An empirical test of models for the evolution of sexual size dimorphism as a correlated response to selection on body size

Jeff P Reeve

A Thesis
in
The Department
of
Biology

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

March, 1995

ISBN 0-612-01328-6

Canada

ABSTRACT

An empirical test of models for the evolution of sexual size dimorphism as a correlated response to selection on body size

Jeff P Reeve

Artificial selection for increasing and decreasing body size was applied to *Drosophila melanogaster* Thorax width was used as an estimator of body size. Two replicates of seven selection regimes were maintained, with selection being for. 1) increased size in both sexes, 2) decreased size in both sexes; 3) increased size in females; 4) decreased size in females, 5) increased size in males, 6) decreased size in males, and 7) a control with flies randomly selected each generation Selection was maintained for 21 generations The correlated response of sexual size dimorphism in each selection regime was compared to the response predicted by existing quantitative genetics models. These models have previously been tested only by one retrospective study on human metric traits, and one examination of the fossil record for a dimorphic trait in extinct peccaries Four models were tested, each of which differed only in assumptions about input parameters Body size responded well to selection, and the models were able to accurately predict these responses The correlated response of sexual size dimorphism, however, was weaker than expected by any of the models. This discrepancy could be due either to insufficiently detailed information for estimating the input parameters, or to some genetic factor not accounted for by the models A significant non-linear allometry for

sexual size dimorphism was found, characterized by faster response in males to selection for increased size, and faster response in females to selection for decreased size. This relationship produced a trend for decreased dimorphism in response to selection in either direction.

ACKNOWLEDGEMENTS

Firstly, I would like to thank Dr Daphne Fairbairn for all of her help and patience throughout the project. The comments of committee members Dr. E Maly and Dr. J Grant have helped to improve this thesis Dr. D. Roff's suggestions for the data analyses were very useful Thanks also to everyone in the lab for providing an enjoyable environment in which to work In particular, I would like to thank my parents and Gina Macintyre for their constant support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION.

Sexual size dimorphism (SSD) is common throughout the animal kingdom, and the observed size difference can favour either sex. Males are larger in most mammals (Ralls, 1977, Alexander et al , 1979) and birds (Selander, 1972, Payne, 1984), while females are larger in most invertebrates (Ghiselin, 1974, Honek, 1993), fish (Andersson, 1994), amphibians (Shine, 1979), and reptiles (Berry and Shine, 1980, Shine, 1994). The magnitude of dimorphism and even its direction can differ among closely related species, and to a lesser extent among populations of the same species (e.g Walker and Corbet, 1975, Fairbairn and Preziosi, 1994) There are a number of theories as to why SSD evolves These include sexual selection (where mating success is a function of body size), fecundity selection (large females produce more and/or larger offspring), and ecological causes (such as niche dimorphism, where different habits of life favour different body sizes in the two sexes) Each of these originated with Darwin (1874), and has been the subject of much research, especially over the last 25 years (e.g. Selander, 1972, Alexander et al , 1979, Bradbury, 1987; Andersson, 1994 [sexual selection], Shine, 1988; Wiklund and Karlsson, 1988 [fecundity selection], Ralls, 1976; Slatkin, 1984 [ecological causes]). These theories share the assumptions that there are different net selective forces acting on the two sexes, and that the dimorphism maximizes fitness and will be retained at equilibrium The observed magnitude of SSD in any species undergoing differential selection between the sexes is not necessarily that which will be seen at equilibrium, since genetic correlations may cause transient but long lasting changes in the non-selected sex (see below) Sexual selection is usually thought to be the main force responsible for male-

biased SSD, while fecundity selection is the usual explanation for SSD in taxa in which females are larger (Andersson, 1994) It is also possible that in some species SSD may exist even though there is no differential selection acting on the two sexes This phenomenon might occur through phylogenetic inertia (current dimorphism being largely due to the dimorphism found in ancestors [Gould and Lewontin, 1979]) or variance dimorphism (differential response to equal intensities of selection [Cheverud et al 1985, Leutenegger and Cheverud, 1985]) In such cases, adaptive explanations for present dimorphisms may be unwarranted

