Genet. Res., Camb. (1993), **62**, *pp.* 63–75 *With 4 text-figures*

Temperature related fertility selection on body size and the sex-ratio gene arrangement in *Drosophila pseudoobscura*

M. D. GEBHARDT*† AND W. W. ANDERSON
Department of Genetics, University of Georgia, Athens, GA 30602

(*Received 13 October 1992 and in revised form 3 February 1993*)

Summary

We measured temperature-dependent fertility selection on body size in *Drosophila pseudoobscura* in the laboratory. One hundred single females of each of the three karyotypes involving the 'sex-ratio' (SR) and the standard (ST) gene arrangement on the sex chromosome laid eggs at either 18 or 24 °C. The experiment addressed the following hypotheses: (*a*) Fertility selection on body size is weaker at the higher temperature, explaining in part why genetically smaller flies appear to evolve in populations at warmer localities. (*b*) Homokaryotypic SR females are less fecund than homokaryotypic ST females, possibly mediated by the effect of body size on fertility, explaining the low frequencies of SR despite its strong advantage due to meiotic drive. The data were also expected to shed light on a mechanism for the evolution of plasticity of body size through fertility selection in environments with an unpredictable temperature regime. Hypothesis (*a*) was clearly refuted because phenotypically larger ST females had an even larger fertility surplus at the higher temperature and, more importantly, the genetic correlation between fertility and body size disappeared at the lower temperature. As to (*b*), we found that temperature affects fertility directly and indirectly through body size such that ST and SR females were about equally fecund at both temperatures, although different in size and size-adjusted fertility. We observed heterosis for both size and fertility, which might stabilize the polymorphism in nature. The reaction norms of body size to the temperature difference were steeper for ST females than for SR females, implying that fertility selection could change phenotypic plasticity of body size in a population. Selection on body size depended not only on the temperature, but also on the karyotypes, suggesting that models of phenotype evolution using purely phenotypic fitness functions may often be inadequate.

1. Introduction

Body size is a trait central to the life history of all organisms (Schmidt-Nielson, 1984). The many implications of body size for biomechanics, physiology, or ecology provide mechanisms of selection for both smaller and larger size which combine to give a stabilizing selection towards an optimal size (Roff, 1981). In the present study this point of view is taken to approach the ecogeographical rule that body size varies in a regular way with environmental temperature (Mayr, 1963). For example, genetic variation in body dimensions among geographic strains has been reported for several *Drosophila* species (Stalker & Carson, 1947; Prevosti, 1955; Misra & Reeve, 1964). Flies were often found to be larger when derived from cooler localities.

There is another line of evidence that indicates that increasing the temperature shifts the balance of selection on body size such that a smaller size confers the highest fitness. Anderson (1966, 1973) reported that experimental populations of *D. pseudoobscura* founded from the same base population diverged genetically with respect to wing length as a response to a difference in temperature at which the populations were kept. Similar observations were made in studies on experimental populations of two other species: *D. melanogaster* that were kept at two different temperatures for many years (Cavicchi *et al.* 1985, 1989) and *D. willistoni* which were tested after 2·5 years (Powell, 1974).

Fecundity (or fertility) is one of the major fitness components in *Drosophila* females. To the degree that it depends on body size, it is likely to be important in

* Present address: Federal Office of Public Health, Epidemiology, Hess-Str. 27E, CH-3097 Bern/Liebefeld, Switzerland.
† Corresponding author.

determining the optimum for this trait since it influences the balance between selection towards larger and smaller size. The phenotypic relationship between body size and fecundity is generally positive (Tantawy & Vetukhiv, 1960), so it is expected that fecundity (or fertility) selection always favours larger body size. However, the strength of fecundity selection on body size could be different at different temperatures, as total lifetime fecundity, irrespective of size, also depends on temperature (Tantawy & Vetukhiv, 1960). If the relationship between fertility and body size were weaker at higher temperatures, one would expect that a smaller body size would become more favourable, because the counterbalancing forces such as selection for shorter developmental time would become relatively more important.

In the present study, different genotypes involving three different stocks polymorphic for the sex-ratio (SR) and standard (ST) gene arrangement on the X-chromosome of *D. pseudoobscura* were used to measure the strength of fertility selection on body size at two temperatures. The SR system was chosen because a strong effect on female fertility associated with the SR gene arrangement is known (Wallace, 1948; Curtsinger & Feldman, 1980). The data promised to give insight into the mechanisms that stabilize this polymorphism in nature and the reasons why the frequency of the SR chromosome shows clines that are probably related to the environmental temperature (Sturtevant & Dobzhansky, 1936; Wallace, 1948).

The most remarkable effect associated with the SR chromosome is that Y-bearing sperms degenerate in males hemizygous for it (Policansky & Ellison, 1970). Such males give rise to progenies that almost exclusively consist of females, all of which carry the SR chromosome. The phenomenon is often referred to as 'meiotic drive', although the bias is due to a post-meiotic mechanism, and it acts to increase the frequency of the SR arrangement in natural populations. Because the observed frequencies are usually low in natural populations (around 10% or less, Wallace, 1948), and because the chromosome tends to disappear in population cages (Beckenbach, 1983), a counterbalancing selective disadvantage of SR has to be postulated. There is some evidence that females homokaryotypic for SR have a fertility disadvantage relative to standard females and that it is temperature-dependent (Wallace, 1948; Curtsinger & Feldman, 1980; Beckenbach, 1983). Our study was designed to examine whether such a dependency could be verified, and whether the mechanism for the reduced fertility of SR females could be related to body size.

Finally, our experiment provided an opportunity to examine whether the different stocks responded to temperature differences differently. It is a necessary condition for the evolution of phenotypic plasticity that genotypes have different responses to environmental variables (Gupta & Lewontin, 1982; Stearns & Koella, 1986; Gebhardt & Stearns, 1988). It is convenient to use the analysis of reaction norms in this context, because they are the functions that describe the phenotypic responses of defined genotypes to environmental variables.

