Molecules versus morphology: the detection of selection acting on morphological characters along a cline in *Drosophila melanogaster*

ANTHONY DOUGLAS LONG*†‡ & RAMA S. SINGH†

†*Department of Biology, McMaster University, Hamilton, Ont., Canada L8S 4K1 and ‡Center for Population Biology,
University of California at Davis, Davis, CA 95616, U.S.A.*

This work examines the nature of north–south clinal variation in morphological characters in *Drosophila melanogaster*. Isofemale lines were established from flies collected along a transect extending from Winnipeg, Manitoba (Canada) to Tampa Bay, Florida (U.S.A.). Offspring from different lines within each location were then cultured under standardized conditions and used to examine phenotypic variation in seven morphological characters along the cline. In addition, allozyme variation at seven polymorphic loci was examined for the same set of clinal populations. Scutellum length and wing length show the strongest clinal trends. Clinal variation is nonmonotonic, with larger flies in the middle latitudes and smaller flies in the north and south. This result contrasts with other studies which have shown monotonic clines. Patterns of population subdivision were different for the different characters. This implies that there are different selective forces acting on the different morphological characters. Based on a comparison of morphological and molecular population subdivision for adjacent populations it is inferred that natural selection is operating to maintain a high level of population subdivision for wing width and the first principal component between one of the sets of populations. A combined approach using molecules and morphology may provide an alternative to retrospective selection analysis for detecting selection in nature.

Keywords: allozymes, cline, *Drosophila*, gene flow, geographical variation, wing length.

Introduction

Attempts to measure selection in the wild on quantitative traits have relied primarily on two methods. The first uses the retrospective selection equation popularized by Lande (1979) to estimate a vector of selection coefficients on a set of quantitative traits if the genetic variance/covariance matrix for the traits and a vector corresponding to the net change in each character between two populations are known. Turelli (1988) showed that this method has several weaknesses. It assumes, for example, that the genetic variance/covariance matrix is constant over both time and space, and is estimated without error. In addition, if the environments in which the characters from the two populations are measured are not identical the differences between populations could be entirely environmentally determined (Falconer, 1981; Gupta &

Lewontin, 1982; James, 1983). The second method estimates selection directly on a set of characters by regressing some measure of 'fitness' on the phenotype of each individual and estimating both the linear and quadratic partial regression coefficients, which correspond to the directional and stabilizing selection operating on each character (Lande & Arnold, 1983). The major disadvantages of this method with respect to the study of spatial variation are that the measure of selection is only as good as the measure of fitness, which is difficult to measure, and that selection is measured over a single generation so that it may not be applicable to measuring selection acting over space or longer time periods than a few generations (Price *et al.*, 1984).

Another technique that has been used in the past to infer qualitatively the patterns of natural selection acting on morphological characters over space is the study of geographical or clinal variation (for review see Endler, 1977; Reyment *et al.*, 1984). Presumably, a trend seen over a transect, or multiple independent

transects, is the result of some underlying environmental selective agent. It follows that studies of clinal variation in characters of likely adaptive importance, such as morphological, physiological, behavioural and life history traits, can provide insight as to how organisms adapt to their environment. A large number of studies have examined geographical variation in morphological characters in laboratory-reared dipterans. This is possible because of the wide distribution, ease of collection, and relative ease of laboratory rearing of many dipteran species. It appears that a great deal of local differentiation is possible even for species with a relatively high dispersal rate such as *Drosophila melanogaster* (Coyne & Milstead, 1987; Coyne *et al.*, 1987). It also appears that north–south (N–S) clinal variation is a general phenomenon in many of the dipterans studied.

Although populations of *Drosophila* have been studied extensively with the objective of estimating levels of gene flow between populations (for reviews see David & Capy, 1988; Lachaise *et al.*, 1988; Singh, 1989; Singh & Long, 1992) only a few studies have combined estimates of gene flow and morphological character differences to understand the nature of selection acting on morphological traits (Prout & Barker, 1993; Spitze, 1993). The test of Prout & Barker (P–B test) is in many ways similar to the earlier Lewontin–Krakauer test (1973), but it avoids the phylogenetic pitfall pointed out by Robertson (1975) by averaging over loci in estimating character and allozyme population subdivision. A problem still exists with the P–B test though. Hidden biases, such as the choice of allozyme loci used for analysis, may result in a difference between morphological and allozyme population subdivision that is not because of the action of natural selection. Here we propose a new test which is similar to the P–B test but is extended to multiple

pairs of populations compared and is insensitive to potential bias introduced by an inappropriate selection of molecular markers. If morphological and allozyme population subdivision are not *proportional* over a number of pairs of populations compared, the past action of natural selection can be inferred. Here we apply this test to estimates of morphological and allozyme population subdivision made from adjacent populations collected from a N–S latitudinal cline extending from Winnipeg, Manitoba (Canada) to Tampa Bay, Florida (U.S.A.). Because flies reared under standard laboratory conditions from a carefully sampled transect are used, the observed differences between populations are genetic in origin and extraneous sources of variation such as time of year of collection, altitude, and longitude are reduced. Thus we believe that the current paper presents a robust approach for inferring the patterns of selection acting between populations in nature.

Materials and methods

Experimental design and character measurements

In the late summer of 1988, adult *Drosophila melanogaster* females were collected in traps baited with fermenting bananas and were used to establish isofemale lines. The flies were collected at approximately 500 kilometre intervals along a N–S transect that skirted west of the Appalachian mountains. At each location traps were, whenever possible, placed at least a kilometre apart to avoid sampling only the descendants of a few females that may have recently colonized a local food source (Hoffmann & Neilsen, 1985). Details of the collection sites are given in Table 1. Within two to four generations of being established in the laboratory, 10 virgin females were mass mated to

Table 1 Details of samples used in the study

Location	N_I^*	Lat. (°N)	Long. (°W)	Date d/m	$N_L^{\dagger}$	Collection site
1 Winnipeg, Man.	83	50.0	97.25	16/7	4	Brewery
2 Windsor, Ont.	50	42.0	83.0	31/8	3	Fruit stand, campsite, apple tree
3 Dayton, OH.	66	40.25	84.25	2/9	9	Fruit stand, pear tree, campsite, apple tree
4 Louisville, KY.	215	37.75	85.75	4/9	9	Campsite, fruit stand, apple tree, bushes
5 Nashville, TN.	50	36.0	86.75	6/9	5	Campsite (hornets in trap), bushes
6 Cartersville, GA.	145	34.0	84.75	8/9	6	Apple tree, second apple tree, third apple tree, pear tree, grape vine
7 Cordele, GA.	155	32.0	83.75	10/9	8	Shade tree, peach tree, bushes, apple tree
8 High Springs, FL.	71	29.75	82.5	12/9	6	Vines, blueberry bush
9 Tampa Bay, FL.	3	28.0	82.25	14/9	1	Orange grove

*Number of male flies measured.
†Number of lines for which male flies were measured.