Whatever the reason for SSD, the evolution of the underlying genetic mechanism is not well understood The two sexes share the same set of autosomal genes, which usually make up the majority of the genome In animals with heteromorphic sex chromosomes, one (for instance the Y in mammals) is nearly genetically inert as far as morphological traits are concerned, and some form of dosage compensation usually equalizes the effects of having either one or two X chromosomes (Bull, 1983) Because the sexes have virtually all the same genes, most traits show a high genetic correlation between the sexes (Lande, 1980), and body size differences are due largely to differential expression of the same genes Body size in the two sexes can be conceptualized either as the expression of the same trait in two different environments, or as two different but correlated traits The standard quantitative genetics formulae used to predict the correlated response of one sex to selection on the other are based on the latter Direct response (R) to selection is given by

$$R=h^2S \qquad (1)$$

where h^2 is the narrow-sense heritability [(additive genetic variance)/(total phenotypic variance)], and S is the selection differential (the difference between the mean of the selected population and the mean of the original population) The selection differential is often standardized by dividing by the phenotypic standard deviation σ_p to give the selection intensity i, so that

$$i = S/\sigma_p \qquad (2)$$

The total response of each sex is due to the direct response to selection on that sex and the correlated response to selection on the other sex Therefore, where the subscripts m and f refer to male and female respectively, and r is genetic correlation between the sexes,

$$R_m = \tfrac{1}{2}(h^2_m\sigma_{pm}i_m + h_m h_f r i_f \sigma_{pm}) \qquad (3a)$$

$$R_f = \tfrac{1}{2}(h^2_f\sigma_{pf}i_f + h_m h_f r i_m \sigma_{pf}) \qquad (3b)$$

(Leutenegger and Cheverud, 1985, based on Falconer, 1981 [1989])

If the genetic correlation is high, selection for a change in body size in one sex will result in a correlated response in the other, and dimorphism will be difficult to alter. This idea was first expressed qualitatively by Fisher (1958), and later quantified by Lande (1980) Lande took the standard quantitative genetics formulae for response and correlated response to selection and converted them into a set of recursion equations that could be used to follow the trajectories of the two sexes under a given set of parameter values His model also incorporated the effect of Gaussian stabilizing selection for some optimum value for natural selection in each sex, which provides a means of producing a response plateau as the extrinsic selective force (e.g sexual selection) moves the organisms away

from their original optimal size (The standard response equation (1) has no built-in constraints on response, so cannot make any predictions about response to long-term selection.) In artificial selection experiments, a plateau is often reached after a relatively small number of generations (Falconer, 1989) When this plateau is found in conjunction with high remaining levels of genetic variance, natural selection in the opposite direction is the most likely explanation for the lack of further response When the plateau is associated with a lack of variance, the lack of response most probably reflects fixation or loss of alleles, caused by selection and relatively small effective population sizes Lande's model assumes that selection is weak enough and/or population sizes large enough that loss of genetic variance is balanced by new alleles created through mutation The conclusion he drew was that SSD evolves in two phases In the first, rapid phase, the evolutionary trajectories of the trait means in the two sexes are "almost parallel" (Lande, 1987), to a point where the net selective forces in each sex are equal in magnitude but opposite in sign In the second phase, the two sexes slowly diverge, with the selected sex moving in the selected direction, and the non-selected sex returning to its initial trait value This slow divergence is accomplished by a breakdown of the genetic correlation between the sexes through the accumulation of sex-limited genes (Fisher,1958), and leads to the prediction that highly dimorphic traits will have low genetic correlations (Although Lande (1980) agrees that the correlation will break down, his model assumes constant genetic correlations (and covariances), and the sexes diverge because of diminishing selection pressure on the opposite sex after phase I) In this scenario, the non-selected sex is temporarily drawn away from its optimum (natural selection) value, while the selected

sex equilibrates at a point dictated by a balance between the forces of natural selection in one direction and sexual selection in the other The relative lengths of time spent in phases I and II are proportional to 1/(1+r) and 1/(1-r) respectively This conclusion is based on the assumption of equal heritabilities and phenotypic variances in the two sexes, high genetic correlations, and different selection intensities.

In Lande's model, the measurement scale is not specified, but is assumed to be in units such that variance is relatively constant with respect to the trait mean. For the results shown in his 1980 paper, he assumes equal heritabilities and phenotypic variances in males and females. Selecting on both sexes simultaneously in the same direction with the same intensity will produce a response of 'x' units in males and 'x' units in females. If the measurement scale is arithmetic, the dimorphism measured as a difference will not change, while the dimorphism measured as a ratio (size of larger sex divided by the size of the smaller sex) increases with downward selection, and decreases with upward selection If the scale is logarithmic, size ratio remains constant while size difference decreases with downward selection, and increases with upward selection. The magnitude of the change in SSD therefore depends on the relative size of the original dimorphism and body size, and on the measurement scale used.