2. Materials and methods

(i) *Flies*

Three stocks of *D. pseudoobscura* (labelled SR9, SR13 and SR16), each containing one version of the SR chromosome, were used for the experiment. The stocks originated from single SR males and wild-type (ST) females collected several years earlier in California. To facilitate the presentation, we depart in the following from the conventional notation for the different karyotypes. We use the letter 'S' to denote an SR chromosome and the letter 'X' to denote an ST chromosome. For example, we use 'SX' instead of 'SR/ST' to denote heterokaryotypic females. Figure 1 shows the crossing scheme that was followed every generation to maintain the stocks. Note that because each SR stock was initially derived from a single SR male, all the SR chromosomes within a stock are identical, but presumably different among stocks. The ST lines of each stock were derived from single inseminated females caught in the same populations and at the same time as the males that provided the SR chromosomes and were never mixed among stocks. Thus, each SR stock was unique by its single version of an SR chromosome and an autosomal background from an isofemale line. Because of the repeated crosses within stock among the SR, mixed and ST lines for many generations, the three lines shared the autosomal background, but differed in the number of SR chromosomes present in the females: two in the SR line, one in the mixed line, and none in the ST line. A within-stock comparison of the different lines thus allowed us to analyse the effect of individual SR

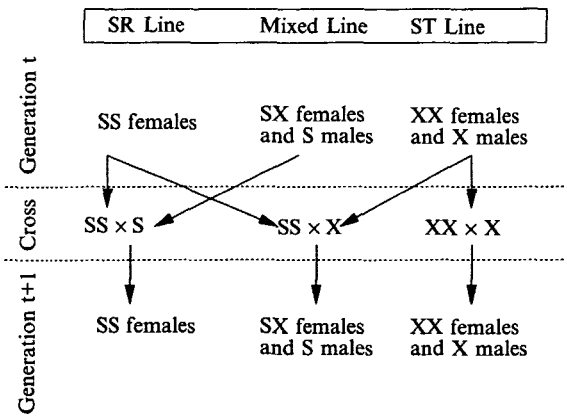

Fig. 1. Crossing scheme for the maintenance of the sex-ratio stocks. A 'stock' was maintained in three lines, as shown. Notation: 'S' designates a sex-ratio (SR) chromosome, 'X' designates a standard sex chromosome. For males, the symbol for the Y chromosomes is left out. The experimental females were obtained in the same way.

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

chromosomes as compared to ST, without confounding effects of different autosomal backgrounds among different karyotypes.

The medium used to keep the stocks and in the experiments was based on cornmeal, baker's yeast and molasses with propionic acid as a mould inhibitor.

(ii) *Experiment*

The aim was to measure the fertility of individual females of each karyotype (SS, SX and XX) in each of the three stocks (SR9, SR13 and SR16) at two temperatures (18 and 24 °C). To avoid the possibility that differences among stocks or karyotypes could be generated by non-genetic factors such as common environment or maternal effects, the females used in the experiment were obtained by following the crossing scheme shown in Fig. 1 for two generations under standardized conditions at the two experimental temperatures. Each cross in the scheme was replicated in at least three bottles in the first round, in at least ten in the second.

The experimental females were collected into bottles during 3 days and aged thereafter for another 4 days to ensure that all were fully mature at the start of the experiment. The females thus were between 4 and 7 days old, which was not different among the stocks or karyotypes, because hatching rates were similar. The experiment consisted of having 100 single females of each type lay eggs at one of the two temperatures on the surface of about 20 ml of medium. The area provided for egg laying was about 7 cm^2. Each female was combined with two ST males of her own stock in a vial and kept at the appropriate temperature. Each vial was checked daily for the first presence of eggs. Only when a female started egg laying within 6 days was she further monitored in the experiment. In all stocks and karyotypes, most females started egg laying between days 2 and 3, more than 90% having started by day 4. The vials with the females were placed into one climate chamber for each temperature ($\pm 0{\cdot}5$ °C), randomized with respect to stock or karyotype. The vials were freshly repositioned in a random fashion every day during egg laying.

(iii) *Fertility estimation*

The large number of females prevented daily egg counts. A different approach was therefore chosen that employed counts of larvae hatching from the eggs. This means that our measure ('fertility') is not strictly referring to the number of eggs laid in a given time (fecundity), but also includes hatchability and also some contribution of early larval viability. Viability had probably a very small influence because the time between hatching and sampling took a small proportion of the total larval development. For hatchability, one could argue that the inclusion is even preferable over direct egg counts for our purposes since it likely has a large component that depends on the genotypes of the females, rather than the genotypes of the zygotes (e.g. provision of the egg with maternal RNA, or efficiency of the fertilization with stored sperm). To the degree that differences among stocks, karyotypes, or the effects of temperature on these components have contributed to the larval counts, the comparison with previous studies using direct egg counts (fecundities) would be hampered. However, since pilot experiments suggested that the difference between our measures and direct egg counts would rarely have exceeded 5% (see also below), we believe that the comparison of the females using the larval counts closely resembled the comparison that would have been obtained with direct egg counts.

It was intended to have a fertility estimate over the longest possible length of time in order to average out the typical day-to-day variation in egg laying rate. The females were allowed to lay eggs for 4 days after the first eggs had been observed. Pilot trials had shown that the larvae hatching from the first eggs would not burrow into the medium during the first 4 days. After that, they would bury eggs that were laid later and prevent them from hatching. Any possible inhibiting effect of larval activity on the females was probably small during the first 4 days because the number and size of the larvae were small. Since the females (regardless of stock or karyotype) were between 5 and 13 days old when they started egg laying (see previous section), the time interval of 4 days coincides with the reported peak of egg laying rate in *D. pseudoobscura* (Tantawy & Vetukhiv, 1960).

The adults were removed from the vials after the egg-laying time. Females were immediately frozen and stored in Eppendorf tubes at -20 °C for later measurements. The males were discarded. The vials were kept for another 2 or 3 days (for the higher or lower temperature) to allow the last larvae to hatch and grow for better countability. The larvae were separated from the medium by using a 27% sucrose solution that caused the larvae to float and the medium (gently stirred) to sink. The floating larvae were collected with a suction device connected to a waterline vacuum and retained on a filter. They were stored in 70% ethanol until they could be counted on a counting grid. From previous trials involving direct egg counts it was estimated that the error of this method (lost larvae) was not larger than about 5%, even for the smallest larvae and did not depend on either genotypes or temperature (unpublished data).

Body size of the females was measured in three ways: abdominal length, thorax length, and fresh weight. The length measurements were made under a binocular scope at 30× magnification with an ocular micrometer of 100 divisions. All measurements were recorded to the nearest unit of the micrometer scale, each of which corresponded to 33·8 μm. For both length measurements, flies were laid on their right sides. Thorax length was taken as the distance between

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

the point where head and thorax meet and the tip of the scutellum. Abdominal length was approximated by the distance between the left haltere joint and the tip of the anal papilla.

Fresh weight was determined on a Cahn C-31 microbalance to the nearest full μg. The flies lost weight due to desiccation after they were taken out of the freezer at a rate of 0·2 μg/min. Only small batches of flies were therefore taken out of the freezer at a time, and a correction was applied to all weight measurements assuming a linear time-dependent weight loss during the time spent for weighing. The largest resulting correction was about 30 μg, which is less than about 1·5%.

