinants of ovariole number. Evolution **52:** 806–815.

Thomas-Orillard, M., 1984 Modifications of mean ovariole number, fresh weight of adult females and developmental time in *Drosophila melanogaster* induced by Drosophila C virus. Genetics **107:** 635–644.

Van't Land, J., P. van Putten, A. Kamping and W. van Delden, 1995 Latitudinal variation in wing length and allele frequencies for *Adh* and α-*Gpdh* in populations of *Drosophila melanogaster* from Ecuador and Chile. Dros. Inf. Serv. **76:** 156.

Vetukhiv, M., 1956 Fecundity of hybrids between geographic populations of *Drosophila melanogaster*. Evolution **10:** 139–146.

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.

Wright, S., 1977 *Evolution and the Genetics of Populations. Vol. 3. Experimental Results and Evolutionary Deductions.* University of Chicago Press, Chicago.

Communicating editor: T. F. C. Mackay