10 males from a given line and allowed to oviposit for 12 h in jars containing approximately 40 mL of standard banana medium. Parents were removed from the jars and eggs were removed from each jar so that only approximately 80 eggs remained. The jars were incubated at 24°C with a 12 h day/night cycle for 12 days, at which time the offspring were collected. This design controlled for the known effects of maternal age, egg density, temperature, and environmental quality on adult body size in *Drosophila* (Ashburner, 1989). Collected flies were anaesthetized with diethyl ether and dry-preserved in vials containing vapour of ethyl acetate for later measurement of external morphological characters. An average of 16 flies was measured per line.

The seven characters considered here are wing length (WL), wing width (WW), femur length (FEM), tibia length (TIB), scutellum width (SCUT), face width (FW), and face length (FL) (Fig. 1). Two additional principal component scores were created and analysed (Chatfield & Collins, 1980). The first principal component can be considered to be a measure of general body size as all component loadings were large and positive on this component (Long, 1991). All dissec-

tions and measurements were made using a set of 'mouse' controlled crosshairs overlaid on a video image of the character (VOCGA card, K-systems of Texas) supplied by a colour video camera fitted to a Zeiss DRC stereomicroscope. The raw data consisted of a set of pixel distances in the X and Y directions (distances in the Y direction are multiplied by a constant) which were used to calculate character sizes in 'pixel units'.

Data analysis

A log tranformation of the raw data was deemed necessary as the variances of the traits under consideration were a function of their means and the data appeared to be distributed in a log normal fashion. A suitable log transformation is $X' = \ln(X + a/b)$, where X is a value of a dependent variable in raw units, X' is the value of the same variable in transformed units, and a and b (47.31 and 0.024, respectively) are the intercept and regression coefficient of a regression of character standard deviations on character means over characters (Falconer, 1981). The natural logarithm will transform the variables to an additive scale, the constants will decouple the relationship between the mean and variance of the characters. All further analyses were carried out on data transformed in this manner with flies subsequently removed that had a character measure more than four standard deviations from the mean for that character. At this point the data appeared to be normally distributed based on a visual examination of normal probability plots for each character combination.

Data were analysed for each character using the GLM (general linear model) procedure of SAS with the model $Y_{ijk} = \mu + Loc_i + Line_{j(i)} + \varepsilon_{k(ij)}$, where Y_{ijk} is the score for the kth individual in the jth line in the ith location, parentheses represent a nesting and all effects are treated as random. In addition, the regression model $Y_{ij.} = \mu + Lat_i + Lat_i^2 + \varepsilon_{ij}$, where $Y_{ij.}$ is the mean score for the jth line at the ith latitude, was fitted to the data. Using line means avoids lines with a large number of individuals having a large effect on the regression. Sums of squares were used to estimate the proportion of the total variation due to location attributable to the linear and quadratic trends over latitude, and the significances of the linear and quadratic terms were assessed by F-tests, using the residual (lack of fit) term from the regression as a denominator.
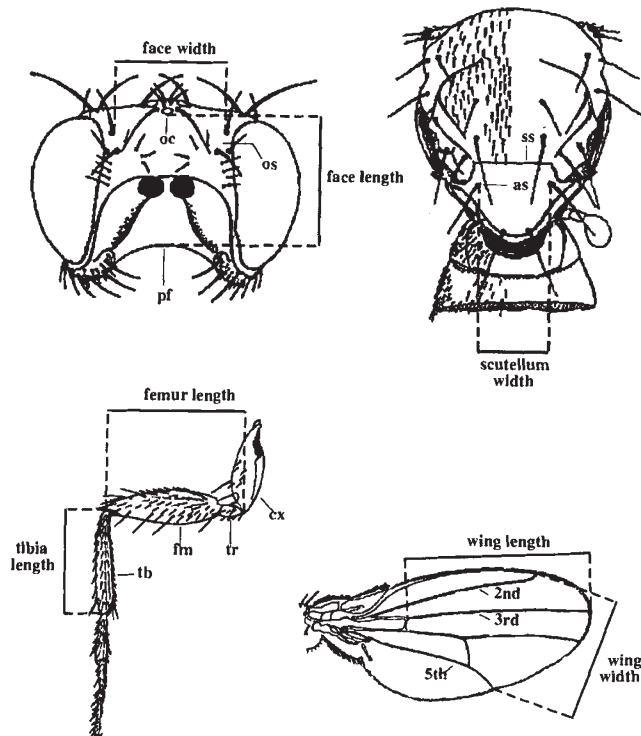

Fig. 1 The seven morphological characters measured on *Drosophila melanogaster*. Abbreviations: 2nd, 3rd, and 5th refer to the longitudinal vein of the same number; cx = coxa; tr = trochanter; fm = femur; tb = tibia; ss = scutoscutellaer suture; as = anterior scuteller setae; oc = ocellus; os = orbital setae (3); pf = edge of prefrons.

Genetic parameter estimation

The experimental design of this study allows the estimation of the genetic variance between locations and

the average genetic variance within locations (Falconer, 1981; Table 15.1 with modifications). This is accomplished in the above design by equating:

$$V_{loc} = V_{Gb},$$

$$V_{line} = V_b + 2fV_{Gw} \text{ and}$$

$$V_{ind} = V_e + (1 + F - 2f) V_{Gw},$$

where V_{Gb} is the additive genetic variance between locations, V_{Gw} is the additive genetic variance between lines within locations, V_b is the variance due to common bottles, V_e is the environmental variance, V_{loc} is the observed variance due to locations, V_{line} is the observed variance due to lines within locations, V_{ind} is the observed variance due to individuals within lines within locations, F is the inbreeding coefficient in isofemale lines from the first generation after establishment to the time individuals are measured, and f (equal to 0.25 for isofemale lines) is the coancestry of individuals in an isofemale line (i.e. identity by descent of the first generation). The inbreeding coefficient in the isofemale lines of this study can be estimated as $0.20 < F < 0.40$ by assuming two to four generations of inbreeding from line establishment to fly measurement at an effective population size of about five individuals.

To equate the variance components to their expectations in this manner it is necessary to assume that the variance between lines and between locations is additive (Falconer, 1981; Prout & Barker, 1993). Body size in *Drosophila* appears to have only a very small nonadditive genetic component (Robertson, 1957), thus this assumption appears to be met. This study will not allow the simultaneous estimation of V_{Gw}, V_b and V_e, but previous studies on wing length in *Drosophila* have shown V_b to be small and approximately constant relative to V_e (i.e. $V_b = kV_e$, where k can be estimated from similar studies of wing length as 0.15 (95 per cent confidence interval is 0.11 to 0.19) (Robertson, 1957; Prout, 1958)). Given the above assumptions: $\hat{V}_{Gb}$ is estimated as V_{loc}, and $\hat{V}_{Gw} = [2(V_{line} - k \ V_{ind})]/[1 - k - 2kF]$.