A complication arose during the egg-laying phase of the experiment. Possibly related to a contamination of part of the foam tops used for plugging the vials, the substrate in many vials was infected by mould and other organisms. Because an effect of the contamination on egg laying appeared to be likely, a qualitative score was recorded to get an estimate of the disturbing effect of the infestation. The score ranged from 0 (no infestation) to 5 (substrate completely overgrown). An analysis of the effect of this score (see Results) lead to the exclusion of all vials with scores higher than 3 and the introduction of the score as a control variable into all ANOVAs in order to correct for its effect.

(iv) *Statistical analyses*

The effects of stocks, female types, temperature and infestation score on the different traits were analysed with fixed effects analyses of variance (ANOVAs) using the GLM procedure in the SAS package (release 6·04, 1990). Because of unequal cell numbers, type III sums of squares were used for statistical tests, and the means reported are least squares means that take the differences in cell sizes among compared groups into account. All variables were checked for a relationship between within-cell variances and means. Non-significant positive trends were found for weight and for the length measurements. A significant positive correlation between means and variances was found for the number of larvae (henceforth also referred to as 'fertility'): at 18 and 24 °C, respectively: $\rho = 0{\cdot}77$ ($P = 0{\cdot}016$) and $\rho = 0{\cdot}89$ ($P = 0{\cdot}001$). All variables were analysed both as original data and after trend-reducing transformations had been applied. There was no effect of the transformations on the conclusions for weight and lengths. Therefore only results based on untransformed data are presented. For fertility, the analyses shown are based on $\log_{10}$-transformed data.

(v) *Measurement of selection*

The strength of fertility selection on body size was measured using the fitness regression method of Lande & Arnold (1983). The contribution of fertility selection to the total gradient on body size can be estimated by regressing relative fertility on body size. The slope of this regression measures how much a female gains in fertility relative to other females if it is one size unit larger. Untransformed fertility data were used in the regressions (Lande & Arnold, 1983). The strength of fertility selection was measured within each category of females (i.e. for each karyotype within a stock at each temperature), and relative fertility was therefore defined as the fertility relative to the average within the category.

3. Results

(i) *Effect of mould*

A score ranging from 0 to 5 was assigned to each vial to describe the degree of infestation by mould or yeasts. The stocks were differently affected: mean scores for the different stocks at 24 °C were 2·1 (SR9), 2·5 (SR13) and 2·8 (SR16) (ANOVA: $P < 0{\cdot}0001$). At 18 °C, most vials had scores less than 2 (0·2 on average). The effect of the scores was different for different traits, but qualitatively similar in the different stocks and at the two temperatures. There was a significant effect on the time until a female began to lay eggs, on fertility, and, less often significant, also on weight. In general, higher scores correlated with later onset of egg laying (mean number of days for scores 0–5: 1·5; 2·3; 2·6; 3·9; 4·1; 4·1; $P < 0{\cdot}0001$ in an ANOVA testing the effect of score), suggesting that a mouldy environment had a negative effect on egg-laying propensity. Surprisingly, fertility was not simply reduced in vials with higher scores. Quite consistently in all stocks and at both temperatures, fertility increased up to score 3. At 24 °C, it dropped for higher scores (average numbers of larvae for scores 1–5: 72·9; 86·8; 99·3; 86·6; 47·6; $P < 0{\cdot}0001$ in an ANOVA testing the effect of score). Weight showed a similar and significant, although numerically less important, pattern. It was decided to exclude the vials with scores higher than 3, because the effect of the score was monotonic and similar for all categories of females in the remaining data set. The number of females in the different categories was thereby considerably reduced, almost by half in the worst case. The mean number of females analysed per category was 91·4 (range: 84–97) at 18 °C, and 67·7 (range: 51–88) at 24 °C.

(ii) *Genetic effects on traits*

All traits differed significantly among stocks and karyotypes at both temperatures (Table 1). The differences among karyotypes were further analysed in two contrasts: between the two homokaryotypes (SS–XX, 'additive genetic' effects) and between the heterokaryotypes and the homokaryotypes (SX–SS/2–XX/2, 'dominance' effects, referred to as 'heterosis' in Table 1). The latter were highly

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

Table 1. *Analyses of variance for body size and fertility at the two temperatures. The analyses included effects of the of the mould scores (and their interactions with other effects) which are not shown*

		18 °C		24 °C	
Source	D.F.	MS	*F*	MS	*F*
Weight					
Stocks	2	0·132	3·9*	0·392	11·6***
Karyotypes	2	0·579	17·1***	0·376	11·1***
Homozygotes	1	0·902	26·6***	0·115	3·4
Heterosis	1	0·301	8·8**	0·743	21·9***
Interaction (stocks × kar.)	4	0·275	8·1***	0·117	3·5**
Error		0·034 (D.F. = 798)		0·034 (D.F. = 585)	
R^2		0·162		0·253	
Abdominal length					
Stocks	2	0·075	10·8***	0·064	8·4***
Karyotypes	2	0·061	8·8***	0·099	13·0***
Homozygotes	1	0·113	16·3***	0·045	5·9*
Heterosis	1	0·012	1·7	0·190	24·9***
Interaction (stocks × kar.)	4	0·030	4·3**	0·024	3·1*
Error		0·007 (D.F. = 798)		0·008 (D.F. = 584)	
R^2		0·143		0·224	
Thorax length					
Stocks	2	0·079	64·1***	0·027	26·1***
Karyotypes	2	0·013	10·0***	0·010	9·6***
Homozygotes	1	0·023	18·5***	0·0	0·0
Heterosis	1	0·001	1·1	0·019	17·8***
Interaction (stocks × kar.)	4	0·015	11·9***	0·016	15·3***
Error		0·001 (D.F. = 798)		0·001 (D.F. = 583)	
R^2		0·315		0·244	
$\log_{10}$ (no. larvae)					
Stocks	2	0·273	4·7**	0·190	2·3
Karyotypes	2	0·166	2·9†	0·619	7·5***
Homozygotes	1	0·122	2·1	0·044	0·5
Heterosis	1	0·195	3·4†	0·970	11·7***
Interaction (stocks × kar.)	4	0·289	5·0**	0·137	1·6
Error		0·508 (D.F. = 799)		0·083 (D.F. = 591)	
R^2		0·315		0·244	

Significance: *** $P < 0·001$; ** $P < 0·01$; * $P < 0·05$ and 0·1 (†).

significant with few exceptions (abdominal length and thorax length at 18 °C). The homokaryotypes also differed in most cases (exceptions: fertility was not different at either temperature, and thorax length was not different at 24 °C).