For the remainder of this paper, SSD_r will refer to the ratio of the larger sex to the smaller, and SSD_d to the size of the larger sex minus the smaller. Where dimorphism in general is being discussed, SSD will be used. It is not clear which measure is better, and both SSD_d and SSD_r are frequently used. Each has its disadvantages; distributions of ratios tend to be leptokurtic and skewed (Atchley et al 1976), while differences are not

intuitively appealing when comparing dimorphisms of groups which differ greatly in body size

Lande's model is the most basic available for the evolution of SSD, with more recent modifications, such as Slatkin's (1984), differing only in the greater detail with which selection differentials are calculated. Although Lande's paper is generally interpreted as predicting that SSD will be slow to evolve (e.g. Rogers and Mukherjee, 1992; Wright, 1993), this is not a fundamental prediction of the model itself, but rather a consequence of the assumptions made about parameter values. Cheverud et al. (1985), again using the standard formulae (eqns. 3a and 3b, i.e. the Lande model minus the constraints due to natural selection), showed algebraically that, given assumptions of equal selection intensities and heritabilities in each sex, but unequal phenotypic variances, SSD_d could evolve rapidly, and that its magnitude is actually enhanced by high genetic correlations. This means that the response of SSD could critically depend on differences between the sexes with regard to the relative and total contributions of additive genes to the phenotypic variance. Cheverud et al. conclude that SSD_d will evolve most rapidly when differential selection pressures and variance dimorphism favouring the same direction of response are found.

For the purposes of this thesis, the Lande model will be understood to refer to his 1980 model with assumptions of equal variances and heritabilities in each sex. What will be referred to as Cheverud's model differs from Lande's model only by the inclusion of sex differences in heritability and variance. The equations in Cheverud et al.'s 1985 paper do not include a term for the stabilizing effect of natural selection. This is because they

were interested in the algebraic relationship between male and female response based on a single generation of selection, where the effects of natural selection could be assumed to be negligible. In this thesis, the effects of natural selection are included in both models, since selection is maintained over many generations.

Sex differences in heritability and genetic variance are difficult to detect, due to the large standard errors usually associated with heritability estimates. Heritability of body size generally does not differ significantly between males and females (e.g. Cowley and Atchley, 1988; Rogers and Mukherjee, 1992; Maria et al., 1993). Estimates of phenotypic variance, while easy to obtain, seem to vary greatly among populations and traits, with no general pattern emerging as to sex differences. Despite these difficulties, it should be kept in mind that 1) high genetic correlations do not necessarily mean slow SSD evolution, and 2) similar phenotypic variances in the two sexes do not necessarily equate to similar heritabilities. Further, while no other mathematical model exists for the evolution of SSD, it has been suggested that evolution may proceed by changes in one sex only, through the evolution of sex-limited genes (Turner, 1978; Rice, 1984; Meagher, 1992) especially in traits which are already slightly dimorphic and subject to developmental canalization (Waddington, 1962; Wright, 1993).

The purpose of this thesis is to test empirically the Lande and Cheverud models by artificially selecting on body size in *Drosophila melanogaster*. Selection is applied to each sex separately, as well as both simultaneously. Using parameters obtained from a half-sib design heritability experiment and the selection experiment, the predictions of the models are compared to the experimental results to see if they are consistent or can be

made consistent by reasonable changes in model parameters Previous tests of the Lande model have been limited to analyses of the fossil record (e g Kiltie,1985, Wright, 1993), and a re-analysis of some old and incomplete data on human populations (Rogers and Mukherjee, 1992)

Most experimental work relevant to the evolution of dimorphism has been done on *Drosophila melanogaster*, where females are about 10% longer than males for most linear traits Many previous papers have reported the results of selection on one sex (e g Frankham, 1968, Alicchio and Palenzona, 1971, Wilkinson, 1993) or on both sexes simultaneously (e g Zeleny, 1921; Robertson and Reeve, 1952, Alicchio and Palenzona, 1971; Higuet, 1991, Partridge and Fowler, 1993), but most have featured one or more of the following characteristics which reduce their value for the purposes of this study

1) Small numbers of mating pairs in the selection lines It was not appreciated during the early days of selection experiments that inbreeding could have serious effects on the interpretation of results Experiments with as few as one mating pair per generation were common, 2) Qualitative reporting of results. Even basic statistical analyses are rare in genetics papers written before the 1970's, 3) Insufficient data with which to test for trends in SSD. Change in SSD is seldom the primary research interest of the authors, so the relevant data are rarely shown

In the present study, I attempt to overcome these deficiencies by using a larger number of mating pairs (15), and selecting over 21 generations. The data are collected and analyzed to examine how well the Lande and Cheverud models are able to predict the changes in body size, SSD_d, and SSD_r, brought about by selection on body size

MATERIALS AND METHODS.