Bootstrapping was used in order to obtain sampling variance on point estimates of variance components, and for subsequent analysis using functions of these variables. For each character and set of adjacent populations, 100 bootstrap replicates of the raw data were performed and five estimates of $\hat{V}_{Gb}$, $\hat{V}_{Gw}$ and $\hat{V}_e$ were made each using a different set of constants for k and F ($\chi = \{k, F\} = \{0.15, 0.30\}, \{0.10, 0.20\}, \{0.10, 0.40\}, \{0.20, 0.20\}, \{0.20, 0.40\}$). Bootstrapping was carried out with simultaneous resampling at both the level of the line and of individuals within a line, and the REML option of the VARCOMP procedure of SAS was used to estimate variance components. The level of popula-

tion subdivision for both allozymes and morphological traits was estimated using Wright's F_{st} (Crow & Kimura, 1970). For each estimate of $\hat{V}_{Gb}$, $\hat{V}_{Gw}$ and $\hat{V}_e$ an estimate of F_{st} was made as $\hat{V}_{Gb}/(\hat{V}_{Gb} + 2 \times \hat{V}_{Gw})$, which was subsequently arcsine transformed ($F'_{st} = \text{arcsine } \sqrt{F_{st}}$) to make estimates of F_{st} approximately normal (Sokal & Rolf, 1981; Prout & Barker, 1993). Analysis of variance was carried out on the data created in this manner in order to determine the effect of the choice of constants (χ) on the estimate of F'_{st}. As χ appeared to have little effect on the variation in estimates of F'_{st} relative to other sources of variation, the effect of χ was statistically removed from the model, and subsequent analyses carried out on residuals generated in this manner.

In order to test the significance of the deviation of character F'_{st} from allozyme F'_{st} the model $F'_{st \ ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$ was fitted for each morphological trait, where α_i is the effect of molecules (allozyme F'_{st}) vs. morphology (morphological F'_{st}), β_j is the effect of the set of populations compared, and ε_{ijk} is the error within a given treatment based on bootstrap replicates. $\alpha\beta_{ij}$ measures the 'lack of fit' from a model which assumes that the two measures of F'_{st} are proportional. We might expect the main effects of α_i and β_j to be significant, but if measures of character and morphological F'_{st} are proportional to one another over the sets of populations compared (i.e. there is no selection acting on the morphological character) the $\alpha\beta_{ij}$ interaction terms should not be significant.

Allozyme electrophoresis and analysis

A rapid method of allozyme analysis using a cellulose acetate electrophoresis system was employed (Hebert & Beaton, 1989). Details of the staining methods used are given in Singh & Rhomberg (1987b). F_{st} was estimated for each locus and set of adjacent populations as previously described, and arcsine transformed as described above (Singh & Rhomberg, 1987b). Errors were placed on estimates of arcsine transformed F_{st} by resampling the loci 500 times with replacement.

Results

Nonmonotonic clinal variation

The means and standard deviations of the male flies measured from each location are given in Table 2. The coefficients of variation for each line based on the raw data (not shown) range from roughly two to five per cent which is consistent with previous studies (Coyne & Beecham, 1987), although they are much lower in the transformed data set. Data are not presented for

Table 2 Means and standard deviations of morphological traits for male *Drosophila melanogaster* in log transformed units*

Location	PC1	PC2	WL	WW	FEM	TIB	SCUT	FW	FL
1	−0.165 (0.025)	−0.030 (0.017)	9.636 (0.012)	9.323 (0.020)	9.033 (0.011)	8.823 (0.010)	8.514 (0.012)	8.544 (0.018)	8.746 (0.015)
2	−0.104 (0.021)	−0.010 (0.019)	9.649 (0.016)	9.350 (0.011)	9.068 (0.012)	8.859 (0.003)	8.544 (0.017)	8.555 (0.025)	8.753 (0.013)
3	−0.077 (0.061)	0.007 (0.026)	9.655 (0.020)	9.357 (0.031)	9.092 (0.019)	8.871 (0.018)	8.573 (0.021)	8.557 (0.031)	8.571 (0.041)
4	−0.068 (0.020)	0.016 (0.015)	9.667 (0.014)	9.358 (0.013)	9.082 (0.010)	8.859 (0.011)	8.587 (0.009)	8.574 (0.009)	8.743 (0.015)
5	−0.017 (0.022)	0.009 (0.010)	9.681 (0.010)	9.390 (0.014)	9.106 (0.009)	8.887 (0.014)	8.591 (0.010)	8.577 (0.006)	8.771 (0.012)
6	−0.088 (0.033)	0.015 (0.010)	9.664 (0.017)	9.337 (0.017)	9.081 (0.012)	8.861 (0.011)	8.574 (0.015)	8.576 (0.006)	8.737 (0.016)
7	−0.115 (0.023)	−0.013 (0.017)	9.649 (0.007)	9.340 (0.014)	9.066 (0.008)	8.852 (0.007)	8.563 (0.011)	8.534 (0.025)	8.753 (0.024)
8	−0.101 (0.039)	−0.009 (0.018)	9.634 (0.022)	9.331 (0.023)	9.077 (0.016)	8.855 (0.014)	8.571 (0.012)	8.561 (0.009)	8.757 (0.021)

*Means are the unweighted average of the mean character measures (see Fig. 1) for each line; standard deviations are calculated over lines within locations; the log transformation is described in the text.

Tampa Bay as the orange groves from which these lines were collected had recently been sprayed and as a result too few lines were collected. Female flies showed the same qualitative patterns of variation as the males (not shown) but they were not considered in further analysis because they were not measured for all locations. Table 3 gives the results of an ANOVA on male flies testing the effects of Location and Line within Location for seven morphological characters and the two principal components. For all characters, the effect of Line within Location is significant, and for all characters except face length, the effect of Location is significant. Thus the average differences between locations were significantly different from one another when tested with mean square error between lines within locations.

Figure 2 shows the phenotypic trends in three characters and the first principal component over distance. There appears to be a nonmonotonic trend which is most apparent for wing length and scutellum width, which are highly correlated with general body size, as defined by loadings on the first principal component (data not shown). Although there is a great deal of variation due to Lines within Locations, there is a clear nonmonotonic trend in the mean of many of the characters studied over the cline. That the morphological cline peaked in the middle latitudes, as opposed to the north, stands in marked contrast to other studies of clinal variation. Table 3 shows that for the majority of characters a model including a quadratic term explains significantly more variation in line means than a linear model. The quadratic coefficient for each character is negative in value, implying a peak in the middle latitudes. In most cases the linear term of the quadratic model was not significant. The proportion of variance in location attributable to the quadratic term of a regression on latitude varied from slightly more than the linear term (e.g. scutellum) to two orders of magnitude larger than the linear term (e.g. wing length and wing width). In contrast to previous studies of clinal variation, the quadratic term of the regression was significant and explained a biologically significant proportion of the total variation due to Location.