Interpretation of the main effects in the table is complicated by the presence of highly significant interactions between stocks and karyotypes. The interactions indicate that the effect of the SR chromosome is not the same in all stocks. In Fig. 2, the means of the four traits can be compared among stocks and karyotypes. The lines connect means for the karyotypes, dotted lines for the separate stocks and a solid line for the average of the stocks. The figure shows that the effect of the SR chromosome differed among stocks mostly in a quantitative way, but that it was qualitatively similar, with a few exceptions. The heterokaryotypes (SX) had higher values than the average of the homokaryotypes in most stocks and most traits, especially at the higher temperature. In most cases this was equivalent to true heterosis or overdominance (the heterokaryotype superior to both homokaryotypes), but sometimes only dominance in the direction of the superior homokaryotype was involved. The other component of the karyotypic effect, the difference between the two homokaryotypes SS and XX, was also mostly consistent among stocks in a qualitative way, although more exceptions can be found here and the pattern is different for the traits and changes across temperatures: SS was always heavier (weight) than XX at 24 °C, but lighter at 18 °C. An analogous pattern was observed for abdominal length. Thorax length was longer in SS than in XX with only one exception

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

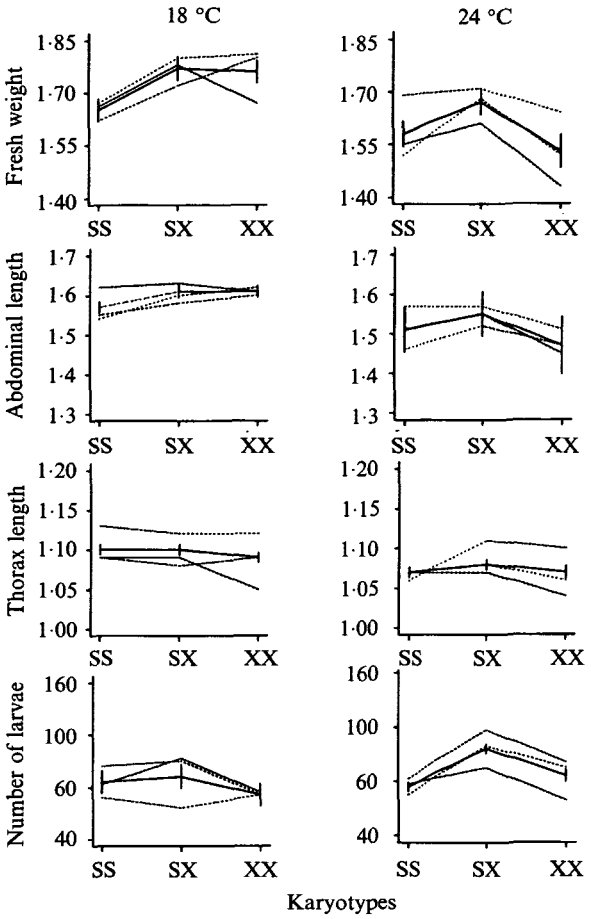

Fig. 2. Mean values for all traits at both temperatures for the three different karyotypes of each stock (notation on abscissa as in Fig. 1). The dotted lines connect the points for each stock separately (without error bars for clarity), and the solid lines connect the means over all three stocks (with bars representing two standard errors). Note the logarithmic scale on the *y*-axis of 'number of larvae' (lowest panels). These means are actually the means of the log-transformed data.

(SR13 at 24 °C). There was no consistent fertility difference between the SS and XX karyotypes.

To summarize: the effect of the SR chromosome on traits related to size appeared to be qualitatively consistent among stocks, but quantitatively different. This effect implied heterosis, or at least dominance in the direction of the larger homokaryotype. The effects on the difference between the two homokaryotypes depended on the traits and on the temperature. There was heterosis for fertility, but the homokaryotypes were not consistently different in their fertilities.

(iii) *Correlations among traits*

It is clear that the different traits are not independent of each other. Correlations were calculated separately for each category of females (9 values for each trait combination) and as partial within-cell correlations computed in a multivariate ANOVA. The latter were very similar to the means of the 9 separate within-category correlations which are given in Table 2 (above the diagonal for 18 °C and below the diagonal for 24 °C). At both temperatures, weight was strongly correlated with abdominal length and moderately correlated with thorax length and with fertility. Abdominal length was moderately correlated with thorax length and with fertility. Only thorax length showed almost no covariation with fertility. With the degrees of freedom applying for these correlations (between 60 and 95 in most cases), all correlations greater than 0·25 are significant at the 5% level and those greater than 0·4 are highly significant ($P < 0{\cdot}001$).

The pattern shown by the mean correlations seemed to hold for most genotypes, but a few exceptions qualified the picture. The ranges of the correlations in Table 2 indicate some heterogeneity among stocks. Closer inspection of the individual correlations revealed that the heterogeneity was confined to a few

Table 2. *Correlations among traits at 18 °C (above diagonal) and 24 °C (below diagonal). For each trait combination, 9 correlations were calculated (one for each female category, defined by stock and karyotype). The cells in the matrix show the means, the ranges (in square brackets) and the number of significant correlations (superscript)*

	Weight	Abdominal length	Thorax length	$\log_{10}$ (no. larvae)
Weight		$0{\cdot}85^{(9)}$ [0·77, 0·90]	$0{\cdot}39^{(9)}$ [−0·28, 0·61]	$0{\cdot}27^{(8)}$ [0·06, 0·37]
Abdominal length	$0{\cdot}88^{(9)}$ [0·77, 0·94]		$0{\cdot}32^{(9)}$ [−0·38, 0·49]	$0{\cdot}27^{(8)}$ [−0·03, 0·39]
Thorax length	$0{\cdot}35^{(7)}$ [0·16, 0·55]	$0{\cdot}30^{(4)}$ [0·05, 0·53]		$0{\cdot}12^{(2)}$ [−0·11, 0·29]
$\log_{10}$ (no. larvae)	$0{\cdot}21^{(5)}$ [−0·03, 0.37]	$0{\cdot}18^{(4)}$ [−0·02, 0·36]	$0{\cdot}04^{(0)}$ [−0·08, 0·12]	

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

Table 3. *Weight-adjusted fertility of homo- and heterokaryotypic females in the three stocks and overall. Shown are the estimated values of the statistical contrasts between the two homokaryotypic genotypes (SS–XX), and between the heterokaryotypes and the mean of the homokaryotypes (SX–SS/2–XX/2) (standard errors in parentheses). Fertility is the log_{10} of the number of larvae; a difference of 0·1 in the contrasts is therefore equivalent to a fertility difference of 25%. (Significance levels as in Table 1.)*