To ensure sufficient genetic variance for selection to act upon, the base stock used for these experiments was derived from crosses among four laboratory strains of *Drosophila melanogaster* (Urbana-S, Swedish-C, Lausanne-S, and Oregon-R) After two generations of reciprocal crosses between these strains, flies had an expected genetic constitution of 25% from each original strain These flies were then randomly bred for four generations prior to commencement of the selection and heritability experiments.

Flies were raised in 3.1 x 10.0 cm plastic vials, and housed in an incubator kept at 25°C with a 12-hour light/dark cycle. An open pan of water was kept in the incubator to maintain humidity Flies were grown in commercial *Drosophila* medium, to which a weighed amount of dried yeast (17 mg per vial) was added Measurements were made on ether-anaesthetized flies, using digitizing software on a computer attached to a 40X microscope. The trait measured was the distance between the inner edges of the insertion points of the two posterior sternopleural bristles on the ventral surface of the thorax. This distance was used as an indicator of body size in preference to the more usual measures of thorax length or width, because of the ease with which the two reference points could be aligned in the same plane by resting the flies on their backs This trait, which spans approximately 80% of the maximum width of the thorax, will be referred to as thorax width for the remainder of the thesis

Heritability experiment

A half-sib design mating experiment was done in order to get initial estimates for

the heritability of thorax width in each sex as well as the genetic correlation between the sexes. These estimates, which are obtained through a nested analysis of variance (Becker, 1984), are later used as parameter values for testing the Lande and Cheverud models

Virgin flies were collected by removing adults from vials in the morning, and collecting those flies which eclosed over the next 6 to 7 hours (females do not mate for the first 8 to 12 hours after eclosion [Ashburner and Thompson, 1978]) Each of 85 virgin males (sires) was placed in a separate vial with 4 virgin females (dams) for a period of 24 hours. The females were then separated and each was placed in its own vial for 24 hours, so that the eggs laid in each vial over this period developed into adults which were full-sibs. Flies raised under the conditions used in this experiment start to eclose about 9 days after eggs are laid On day 13, when virtually all of the flies had eclosed, three male and three female offspring were measured from three of the four families sired by each male The offspring sired by a male were included in the analysis only if at least three of his dams produced enough offspring for measuring. Ten of the sire's families did not have sufficient numbers, leaving 75 half-sib families for the analysis of variance

Selection experiment

Artificial selection on thorax width was maintained for 21 generations in order to examine the change in sexual size dimorphism as a correlated response to direct selection on thorax width Two replicate sets of seven selection lines were kept, for a total of fourteen lines The selection regimes were as follows 1) both sexes increasing (up) (BU), 2) both sexes decreasing (down) (BD), 3) females up (FU), 4) females down (FD), 5)

males up (MU); 6) males down (MD); and 7) a control (C). In the selected sex(es), 60 flies were measured and the 15 most extreme phenotypes in the appropriate direction selected In the unselected sexes, 15 flies were randomly selected and measured. For each line, the 15 pairs of flies were placed together in a vial for 24 hours, then transferred to another vial for 24 hours, and then again a third time. No attempt was made to control egg density, other than by collecting for a fixed period of time. After transfer from the third vial, the remaining flies were kept on a shelf in the laboratory (at room temperature), as a reserve in case of incubator failure. When offspring from these matings began to eclose after about 9 days, virgin flies were collected and separated by sex, under ether anaesthesia. The virgin flies from each vial (3 vials per selection line) were kept separate from those from other vials until they were measured. After four days, or longer if necessary to collect sufficient numbers, flies were measured and selected. Where possible, roughly equal numbers were measured and selected from each of two of the three vials per line (e.g. 30 measured and 8 selected from the first vial, and 30 measured and 7 selected from the second). The entire cycle from one measuring day to the next was usually 14 days. Some flies were still eclosing after 4 days of collecting, which caused the distribution of flies measured from the second day's mating vial to be truncated to a slightly greater extent at the slow developing extreme than was that of the flies from day one's vials This extended period of eclosion (compared to that seen in the half-sib heritability experiment, where all the offspring usually eclosed within 48 hours of the first eclosion) was presumably due to increased competition from the much larger number of offspring produced by 15 mating pairs versus 1 pair in the heritability experiment. Vials

from the first two days of egg laying were used whenever possible. When the number of offspring produced by two vials was not great enough to make up the required 60, all three vials were used. The proportion of flies measured and selected from each vial was the same for both sexes, so that measures of dimorphism were not influenced by unequal sampling between the sexes.