Population structure for molecules and morphology

An analysis of variance on and mean values of the estimates of arcsine transformed F_{st} for adjacent sets of populations derived from the bootstrapping of the original data are presented in Tables 4 and 5, respectively. In this analysis Tampa Bay was dropped because of the small number of flies representing this population. The analysis is of a model that includes the effects of: adjacent populations compared, the choice of

Table 3 ANOVA table for male *Drosophila melanogaster* showing the effects of Location, Line within Location, and Latitude on variation in male morphological traits and principal components

Source	d.f.	PC1	PC2	WL	WW	FEM	TIB	SCUT	FW	FL
Sum of squares: complete model†										
Location	8	101.52***	20.24***	15.28***	22.64***	25.20***	17.36***	39.28***	21.44***	6.40
Line (Location)	42	54.18***	16.80***	13.86***	18.48***	8.40***	7.56***	12.18***	19.74***	23.94***
Error	786	152.48	94.32	44.80	41.66	36.94	37.73	41.66	40.09	117.11
Sum of squares: Line means model										
Location	7	61.03	1.04	0.99	1.60	1.54	1.09	1.99	1.12	0.39
Latitude	1	0.46	0.05	0.00	0.00	0.20	0.11	0.66*	0.01	0.02
(Latitude)²	1	3.45*	0.70*	0.60*	0.88	0.87*	0.62*	0.93*	0.28	0.00
Residual	5	2.20*	0.29	0.39*	0.72*	0.46***	0.36***	0.40**	0.83**	0.37
Pure Error	42	5.44	1.33	1.04	1.65	0.73	0.63	0.84	1.56	2.46
Parameter estimates‡										
Intercept		−1.06	−0.45	9.23	8.82	8.60	8.45	8.11	8.27	8.78
Latitude		5.23	2.18	2.25	2.74	2.60	2.20	2.60	1.53	−0.14
(Latitude)²		−6.91	−3.12	−2.88	−3.49	−3.48	−2.92	−3.60	−1.98	0.14
Percentage of variation in Location due to:										
Latitude		0.75	4.57	0.09	0.05	13.09	10.23	32.96	0.46	3.87
(Latitude)²		5.65	67.51	60.53	55.15	56.80	56.69	46.90	25.27	0.35

†Sum of squares times 100 are listed under their respective characters in the upper part of the table. The Line means model was fitted with Tampa Bay flies removed. Location (complete model) is tested with $F_{8,42}$, Line (Location) with $F_{42,786}$, Latitude and (Latitude)² with $F_{1,5}$, and the Residual with $F_{5,42}$. $*P \leqslant 0.05$, $**P \leqslant 0.01$, $***P \leqslant 0.001$. The significance of the Latitude squared term after fitting the Latitude term indicates that a quadratic model explains significantly more variation in the data set than a model that only includes Latitude.

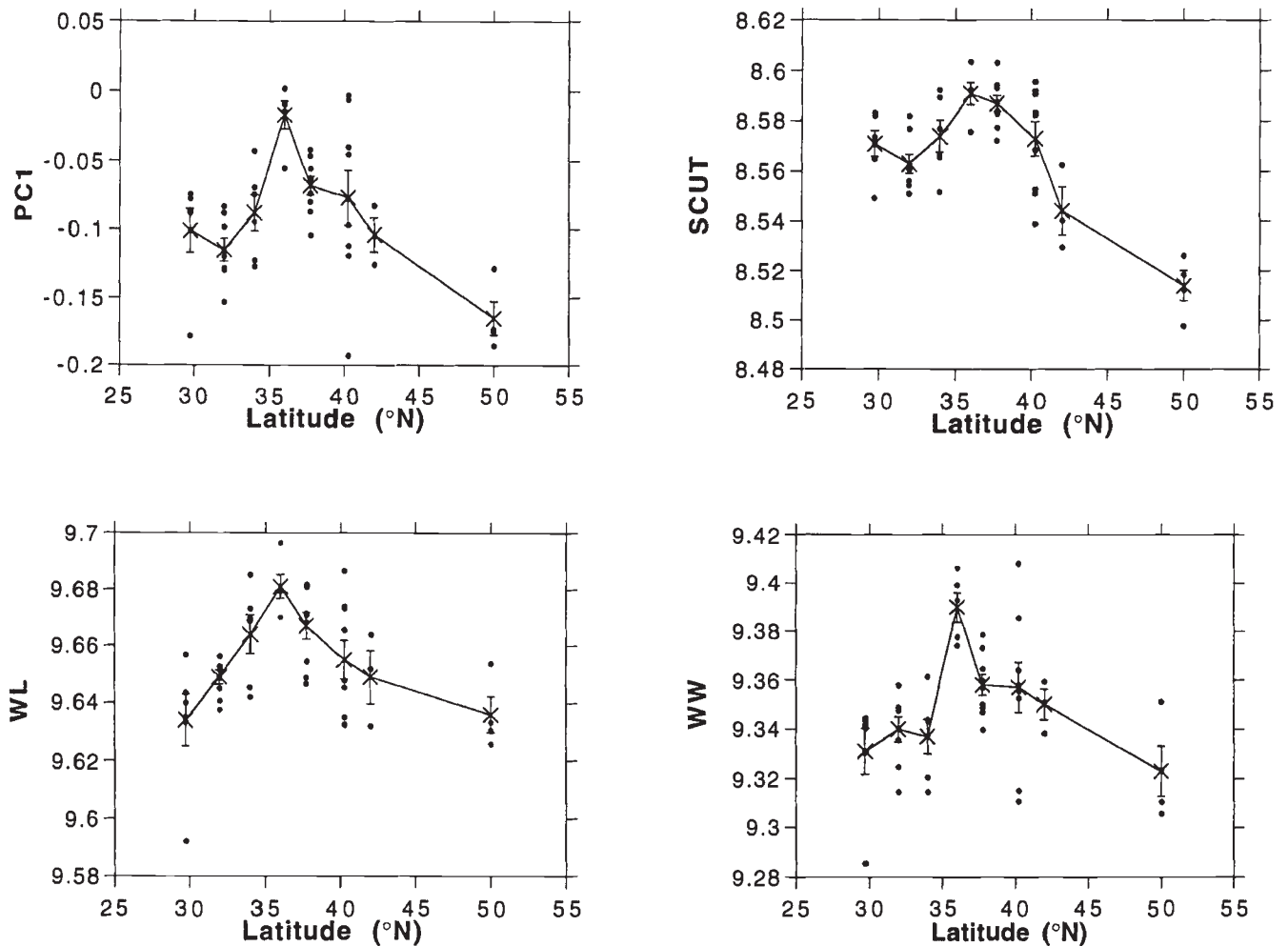‡Estimates of coefficients for Latitude are times 100, and estimates of coefficients for (Latitude)² are times 10^4.

Fig. 2 Graphs of the trend in phenotypes over distance for selected male characters in *Drosophila melanogaster*. Phenotypic values are in transformed units as described in the text; latitudes and sample sizes correspond to those given in Table 1. ● represent the line means for each location, and 'X' is the average phenotype over the line means for that location, with bars representing standard errors.

constants $(\chi = \{k, F\})$, each of the ten morphological characters examined, and all interactions. This model accounts for 45 per cent of the total observed variation in F_{st}, the rest of the variation representing the actual error in F_{st} estimates within a given 'treatment'. Although all effects and interactions are significant, the choice of constants, χ, and interactions involving χ account for only 4.9 per cent of the total variation in the model. Thus, the effect corresponding to the choice of constants was removed from the model, and subsequent analyses are on the residuals generated in this manner (Sokal & Rolf, 1981). The interaction between character and comparison is the largest effect, accounting for 50 per cent of the total variation in the model. That is, the pattern of population subdivision between adjacent populations observed in this study is different for each character, and is highly statistically significant.