Contrast	SR9	SR13	SR16	All stocks
18 °C				
Homo-	−0·009	0·204***	0·087	0.282*
karyotypes	(0·042)	(0·053)	(0·059)	(0·091)
Hetero-	0·083*	0·078	0·012	0·173*
karyotypes	(0·036)	(0·058)	(0·058)	(0·086)
24 °C				
Homo-	−0·013	−0·120*	−0·086	−0·220**
karyotypes	(0·047)	(0·051)	(0·052)	(0·087)
Hetero-	0·060*	0·078	0·186***	0·324***
karyotypes	(0·040)	(0·047)	(0·047)	(0·078)

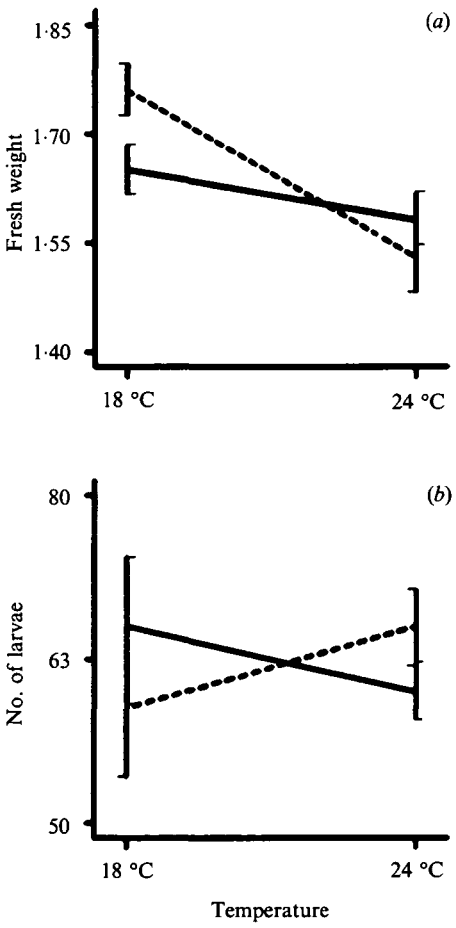

Fig. 3. Reaction norms for (*a*) weight and (*b*) fertility for the homokaryotypes (solid: SS, dashed: XX, heterokaryotypes left out for clarity; they were intermediate). For each karyotype, the means of the three stocks are plotted. Bars represent two standard errors.

categories of females, which caused the very low or negative values in the ranges for the different size–fertility correlations. The atypical correlations were not related to the infection scores and there is no obvious explanation for the deviations.

Because weight and abdominal length were so highly correlated, they were also very similar in all the analyses. Thorax length, on the other hand, was so weakly correlated with fertility that a detailed analysis of fertility selection on this trait was not undertaken. For these reasons, only analyses involving weight and fertility will be presented below.

(iv) *Weight-corrected fertility*

Because weight is positively correlated with fertility (Table 2) and the different categories of females have different weights on average (Table 1, Fig. 2), the analysis of genetic effects on fertility above is confounded with the weight differences among females of different categories. An analysis of covariance was used to estimate the genetic effects on fertility without the confounding effects of weight. Table 3 shows the results of an analysis of the effect of the SR chromosomes within the three stocks. As in Table 1, the differences among the three karyotypes (SS, SX and XX) were analysed as the contrast between the two homokaryotypes (SS–XX) and the contrast between hetero- and homokaryotypes (SX–SS/2–XX/2). The three stocks show qualitatively a similar pattern, although they differ quantitatively and the within-stock effects were not always significant. Overall, there were significant differences between the homokaryotypes at both temperatures and heterosis was also observed. The difference between homokaryotypes had a different sign at the two temperatures: at 18 °C, the SS females were more fecund than the XX females of the same size; at 24 °C, the

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

Table 4. *Environmental sensitivity of weight and fertility (analysed as log_{10} of the number of larvae). The table shows the ANOVA for the regression of the traits on temperature. The 'remainder' item represents the differences in slopes among stocks and karyotypes. The differences among karyotypes are further broken down into differences in slopes between the two homokaryotypes (SS–XX) and between the heterokaryotypes and the mean of the homokaryotypes (SX–SS/2–XX/2). (Significance levels as in Table 1.)*

Source	D.F.	Weight		Fertility	
Stock	2	0·177	5·2**	0·226	3·3*
Karyotypes (within stocks)	6	0·456	13·5***	0·436	6·3***
Temperature	1	2·098	61·9***	0·160	2·3
Remainder (differences of slopes)					
Among stocks	2	0·221	6·5**	0·255	3·7*
Among karyotypes	6	0·430	12·7***	0·465	6·8***
Homozygotes	3	0·764	22·5***	0·379	5·5***
Heterosis	3	0·205	6·1***	0·195	2·8*
Error		0·034	(D.F. = 1394)	0·069	(D.F. = 1401)
R^2		0·253		0·118	

reverse was observed. This contrasts with the weight-confounded analysis above : SS and XX females were not different in their fertilities at both temperatures. However, they had different weights: SS females were lighter than XX females at 18 °C, but heavier at 24 °C. It can be concluded that the SR chromosomes have temperature-dependent effects on both weight and fertility. Because fertility depends on weight, there are therefore two effects of the SR chromosome on fertility: a direct effect, and an indirect one mediated by weight. The two effects tend to cancel each other.

(v) *Environmental sensitivity*

The phenotypic response to temperature was analysed in terms of the slope of the reaction norms of weight and fertility. Since our experiment involved only two temperatures, this is equivalent to analysing the differences between temperatures. We preferred the reaction norm approach, because it directly addresses phenotypic plasticity.

Because the differences in weight among karyotypes depended on the temperature (see above), it could be anticipated that the karyotypes also differed in their phenotypic response to temperature. Fig. 3 shows the reaction norms of the homokaryotypes (SS and XX) for weight and fertility, and Table 4 gives the results of an ANOVA testing the effects of temperature on the same two traits. As expected, the overall response of weight to temperature was significant (Table 4). The reaction norms had a negative slope (Fig. 3*a*), implying that the flies were smaller at the higher temperature. There were significant differences in the sensitivity of weight to temperature and among stocks and karyotypes (i.e. differences among slopes, Table 4). As before, the differences among karyotypes were further analysed into homokaryotypic differences (SS–XX) and heterosis (SX–SS/2–XX/2). The difference between SS and XX females was significant in all stocks, and Fig. 3*a* shows that this was because XX females had steeper reaction norms than SS females, meaning that the latter were less sensitive to the temperature gradient. The heterokaryotypic females were in general intermediate (not shown in figure for clarity). Heterosis was not pronounced, although significant, which was due to one of the stocks, SR13.