Death from anaesthesia was very infrequent. Dead flies were never mistakenly measured since ether-induced death causes the wings to project upwards at a 45-degree angle, making them impossible to measure. Ether in low doses does not adversely effect the flies, although prolonged exposure can reduce fertility in males (Ashburner and Thompson, 1978). It was rare for more than 5 of the original 30 flies to die (or escape during transfers) before transfer out of the third day's vial.

The lines were originally set up as 2 replicates, with the second set started one week after the first. In each replicate the order of line initiation was random. No attempt was made to maintain the sequence of lines. The determining factor was always which line had sufficient numbers of offspring for measuring After generation 0, one line was measured per day on average. By the end of the selection experiment, the down and control lines were developing about one day faster than the up lines. A decline in size in the control lines, as seen in the present study, is often found in selection experiments (e g Partridge and Fowler, 1993), and is probably caused by natural selection favouring smaller flies due to the restriction of development time to 14 days. Since large flies generally take longer to develop than small flies (Ashburner and Thompson, 1978), a reduction in average body size over time under these conditions is not unexpected.

RESULTS.

1) Half-sib estimates of heritability and genetic correlation

Heritabilities were estimated as four times the intraclass correlation of half-sibs (Becker, 1984). Full-sib estimates were not used, since these contain dominance, maternal, and common-environment effects (Falconer, 1989). The estimate of heritability of thorax width is greater in females than in males (Table 1) Both are significantly different from zero, but not from each other. The genetic correlation (r) between the sexes was calculated from an analysis of covariance, again from the nested design described by Becker (1984). It is high and significantly different from zero (Table 1).

2) Response of thorax width to selection

Female and male generation means for both replicates of the seven selection lines are given in Appendix Tables A1-A4. Figures 1 and 2 show the effects, on females and males respectively, of thorax width selection. Each data point is the average of the two replicates for that selection type. Although response was greater in the down than in the up lines, the control lines also decreased in size. The responses up and down, after adjusting by subtracting the mean of the two control lines, are approximately symmetrical for selection on both sexes simultaneously, and for selection on males (Figures 3 and 4). All selection lines had clearly responded in the expected direction after 21 generations of selection, indicating that the trait measured as an estimate of thorax length had a significant heritability. An analysis of covariance was done in order to test for significant changes in thorax width, and to compare the response in each sex for all selection lines. Thorax size was the dependent variable, generation the independent variable, and sex the

Table 1. Anova results for half-sib design heritability estimates.

	Source	df	SS	MS	F
Female	Sires	74	3.20×10^{2}	4.32×10^{1}	2.03***
offspring	Dams	150	3.20×10^{2}	2.13×10^{1}	2.70***
	Progeny	450	3.54×10^{2}	7.88×10^{0}	
Male	Sires	74	2.56×10^{2}	3.45×10^{-1}	1.59***
offspring	Dams	150	3.26×10^{2}	2.17×10^{-1}	3.07***
	Progeny	450	3.18×10^{2}	7.06×10^{0}	