Table 6 gives the estimated allele frequencies of the faster migrating allele at each of the seven allozyme loci examined, and the number of isofemale lines surveyed. Figure 3, a subset of the data presented in Table 5, gives plots of F_{st} estimated for the morphological characters against F_{st} estimated for the allozymes, with corresponding standard deviations derived from the bootstrapping of the data. The standard deviations on F_{st} are large, as one would expect the variance of a variance estimate to be. Both PC1 and WW have one population comparison for which the morphological F_{st} is not proportional to the allozyme F_{st}, the aberrant comparison being between the Nashville and Louisville populations. This discordance does not exist for all characters, and is absent in characters such as WL and SCUT. It is interesting to contrast this pattern with the patterns of clinal variation shown in Fig. 2. It appears

Table 4 ANOVA table showing results of analysis of bootstrapped estimates of Wright's F_{st}

Source	d.f.	Sum of squares	Mean square	F
Cmp	6	1120.71	186.78	1387.30***
χ	4	15.05	3.76	27.94***
Chr	9	568.93	63.21	469.51***
Cmp $\times \chi$	24	8.93	0.37	2.76***
Cmp $\times$ Chr	54	1887.87	34.96	259.66***
$\chi \times$ Chr	36	51.67	1.44	10.66***
Cmp $\times \chi \times$ Chr	216	108.80	0.50	3.74***
Error	34 650	4665.23	0.13	
Corr. Total	34 999	8427.19		

Cmp is the effect of adjacent populations Compared, χ is the effect of the set of constants chosen for the parameters k and F (see text), and Chr is the effect of various morphological Characters examined. All ANOVA was carried out on arcsine transformed data derived from 100 bootstrap replicates of the data set per comparison, character, constant combination (see text).
***indicates significance of F-tests at $P < 0.001$.

that WW and PC1 have relatively high levels of population structure for the Nashville/Louisville comparison relative to a 'neutral' estimate of population subdivision derived from allozyme frequencies.

Table 5 also gives the percentage of variation in the data attributable to the 'populations compared' by 'molecules versus morphology' $(\alpha\beta_{ij})$ interaction effect of an ANOVA. The significance of this interaction effect reflects the lack of proportionality between allozyme and character F_{st} over the pairs of populations compared and is significant for all characters. This term is also biologically relevant as it accounts for 5 (SCUT) to 41 (WW) per cent of the total variation in the data set after removing the main effect of the comparison of allozyme vs. morphological F_{st} (whose effect is somewhat arbitrary after the removal of choice of constants (χ)). In a number of cases the variation attributable to this interaction is about the same size as the variation attributable to the effect of populations compared (e.g. PC1, WL, WW, FW, FL). This implies that for a given set of populations compared the difference between the morphological and allozyme F_{st} can be as large as the average difference between pairs of populations compared within a given trait.

Discussion

The present study examines the extent of clinal variation in seven morphological characters in laboratory-reared *Drosophila melanogaster*. This study differs from previous studies in that collection sites were carefully chosen to minimize extraneous sources of varia-

tion such as season and longitude. In addition, the length of the transect over which the flies were collected was maximized in order to observe variation over much of *D. melanogaster*'s North American distribution. Congruently, the distance between collection sites was approximately uniform so that nonmonotonic trends, if present, could be observed. In short, a transect was sampled which was comparable to that of Coyne & Beecham (1987), but further inland so that the two transects are independent. The above sampling scheme allows an in-depth study of the nature of clinal variation, and, indirectly, adaptation in *D. melanogaster*.

Nonmonotonic clinal variation

Extensive studies on both dispersal rates (Coyne et al., 1987; Coyne & Milstead, 1987) and levels of genetic differentiation at molecular marker loci (see reviews by: David & Capy, 1988; Singh, 1989; Singh & Long, 1992) indicate that migration rates are high among various populations of *D. melanogaster* within North America. The average absolute number of migrants between any two populations per generation (Nm) within North America has been estimated to be between two and six (Singh & Rhomberg, 1987a,b). Based on the average F_{st} between adjacent populations, the estimate of Nm between adjacent populations from this study is 9.4. This is consistent with earlier studies, as the average distance between adjacent populations in this study is smaller than the average distance between populations in Singh &

Table 5 Values of arcsine transformed Wright's F_{st} for morphological characters and allozymes of Drosophila melanogaster

Cmp	Allozyme	Character								
		PC1	PC2	WL	WW	FEM	TIB	SCUT	FW	FL
1 & 2	0.27 (0.070)	0.99 (0.562)	0.41 (0.533)	0.40 (0.555)	0.61 (0.523)	1.05 (0.557)	1.16 (0.614)	0.73 (0.488)	0.18 (0.313)	0.06 (0.253)
2 & 3	0.15 (0.056)	0.16 (0.240)	0.15 (0.209)	0.11 (0.253)	0.10 (0.242)	0.49 (0.508)	0.25 (0.351)	0.51 (0.317)	0.11 (0.246)	0.06 (0.228)
3 & 4	0.13 (0.035)	0.10 (0.167)	0.26 (0.417)	0.28 (0.312)	0.09 (0.168)	0.28 (0.321)	0.38 (0.369)	0.30 (0.272)	0.19 (0.169)	0.23 (0.406)
4 & 5	0.11 (0.018)	1.04 (0.459)	0.12 (0.331)	0.44 (0.364)	0.97 (0.299)	1.00 (0.385)	0.93 (0.377)	0.14 (0.359)	0.10 (0.261)	0.55 (0.645)
5 & 6	0.21 (0.029)	0.81 (0.233)	0.04 (0.237)	0.38 (0.281)	0.98 (0.205)	0.88 (0.334)	0.84 (0.336)	0.52 (0.408)	0.04 (0.206)	0.81 (0.663)
6 & 7	0.13 (0.029)	0.34 (0.281)	0.75 (0.655)	0.55 (0.470)	0.11 (0.212)	0.68 (0.428)	0.43 (0.495)	0.28 (0.326)	0.73 (0.216)	0.32 (0.468)
7 & 8	0.13 (0.031)	0.20 (0.366)	0.23 (0.394)	0.35 (0.385)	0.17 (0.237)	0.14 (0.324)	0.08 (0.252)	0.21 (0.393)	0.41 (0.283)	0.12 (0.287)
Total	0.56 (0.033)	0.67 (0.159)	0.75 (0.564)	0.62 (0.186)	0.58 (0.145)	1.02 (0.247)	1.05 (0.335)	0.90 (0.197)	0.47 (0.145)	0.38 (0.440)
% $\alpha\beta_{ij}$	28	28	12	12	41	18	20	5	31	28

Values of arcsine transformed F_{st} with their corresponding standard deviations (500 bootstrap replicates per estimate, χ removed) following in parentheses listed for the different pairs of populations compared. Total, is F_{st} calculated over all populations. The bottom portion of the table gives the percentage of variation due to the $\alpha\beta_{ij}$ interaction term (d.f. = 6) relative to the total variation in the data set less the variation due to the main effect of morphological vs. allozyme F_{st} (see text). The $\alpha\beta_{ij}$ interaction term reflects the lack of proportionality between morphological and allozyme F_{st} and is significant at $P < 0.001$ for all characters.