The response of fertility to temperature was not like that of weight. There was no significant net effect of temperature on fertility (Table 4). Figure 3*b* shows that there were very different temperature responses of fertility for different karyotypes. This was also true when the stocks were compared separately. Although the differences in sensitivity among stocks and karyotypes were significant (Table 4), they could not be consistently related to either the stocks (i.e. the autosomal background) or the SR chromosomes.

(vi) *Selection on weight*

Fertility selection on weight was measured by the regression of relative fertility on weight. Table 5*a* gives estimates for the selection coefficients, and Table 5*b* shows the ANOVA tables for the comparison of different groups. With few exceptions, the selection coefficients were positive and significant, showing that there is fertility selection for larger body weight in almost all types of females and at both temperatures. The average selection coefficient was very similar at

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

Table 5. *Fertility selection on weight. Panel (a) gives estimates of the selection coefficients (one standard error in parentheses), which are the regression slopes of relative fertilities (numbers of larvae divided by the average number of larvae of females of the same type) on the weights. The marginal values (for karyotypes: row margins; for stocks: column margins) are the least square means computed with an analysis of covariance that allows for different slopes in each cell (corresponding to arithmetic means of the cell entries in completely balanced designs). Panel (b) gives the ANCOVA results on the differences among slopes. (Significances as in Table 1.)*

(*a*) Estimates of selection coefficients

	18 °C				24 °C			
	SR9	SR13	SR16	All stocks	SR9	SR13	SR16	All stocks
SS	0·89 (0·35)	0·27 (0·17)	0·62 (0·21)	0·60 (0·15)	0·83 (0·37)	0·72 (0·33)	−0·16 (0·29)	0·50 (0·19)
SX	0·71 (0·20)	0·67 (0·23)	0·18 (0·25)	0·52 (0·12)	0·46 (0·21)	0·51 (0·21)	−0·23 (0·30)	0·27 (0·17)
XX	0·58 (0·18)	0·66 (0·21)	0·41 (0·18)	0·56 (0·12)	1·20 (0·36)	0·93 (0·29)	0·57 (0·38)	0·92 (0·18)
All types	0·74 (0·15)	0·53 (0·12)	0·40 (0·12)	0·56 (0·08)	0·86 (0·18)	0·71 (0·17)	0·13 (0·19)	0·57 (0·10)

(*b*) ANCOVA results

		18 °C		24 °C	
Source	D.F.	MS	*F*	MS	*F*
Regression	1	7·68	55·2***	5·74	29·4***
Remainder (differences of slopes)					
Among stocks	2	0·21	1·6	0·83	4·2*
Among karyotypes	2	0·01	0·1	0·68	3·4*
Residual	4	0·17	1·2	0·06	0·3
Error (D.F.)		0·14 (785)		0·20 (564)	
R^2		0·101		0·174	

the two temperatures: slightly more than 0·5, meaning that if a female is 1 mg heavier than the average, it may lay eggs at an about 50% higher rate and therefore contribute more eggs to the next generation.

At 18 °C, there were no detectable differences among the selection coefficients of different stocks or karyotypes (Table 5*b*); in contrast, there were consistent differences among stocks and karyotypes at 24 °C. In particular, it appeared that fertility selection was much stronger in XX females than in SS females. This pattern could be seen in all three stocks.

As mentioned previously, the selection coefficients were on average very similar at the two temperatures. However, if one looks at the three karyotypes separately, temperature had an effect: for XX females, stronger selection was measured at the higher temperature (0·92 versus 0·55, $P = 0·09$), while for the other two karyotypes, the coefficients tended to be smaller at the higher temperature.

(vii) *Genetic correlations*

The data presented in the previous section show that fertility selection favours phenotypically larger females at both temperatures. Without counterbalancing forces, this would lead to the evolution of larger flies if some of the phenotypic variation had additive genetic causes and if genetically larger flies also had a higher fertility on average. It is therefore of interest to know whether the phenotypic relationship between fertility and size has a genetic parallel. The selection coefficients calculated above are based on within-category data of females. It is not known how much genetic variation for size and fertility there is within the different stocks. However, some insight into genetic correlations can be obtained by comparing the means of different categories of females, since the differences among them are genetic. Such an analysis, however, must be preliminary given the small number

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

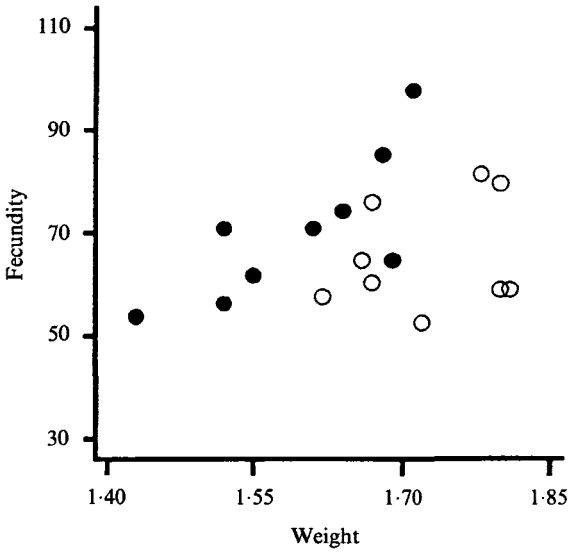

Fig. 4. Scatterplot of mean fertilities (number of larvae) versus mean weights of the nine categories of females at each temperature (three karyotypes in three stocks). Solid circles from 24 °C, open circles from 18 °C.

Table 6. *Genetic correlations between fertility and weight. Fertilities are mean numbers of larvae for the nine different categories of females calculated by back transformation from their respective least squares means of the transformed data. For each temperature, the first value is the usual Pearson correlation through the nine points. What follows are two decompositions of this overall correlation: three separate correlations within karyotypes (for each stock), and three correlations within stocks (for each karyotype). The means for each of the triples are also given for comparison with the overall correlation. With 3 degrees of freedom (applying for the mean correlations and the overall correlation), the 5% critical value is 0·878*