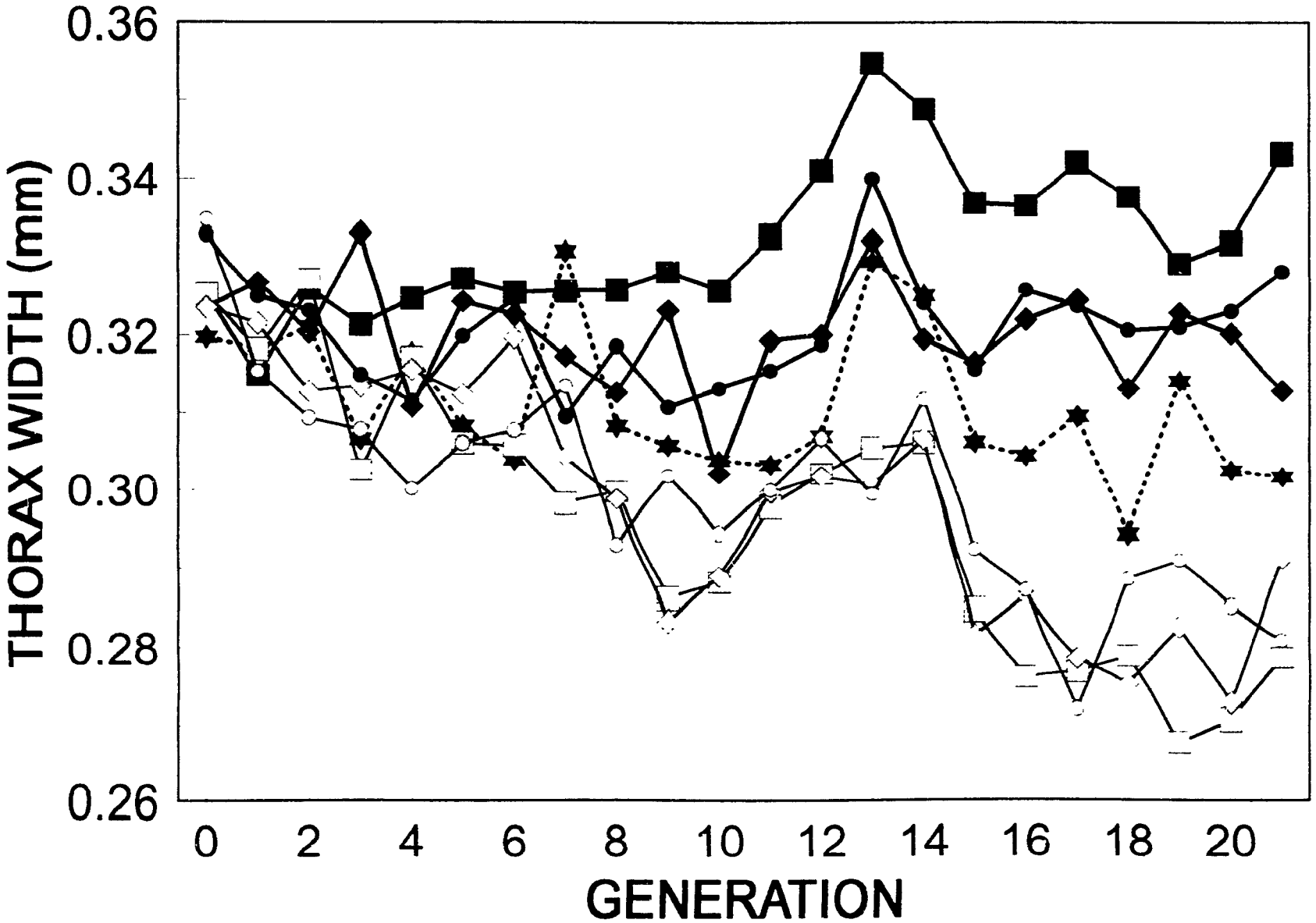
*** $p < 0.001$

h^2 females = 0.66 SE=0.22

h^2 males = 0.43 SE=0.20

Additive genetic correlation between the sexes (r_A) = 0.93 SE = 0.014

Figure 1. Effect on *Drosophila melanogaster* females of selection for thorax width. Control (★), both sexes selected up (■), both sexes selected down (▩), females selected up (◆), females selected down (◈), males selected up (●), males selected down (◉) Each data point is the average of the two replicates per selection regime.


0.36
0.34
0.32
0.30
0.28
0.26
THORAX WIDTH (mm)
0
2
4
6
8
10
12
14
16
18
20
GENERATION

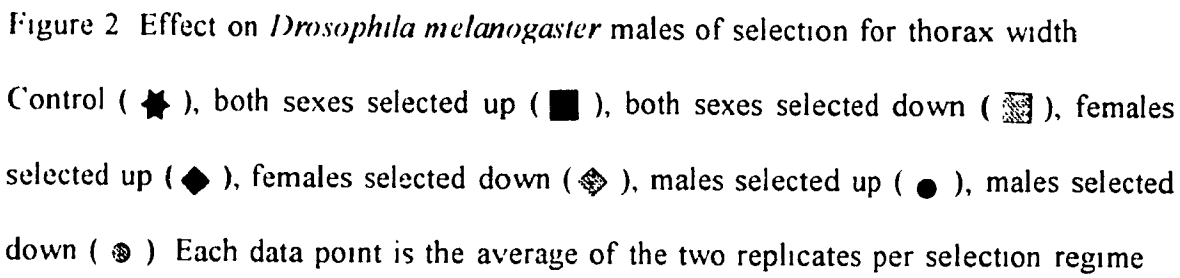


Figure 2 Effect on *Drosophila melanogaster* males of selection for thorax width Control ( ), both sexes selected up ( ), both sexes selected down ( ), females selected up ( ), females selected down ( ), males selected up ( ), males selected down ( ) Each data point is the average of the two replicates per selection regime