Rhomberg (1987a,b). A result of classical population genetics theory is that gene flow much less than that observed here is sufficient to prevent genetic differentiation of local populations in the absence of selection (Wright, 1931). Thus it appears unlikely that the non-monotonic cline in morphological characters observed in this study is simply the result of random genetic drift between isolated populations, as gene flow between populations appears to be extensive. Previous studies have also shown genetical distance to be a monotonic function of geographical distance in *D. melanogaster* (Singh & Rhomberg, 1987b). The allozyme data of this study confirm this result, as a regression of Nei's genetic distance on geographical distance showed a linearly increasing trend (data not shown). This result implies that the resemblance of northern and southern populations in morphology is probably not due to re-establishment of northern populations by southern populations each year, and that northern populations are probably endemic.

Temperature may act directly on the fly's physiology or indirectly on traits such as desiccation resistance in larval flies (Levins, 1969; Coyne *et al.*, 1983). That temperature is the selective force maintaining an observed cline is supported by selection cage experiments which show a change in the body size of flies reared at different temperatures (Anderson, 1966, 1973; Powell, 1974). If temperature is responsible for the maintenance of the observed cline, it is expected that the trend in either yearly temperatures or seasonal temperatures will correspond to the trend observed in the morphological characters. No correspondence suggests that the relationship between temperature and selection on quantitative traits is more complex, or that the characters are responding to a second selective agent. The results of Coyne & Beecham (1987), who observed a monotonic cline extending as far north as the current study, appear to confirm this hypothesis. Nonetheless, an examination of the trends in temperature (i.e. monthly average daily mean, minimum, maximum, and maximum minus minimum for all three months prior to collection; a three-month average of the above; a three-year average of the same values; and 30-year normals for the same) reported by the nearest weather station to the collection site showed all measures of temperature to be monotonic decreasing from south to north (data not shown but available from the authors). As the above temperature measures are all highly correlated with latitude it was decided not to analyse weather patterns further. The observed weather patterns along the transect west of the Appalachians of this study and the observed nonmonotonic cline suggest that the mechanism responsible for maintaining the cline is not simply temperature.

Table 6 Allele frequencies at seven allozyme loci for eight collections of
Drosophila melanogaster

	Adh	*Fuc*	*G6pd*	*Gpd*	*Odh*	*6Pgd*	*Pgm*
High Springs (28)	0.679	0.571	1.000	0.179	0.214	0.571	0.679
Cordele (32)	0.688	0.500	1.000	0.067	0.094	0.563	0.875
Cartersville (42)	0.595	0.357	0.976	0.262	0.048	0.450	0.857
Nashville (24)	0.708	0.625	0.792	0.083	0.208	0.292	0.833
Louisville (12)	0.542	0.625	0.875	0.125	0.182	0.167	0.708
Dayton (30)	0.667	0.333	0.900	0.133	0.100	0.333	0.667
Windsor (16)	0.625	0.438	0.600	0.125	0.000	0.313	0.688
Winnipeg (10)	0.300	0.700	0.600	0.200	0.400	0.100	0.700

Frequencies of the faster migrating alleles are given with the exception of *6Pgd* and
Pgm, where middle allele frequency is given. Numbers in parentheses represent the
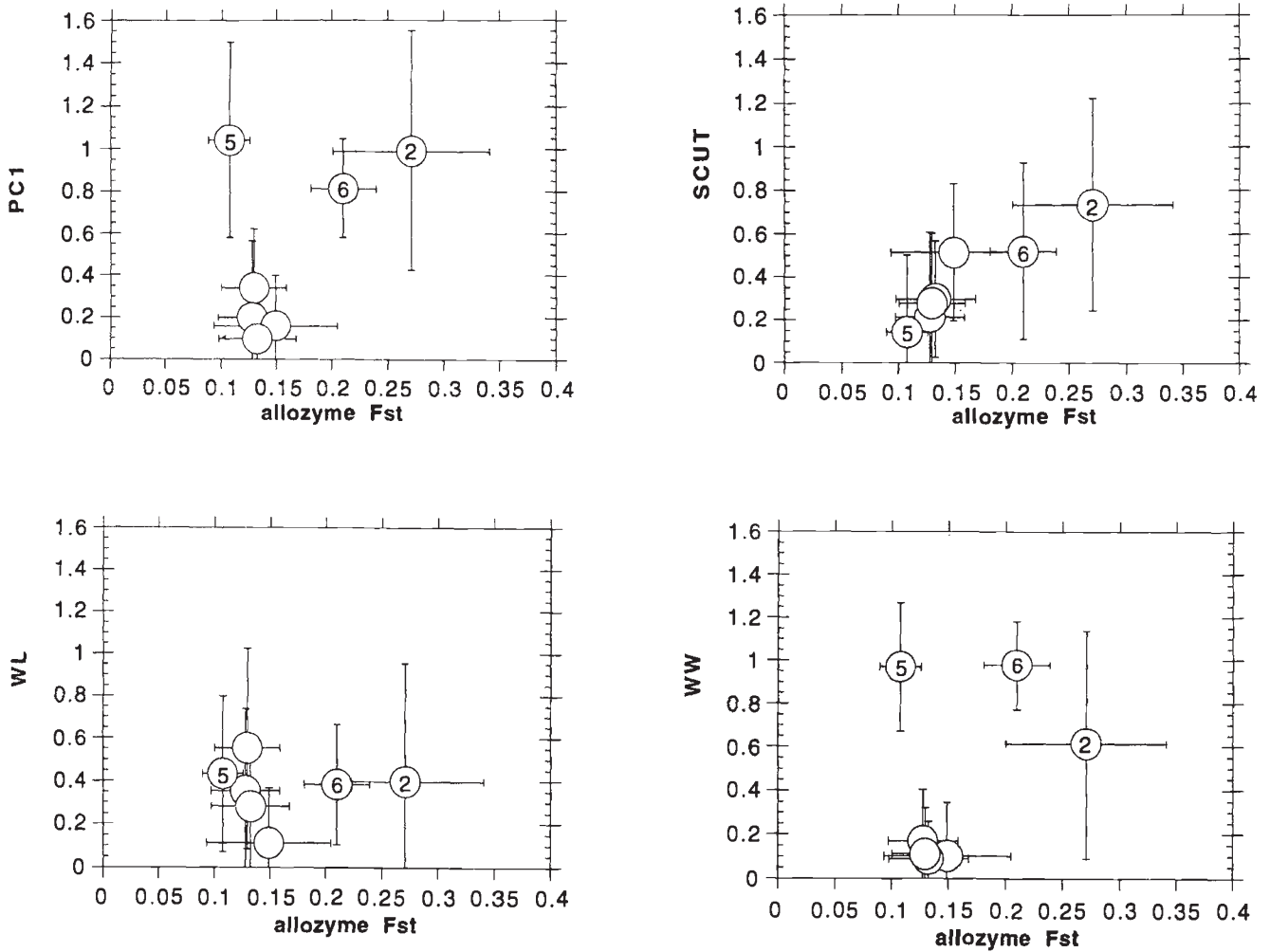number of isofemale lines sampled per location.