		Temperature	
		18 °C	24 °C
Overall correlation		0·247	0·771
Correlations			
By stock	SR9	0·964	0·974
	SR13	−0·414	0·863
	SR16	0·145	0·499
	Mean	0·231	0·779
By karyotype	SS	0·894	0·861
	SX	0·955	0·982
	XX	−0·998	0·898
	Mean	0·284	0·914

of different female types. In Fig. 4, the mean fertilities are plotted against the mean weights for the nine categories of females at each temperature. At 24 °C, a positive relationship appeared showing that genetically larger females (females from categories with a higher mean weight) also were more fecund at that temperature. At 18 °C, however, it appeared that genetically larger females did not lay more eggs on average, in contrast to the observation reported above that phenotypically larger females lay more eggs also at the lower temperature. This implies that at 18 °C, the positive correlation between weight and fertility is mostly due to environmental effects without evolutionary consequences. To have a more quantitative picture, correlation coefficients between mean fertilities and weights were calculated (Table 6). Statistical tests of the correlations cannot be made as usual, since the nine points (at each temperature) are not independent. The total correlation (first row) is composed of three separate correlations computable from three points (therefore with 1 D.F. each). The total correlation therefore has approximately 3 D.F. The three separate correlations can be computed in two ways: either within stock, giving the correlations due to the dosage of SR chromosomes, or within karyotype, giving the correlations due to genetic stock differences. All these correlations are listed in Table 6, along with their means. The correlations were high and close to significance at 24 °C (the 5% critical value is 0·88, which was surpassed in one case). The overall and mean correlations in the 18 °C data were low, and the separate correlations inconsistent. The data suggest that at 24 °C, the more fecund females would be the genetically larger ones, and, without counteracting forces, evolutionary change towards increased size would be expected. This is not the case at 18 °C. If the same pattern holds within stocks as among stocks, then the phenotypic selection shown above would be without evolutionary effect at 18 °C.

4. Discussion

We discuss our data with regard to three hypotheses that were outlined in the Introduction:

(1) There is fertility selection for larger body size; the strength of selection depends on the temperature and is weaker if the temperature is higher. This would explain the observation that smaller flies evolve at higher temperatures.

(2) Homokaryotypic sex-ratio females have a lower fertility than females with the standard gene arrangement on both sex chromosomes. This effect could be mediated by smaller body size. The fertility disadvantage of SS females should be temperature dependent, possibly explaining part of the geographical variation in SR frequencies that appears to be related to temperature.

(3) Stocks and/or karyotypes have different phenotypic responses to temprature.

(i) *Selection on body size*

The data clearly show that there is fertility selection for larger body size in *D. pseudoobscura* at both

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

temperatures and for a variety of genotypes. This agrees with Tantawy & Vetukhiv (1960), who found a positive relationship between wing length and lifetime fecundity, which includes longevity as a component.

There is also evidence that temperature had an effect on the strength of fertility selection on body size. However, the hypothesis that selection is weaker at the high temperature is clearly not supported. The effect of temperature depended on the karyotypes: while it was weak or inconsistent among stocks in SS and SX females, there was evidence that fertility selection could be much stronger at the higher temperature in XX females, contrary to the initial hypothesis. Since the standard arrangement is the most common in natural populations, one would expect that fertility selection is also stronger at higher temperatures in natural populations. It is therefore unlikely that differences in fertility selection across temperatures contribute to the evolution of smaller body size at higher temperatures.

This conclusion is further strengthened by the observation that a positive genetic correlation between fertility and weight, present at 24 °C, disappeared at 18 °C in our experiment. This result can clearly not easily be extrapolated to natural populations, since the correlations depend on the frequencies of the different genotypes which were all equal by design in the experiment. However, the consistency of the sub-correlations at 24 °C (Table 6) suggests that the pattern is not just coincidental. If it is general, then the hypothesis is even more strongly contradicted: even though there was phenotypic selection for larger body size at both temperatures, selection effective for evolution would occur *only* at the higher temperature.

An important implication of the present results is that the effect of temperature on fertility selection may depend on genetic factors such as the gene arrangement. The evolution of body size in relation to temperature could depend on the frequencies and the types of gene arrangements present in the population. To what extent did the evolution of smaller flies in Vetukhiv's cages depend on the monomorphism for the Arrowhead gene arrangement (Anderson, 1966)? Geographical observations (e.g. Stalker & Carson, 1947) and observations in laboratory populations even of different species (Cavicchi *et al.* 1985, 1989; Powell, 1974) suggest that the rule of genetically smaller flies at higher temperatures may be quite general. However, there are exceptions (Sokoloff, 1965), and the mechanisms underlying the rule appear not to be simple.

(ii) *The sex-ratio arrangement*

The experiment detected a clear effect of the SR arrangement on most traits. In contrast to previous work (Wallace, 1948; Curtsinger & Feldman, 1980; Beckenbach, 1983), individual versions of the arrangement were tested, each in a separate genetic background (the different stocks). Qualitatively similar effects of the SR arrangement were observed on most traits and in all stocks, although there were quantitative differences among stocks indicating that an important proportion of loci within the inversions segregate for alleles that have different effects on the traits, or that the pleiotropic effects of alleles at the SR loci depend on the genetic background.

In agreement with the findings of the previous authors we found a superiority of the heterokaryotypic females relative to both homokaryotypes for fertility and body size. The question arises whether the effect of the karyotype on fertility is completely through the effect on size, because larger females can lay more eggs. The answer is no; even if the fertility data were corrected for the effect of weight, or abdominal length, there was still a heterosis effect (Table 3).

The difference between the homokaryotypes was more complicated. According to the hypothesis, the SS females should have a lower fertility than the XX females. This was not consistently the case at either temperature. In contrast, Wallace (1948) found a slight disadvantage of SS females at 25 °C and a greater disadvantage at 16·5 °C. However, he did not control for body size.

Our study shows that the effect of SR on fertility is complicated because it was twofold: direct, and indirect through body size. The direct effect was negative at 24 °C (negative contrasts between the homokaryotypes, Table 3) agreeing with Wallace's results. The way the contrasts changed between temperatures was different, however. Wallace (1948) found that the sex-ratio females performed even worse at the lower temperature, whereas we found that they had an advantage over the standard females. The SR arrangement also has a temperature dependent effect on weight. SS females were significantly heavier than XX females at 24 °C, but lighter at 18 °C. This pattern tends to cancel the direct effects on fertility, and the net fertilities were therefore not significantly different (Fig. 2). It is not surprising that a complicated interaction such as this can result in diverging observations when the environmental conditions are not exactly the same.

How do these results relate to the population dynamics of the sex-ratio chromosome? It has been postulated that SS females should suffer a fertility disadvantage relative to XX females. Although the tendency was in the expected direction in some circumstances, it was very weak and not important for yet another reason: the superiority of heterokaryotypic females. Because the observed frequencies of the SR chromosome are low in natural populations, SS females are quite rare and their reduced fertility would contribute very little to the selection against the arrangement even if the effect were stronger. The SX females, which are much more common, contribute towards stabilizing the frequency of SR by their superiority. Therefore, it can be concluded that it is

https://doi.org/10.1017/S0016672300031578 Published online by Cambridge University Press

very unlikely that the frequencies of the SR arrangement are low in natural populations because of fertility selection.