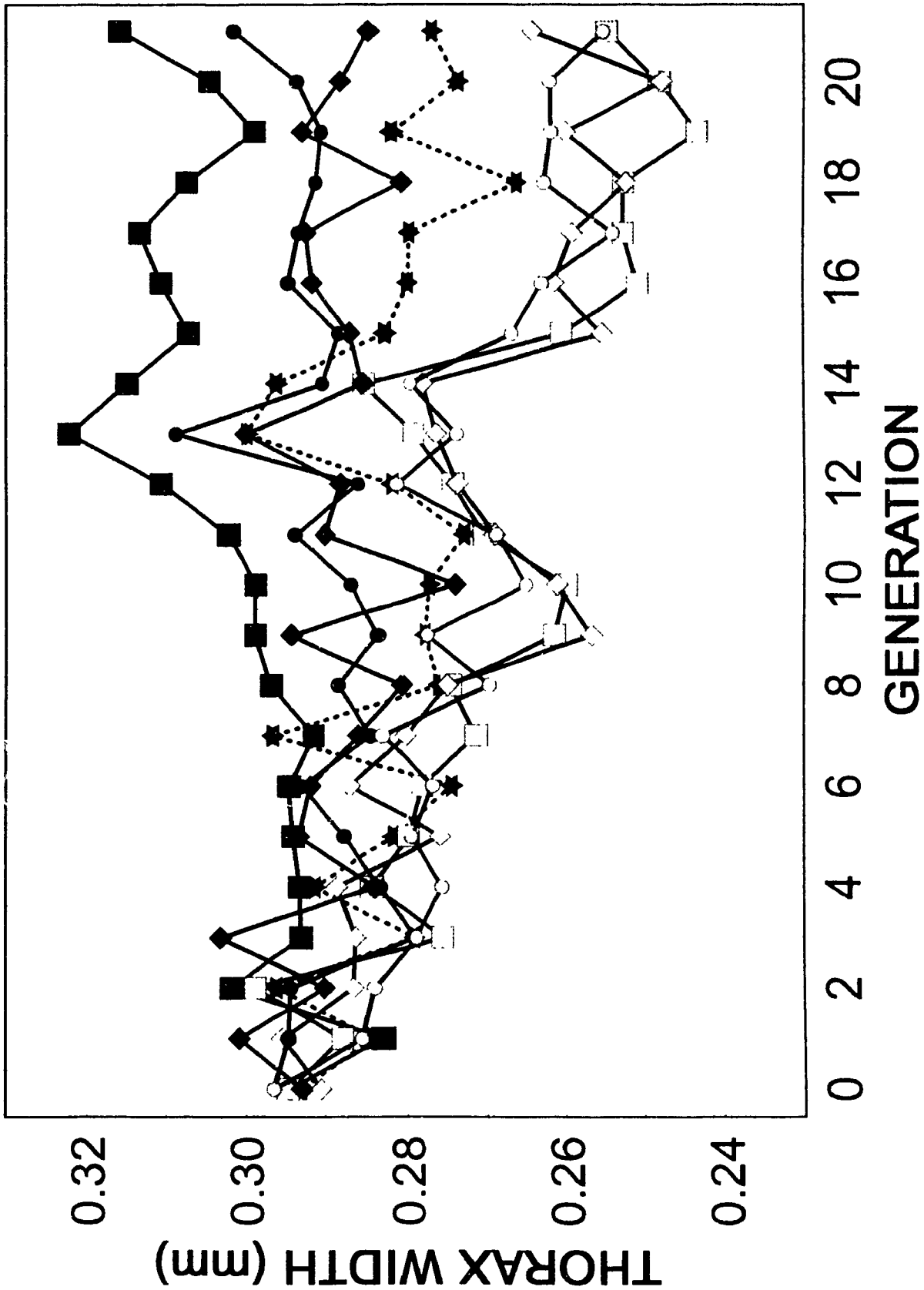

THORAX WIDTH (mm)
0.32
0.30
0.28
0.26
0.24
0
2
4
6
8
10
12
14
16
18
20
GENERATION

Figure 3. Response of *Drosophila melanogaster* females to selection on thorax width. Response is the mean of the 2 control replicates subtracted from the mean of the 2 replicates for each selection regime. Both sexes selected up (■), both sexes selected down (▩), females selected up (◆), females selected down (◈), males selected up (◍), males selected down (●).