Fig. 3 Plots of arcsine transformed F_{st} for morphological characters against allozyme F_{st} for selected characters of *Drosophila melanogaster*. Error bars represent one standard deviation as estimated from the bootstrap analysis. Numbers in circles are the population number of the southernmost of a given set of adjacent populations compared.

The present study did not observe the same monotonic clines that have been previously reported. This is not surprising when one considers the sampling schemes used in other studies where flies were collected: at irregular intervals (Stalker & Carson, 1947; Sokoloff, 1965), from too few locations (Hyytia *et al.*, 1985), without regards for longitude or altitude (Stalker & Carson, 1947; Sokoloff, 1965), and at different times of the year (Stalker & Carson, 1947; Sokoloff, 1965; Hyytia *et al.*, 1985; Watada *et al.*, 1986). Any or all of these factors could have added variation which would tend to obscure a nonmonotonic trend (Stalker & Carson, 1948, 1949). The study of Coyne & Beecham (1987), which sampled flies along the eastern seaboard, and the present study have none of these sampling problems: one shows a monotonic trend, the other a nonmonotonic trend. Figure 4 compares the mean wing length measures (the only comparable character) in laboratory-reared flies in Coyne & Beecham's study (averaged over three temperatures) and our own. It is apparent that the cline of Coyne & Beecham is much less regular than our cline and the three northernmost points in the Coyne & Beecham study could be construed as starting a downward trend. Nonetheless, this is impossible to ascertain, because there is a great deal of variability in Coyne & Beecham's trend, and their transect was not extended further north. The maximum wing length in Coyne & Beecham's study is also further north than in the present study. This is consistent with the yearly average temperature being higher and more variable on the coast than inland at any given latitude (data not

shown). It is possible that nonmonotonic clines are not infrequent and have not been detected in the past simply because of limitations in sampling schemes.

Population structure for molecules and morphology

This study shows considerable differences in the patterns of population subdivision for each morphological character. This implies that not all characters are responding to the same selective forces between different sets of adjacent populations. It is apparent that PC1 and WW have different patterns of population subdivision from WL and SCUT. The relatively high levels of F_{st} for comparisons involving population 5 (Nashville) could be the result of selection maintaining a large difference in PC1 and WW in the face of high levels of gene flow. Comparing levels of F_{st} between adjacent populations along a cline provides a novel method of detecting selection acting on quantitative traits. The power of this approach comes from the selection of multiple populations which can be compared. This approach has different strengths from retrospective selection analysis, where usually only two populations are compared, and where accurate estimates of the genetic covariance matrix are needed (Lande, 1979). Unlike retrospective selection analysis, estimates of the actual selection coefficients on a given character are not measured, and hence our method is a qualitative one. But, as the errors on selection coefficients estimated by retrospective selection analysis are large, their interpretation is often considered to be qualitative at best.

Estimates of gene frequencies at a series of allozyme loci allow estimates of population subdivision for morphological characters to be compared with 'neutral' estimates of the same parameter. It is evident from Fig. 3 that WL and SCUT show patterns of population subdivision that are proportional to the levels of subdivision seen in the allozymes. This relationship is true even though each allozyme generally showed either no clinal trend, or monotonically increasing or decreasing trends over the cline (Table 6). These results stand in contrast to those for WW and PC1, in which one particular population comparison showed a relatively high level of subdivision relative to the estimate of 'neutral' variation measured from the allozymes. This pattern was confirmed by the statistical analysis, as the 'populations compared' by 'molecules vs. morphology' effect accounted for more variation in F_{st} for PC1 and WW than for WL and SCUT. It thus seems likely that if the pattern of variation in allozymes is neutral, there is strong natural selection maintaining differences in WW and PC1 between Nashville and Louisville (populations 4 and 5). As the F_{st} for other population comparisons in morphological characters is largely
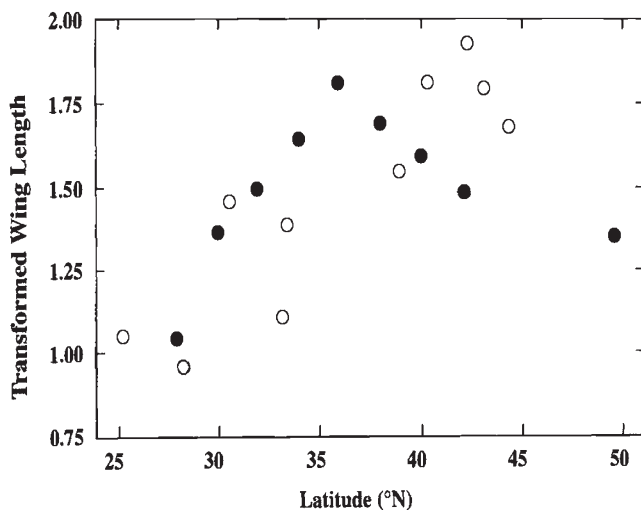

Fig. 4 A comparison of *Drosophila melanogaster* wing length against latitude for the present cline (●) and Coyne and Beecham's (1987) cline (○). Both measures were linearly transformed to put them on the same scale. Coyne and Beecham's data are averages over three temperatures.

proportional to that for allozymes it is difficult to detect selection acting elsewhere in the cline. Thus it appears that this technique is limited by the large errors implicit in the estimation of F_{st}.

It is possible that the allozymes of this study are not actually neutral markers, and that their spatial variation is the result of selective forces (Berry & Kreitman, 1993). This is particularly true for allozymes which have shown clines in previous studies and this study (i.e. *Adh, G6pd, 6Pgd*) (Singh & Rhomberg, 1987b). If this is indeed the case then truly neutral data, from silent sites for example, would be desirable in future studies. Nonetheless, it is possible that the behaviour of selected loci, if selection is weak and adjacent populations have high levels of gene flow between them, will approach that of the neutral case. If the pattern of F_{st} for allozyme loci between adjacent populations is largely the result of selective forces, then one could still infer selection acting along the cline. In this situation, one would conclude that the selective forces maintaining the cline act *differently* on allozymes and morphological characters.

The results of this paper show that in nature two populations can experience considerable gene flow between them yet *still* be different for a quantitative trait. A second important result is that not all characters are responding in the same manner to natural selection acting to differentiate two populations. In particular, patterns of population subdivision for a neutral marker differ from those for morphological characters. Such differences are not expected if the morphological characters are neutral. Further experimental and theoretical work is needed in order to determine what proportion of genetic variation is maintained by spatial variation in the selective forces governing quantitative traits, how much selection is needed to maintain phenotypic differences in spite of high migration, and how robust such models are to the genetical assumptions underlying quantitative traits. Specifically, it is important to determine if spatial variation in phenotypic optima coupled with high migration is a viable alternative to mutation–selection balance for maintaining variation in nature (Barton & Turelli, 1989). Additional refinements in the statistical techniques used to relate morphological to molecular F_{st} may also increase the power of this method.

Acknowledgements

References

ANDERSON, W. W. 1966. Genetic divergence in M. Vetukhiv's experimental populations of *Drosophila pseudoobscura*. 3. Divergence in body size. *Genet. Res.*, 7, 255–266.