Another aspect of the hypothesis was motivated by geographical surveys of SR frequencies in natural populations suggesting that the fitness of the arrangement is reduced at cooler localities (Sturtevant & Dobzhansky, 1936; Wallace, 1948). In contrast to this, the arrangement tends to be eliminated more quickly in population cages kept at warmer temperatures. Our results on weight-adjusted fertilities, indicating reduced relative fertility for SS at 24 °C and increased fertility at 18 °C, suggest that the SS females suffer more under warm conditions in the laboratory and may partly explain the observations made on population cages. The discrepancy with the geographical pattern could be due to other fitness components, as Wallace (1948) suggested. Our results on weight show that even without invoking other fitness components, the apparent contradiction could be resolved: since SS females tended to be smaller than XX females at lower temperatures, their 'weight-adjusted' fitness advantage could easily be offset. What counts, after all, is net fertility.

(iii) *Environmental sensitivity*

The present study confirms previous results by demonstrating the influence of temperature on weight and on fertility, and of the relationship between weight and fertility. It was also found that the different genotypes differed in the slopes of their reaction norms (Table 4). Such data are necessary for the study of the evolution of the reaction norms of weight.

Besides the genetics of reaction norms, a model for the evolution of reaction norms also requires information about the selection processes that affect the plastic trait, and whether fitness changes across environments. We measured a component of this function in the form of the regression of relative fertility on weight. There was evidence that the relationship between the fitness component and the trait may change across environments, as envisaged by most theoretical models. More importantly, however, there was also evidence that the relationship between fitness and the trait, and the way this relationship changes across environments, depends on the genotypes. The selection gradient for weight was much steeper at the higher temperature for XX females in all three stocks, which was not the case for females with SR chromosomes. This is a result that is not usually considered in models of phenotypic evolution, or of phenotypic plasticity in particular. It is generally assumed that a general fitness-landscape can be superimposed over the reaction norms of the different genotypes, which would in principle allow calculation of the evolutionary dynamics of reaction norms (de Jong, 1989, 1990), or of the traits in the separate environments (Via & Lande, 1985). A complete model of phenotypic plasticity evolution needs to include the possibility that fitness depends not only on the phenotypes and the environment, but on the genotypes as well. A possible explanation for this dependency may be that the traits under study are genetically correlated with some other traits that are also important for fitness. It should be an interesting task for future research to work out the mechanisms for the interaction between selection, environment, phenotypes and genotypes.

We wish to acknowledge the contributions of H. Alavi who organized the laboratory, Y. Lay and D. Lay who assisted in breeding the flies and preparing the food, J. Stewart who provided access to his microbalance, and S. Henrich and anonymous reviewers who made helpful suggestions on the manuscript. M.G. was financially supported by the Swiss National Science Foundation.

References

Anderson, W. W. (1966). Genetic divergence in M. Vetukhiv's experimental populations of *Drosophila pseudoobscura*. 3. Divergence in body size. *Genetical Research* **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.

Beckenbach, A. (1983). Fitness analysis of the 'sex-ratio' polymorphism in experimental populations of *Drosophila pseudoobscura*. *American Naturalist* **121**, 630–648.

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

Cavicchi, S., Guerra, D., Natali, V. & Pezzoli, C. (1989). Temperature-related divergence in experimental populations of *Drosophila melanogaster*. II. Correlation between fitness and body dimensions. *Journal of Evolutionary Biology* **2**, 235–251.

Curtsinger, J. W. & Feldman, M. W. (1980). Experimental and theoretical analysis of the 'sex-ratio' polymorphism in *Drosophila pseudoobscura*. *Genetics* **94**, 445–466.

Gebhardt, M. D. & Stearns, S. C. (1988). Reaction norms for developmental time and weight at eclosion in *Drosophila mercatorum*. *Journal of Evolutionary Biology* **1**, 335–354.

Gupta, A. P. & Lewontin, R. C. (1982). A study of reaction norms in natural populations of *Drosophila pseudoobscura*. *Evolution* **36**, 934–948.

Jong, G. de (1989). Phenotypically plastic characters in isolated populations. In *Evolutionary Biology of Transient Unstable Populations* (ed. A. Fontdevila), pp. 3–18. Heidelberg: Springer.

Jong, G. de (1990). Quantitative genetics of reaction norms. *Journal of Evolutionary Biology* **3**, 447–468.

Lande, R. & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226.

Mayr, E. (1963). *Animal Species and Evolution*. Cambridge, Massachusetts: Belknap Press.

Misra, R. K. & Reeve, E. C. R. (1964). Clines in body dimensions in populations of *Drosophila subobscura*. *Genetical Research* **5**, 240–256.

Policansky, D. & Ellison, J. (1970). Sex-ratio in *Drosophila*

pseudoobscura: spermiogenic failure. *Science* **169**, 888–889.
Powell, J. R. (1974). Temperature related genetic divergence in *Drosophila* body size. *Journal of Heredity* **65**, 257–258.
Prevosti, A. (1955). Geographical variability in quantitative traits in populations of *Drosophila subobscura*. *Cold Spring Harbor Symposium on Quantitative Biology* **20**, 294–299.
Roff, D. A. (1981). On being the right size. *American Naturalist* **118**, 405–422.
SAS Institute Inc. (1990). *SAS User's Guide: Statistics*, version 6.04 edition. Cary, North Carolina.
Schmidt-Nielson, K. (1984). *Scaling: Why is Animal Size so Important?* Cambridge, England: Cambridge University Press.
Sokoloff, A. (1965). Geographic variation of quantitative characters in populations of *Drosophila pseudoobscura*. *Evolution* **19**, 300–310.
Stalker, H. & Carson, H. (1947). Morphological variation in natural populations of *Drosophila robusta* Sturtevant. *Evolution* **1**, 237–248.
Stearns, S. C. & Koella, J. C. (1986). The evolution of phenotypic plasticity in life-history traits: predictions of reaction norms for age and size at maturity. *Evolution* **40**, 893–913.
Sturtevant, A. H. & Dobzhansky, Th. (1936). Geographical distribution and cytology of 'sex-ratio' in *Drosophila pseudoobscura* and related species. *Genetics* **21**, 473–490.
Tantawy, A. O. & Vetukhiv, M. O. (1960). Effects of size on fecundity, longevity and viability in populations of *Drosophila pseudoobscura*. *American Naturalist* **94**, 395–403.
Via, S. & Lande, R. (1985). Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* **39**, 505–522.
Wallace, B. (1948). Studies on 'sex-ratio' in *Drosophila pseudoobscura*. I. Selection and 'sex-ratio'. *Evolution* **2**, 189–217.