ANDERSON, W. W. 1973. Genetic divergence in body size among experimental populations of *Drosophila pseudoobscura* kept at different temperatures. *Evolution*, 27, 278–284.

ASHBURNER, M. 1989. *Drosophila: A Laboratory Handbook.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

BARTON, N. H. AND TURELLI, M. 1989. Evolutionary quantitative genetics: how little do we know? *Ann. Rev. Genet.*, 23, 337–370.

BERRY, A. AND KREITMAN, M. 1993. Molecular analysis of an allozyme cline: alcohol dehydrogenase in *Drosophila melanogaster* on the east coast of North America. *Genetics*, 134, 869–893.

CHATFIELD, C. AND COLLINS, A. J. 1980. *Introduction to Multivariate Analysis.* Chapman and Hall, New York.

COYNE, J. A. AND BEECHAM, E. 1987. Heritability of two morphological characters within and among natural populations of *Drosophila melanogaster. Genetics*, 117, 727–737.

COYNE, J. A., BRYANT, S. N. AND TURELLI, M. A. 1987. Long-distance migration of *Drosophila*. 2. Presence in desolate sites and dispersal near a desert oasis. *Am. Nat.*, 129, 847–861.

COYNE, J. A. BUNDGAARD, J. AND PROUT, T. 1983. Geographic variation of tolerance to environmental stress in *Drosophila pseudoobscura. Am. Nat.*, 122, 474–488.

COYNE, J. A. AND MILSTEAD, B. 1987. Long-distance migration of *Drosophila*. 3. Dispersal of *D. melanogaster* alleles from a Maryland orchard. *Am. Nat.*, 130, 70–82.

CROW, J. F. AND KIMURA, M. 1970. *An Introduction to Population Genetics Theory.* Burgess Publishing Company, Minneapolis, MN.

DAVID, J. R. AND CAPY, P. 1988. Genetic variation of *Drosophila melanogaster* natural populations. *Trends Genet.*, 4, 106–111.

ENDLER, J. A. 1977. *Geographic Variation, Speciation, and Clines.* Monographs in Population Biology. Princeton University Press, Princeton, NJ.

FALCONER, D. S. 1981. *Introduction to Quantitative Genetics*, 2nd edn. Longman, New York.

GUPTA, A. P. AND LEWONTIN, R. C. 1982. A study of reaction norms in natural populations of *Drosophila pseudoobscura. Evolution*, 36, 934–948.

HEBERT, P. D. N. AND BEATON, M. J. 1989. *Methodology for Allozyme Analysis using Cellulose Acetate Electrophoresis.* Helena Laboratories, Beaumont, TX.

HOFFMAN, A. A. AND NIELSEN, K. M. 1985. The effect of resource allocation on genetic variation in *Drosophila. Am. Nat.*, 125, 421–430.

HYYTIA, P., CAPY, P., DAVID, J. R. AND SINGH, R. S. 1985. Enzymatic and quantitative variation in European and African populations of *Drosophila simulans. Heredity*, 54, 209–217.

JAMES, F. C. 1983. Environmental component of morphological differentiation in birds. *Science*, 221, 184–186.

LACHAISE, D., CARIOU, M.-L., DAVID, J. R., LEMEUNIER, F., TSACAS, L. AND ASHBURNER, M. 1988. Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evol. Biology*, **22**, 159–225.

LANDE, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution*, **33**, 402–416.

LANDE, R. AND ARNOLD, S. J. 1983. The measurement of selection on correlated characters. *Evolution*, **37**, 1210–1226.

LEVINS, R. A. 1969. Thermal acclimation and heat resistance in *Drosophila* species. *Am. Nat.*, **103**, 483–499.

LEWONTIN, R. C. AND KRAKAUER, J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.

LONG, A. D. 1991. *Quantitative Genetics of Clinal Variation*. MSc. Thesis, McMaster University.

MISRA, R. K. AND REEVE, E. C. R. 1964. Clines in body dimensions in populations of *Drosophila subobscura*. *Genet. Res.*, **5**, 240–256.

POWELL, J. R. 1974. Temperature related genetic divergence in *Drosophila* body size. *J. Hered.*, **65**, 257–258.

PRICE, P. T., GRANT, P. R., GIBBS, H. L. AND BOAG, P. T. 1984. Recurrent patterns of natural selection in a population of Darwin's finches. *Nature*, **309**, 787–789.

PROUT, T. 1958. A possible difference in genetic variance between wild and laboratory populations. *Drosoph. Inf. Serv.*, **32**, 148–149.

PROUT, T. AND BARKER, J. S. F. 1993. *F* statistics in *Drosophila buzzatii*: selection, population size and inbreeding. *Genetics*, **134**, 369–375.

REYMENT, R. A., BLACKITH, R. C. AND CAMPBELL, N. A. 1984. *Multivariate Morphometrics*, 2nd edn. Academic Press, Toronto.

ROBERTSON, A. 1975. Gene frequency distributions as a test of selective neutrality. *Genetics*, **81**, 775–785.

ROBERTSON, F. W. 1957. Studies in quantitative inheritance. XI. Genetic and environmental correlations between body size and egg production in *Drosophila melanogaster*. *J. Genet.*, **55**, 428–443.

SINGH, R. S. 1989. Population genetics and evolution of species related to *Drosophila melanogaster*. *Ann. Rev. Genet.*, **23**, 425–453.

SINGH, R. S. AND LONG, A. D. 1992. Geographic variation in *Drosophila*: from molecules to morphology and back. *Trends Ecol. Evol.*, **7**, 340–345.

SINGH, R. S. AND RHOMBERG, L. R. 1987a. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. I. Estimates of gene flow from rare alleles. *Genetics*, **115**, 313–322.

SINGH, R. S. AND RHOMBERG, L. R. 1987b. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. II. Estimates of heterozygosity and patterns of geographic differentition. *Genetics*, **117**, 255–271.

SOKAL, R. S. AND ROLF, F. J. 1981. *Biometry*, 2nd edn. W. F. Freeman and Co., San Francisco.

SOKOLOFF, A. 1965. Geographic variation in quantitative characters in populations of *Drosophila pseudoobscura*. *Evolution*, **19**, 300–310.

SPITZE, K. 1993. Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.

STALKER, H. D. AND CARSON, H. L. 1947. Morphological variation in natural populations of *Drosophila robusta* Sturtevant. *Evolution*, **1**, 237–248.

STALKER, H. D. AND CARSON, H. L. 1948. An altitudinal transect of *Drosophila robusta* Sturtevant. *Evolution*, **2**, 295–305.

STALKER, H. D. AND CARSON, H. L. 1949. Seasonal variation in the morphology of *Drosophila robusta* Sturtevant. *Evolution*, **3**, 330–343.

TURELLI, M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variation. *Evolution*, **42**, 1342–1347.

WATADA, M., OHBA, S. AND TOBARI, Y. N. 1986. Genetic differentiation in Japanese populations of *Drosophila simulans* and *D. melanogaster*. *Jap. J. Genet.*, **61**, 469–480.

WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics*, **16**, 97–159.