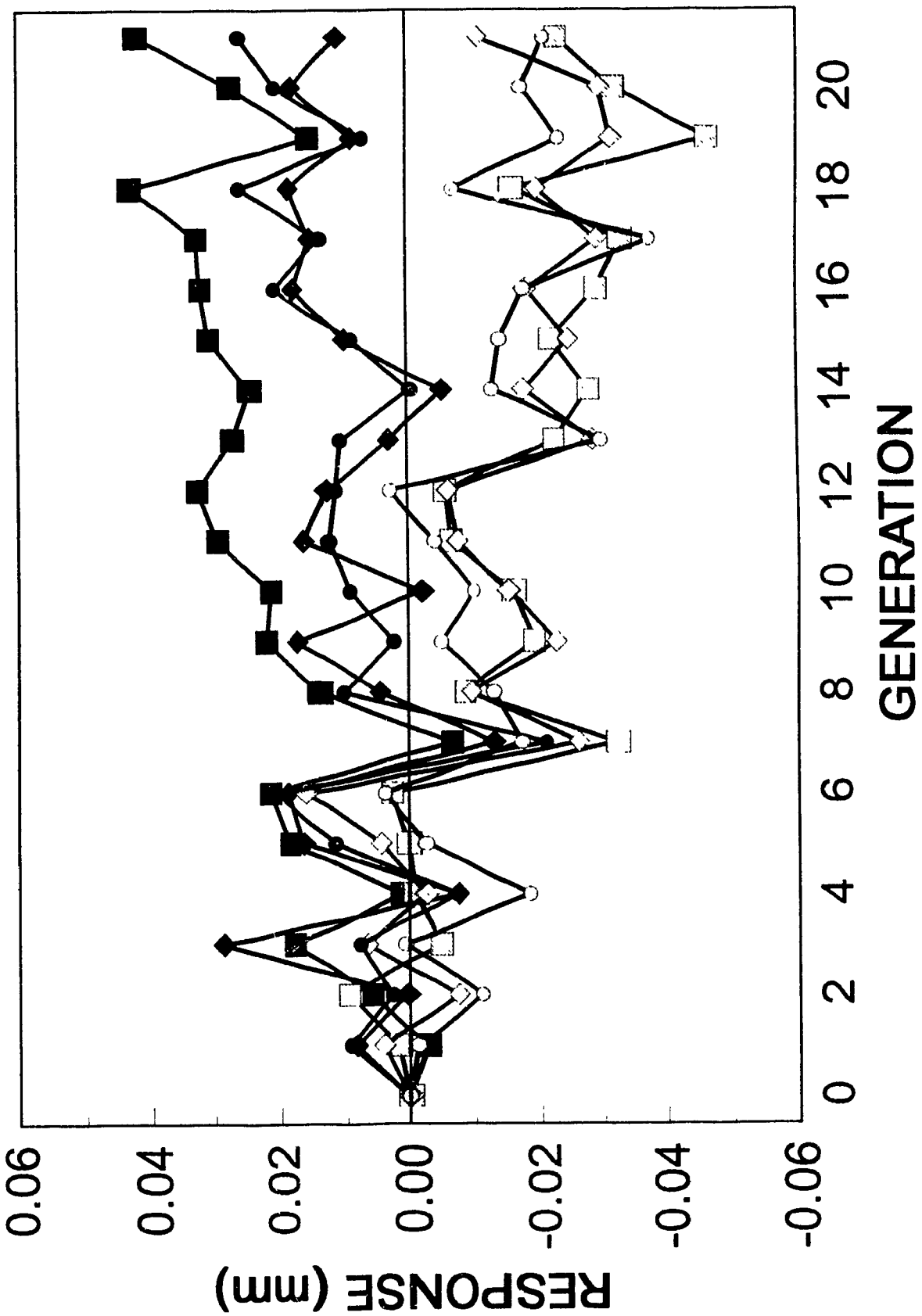
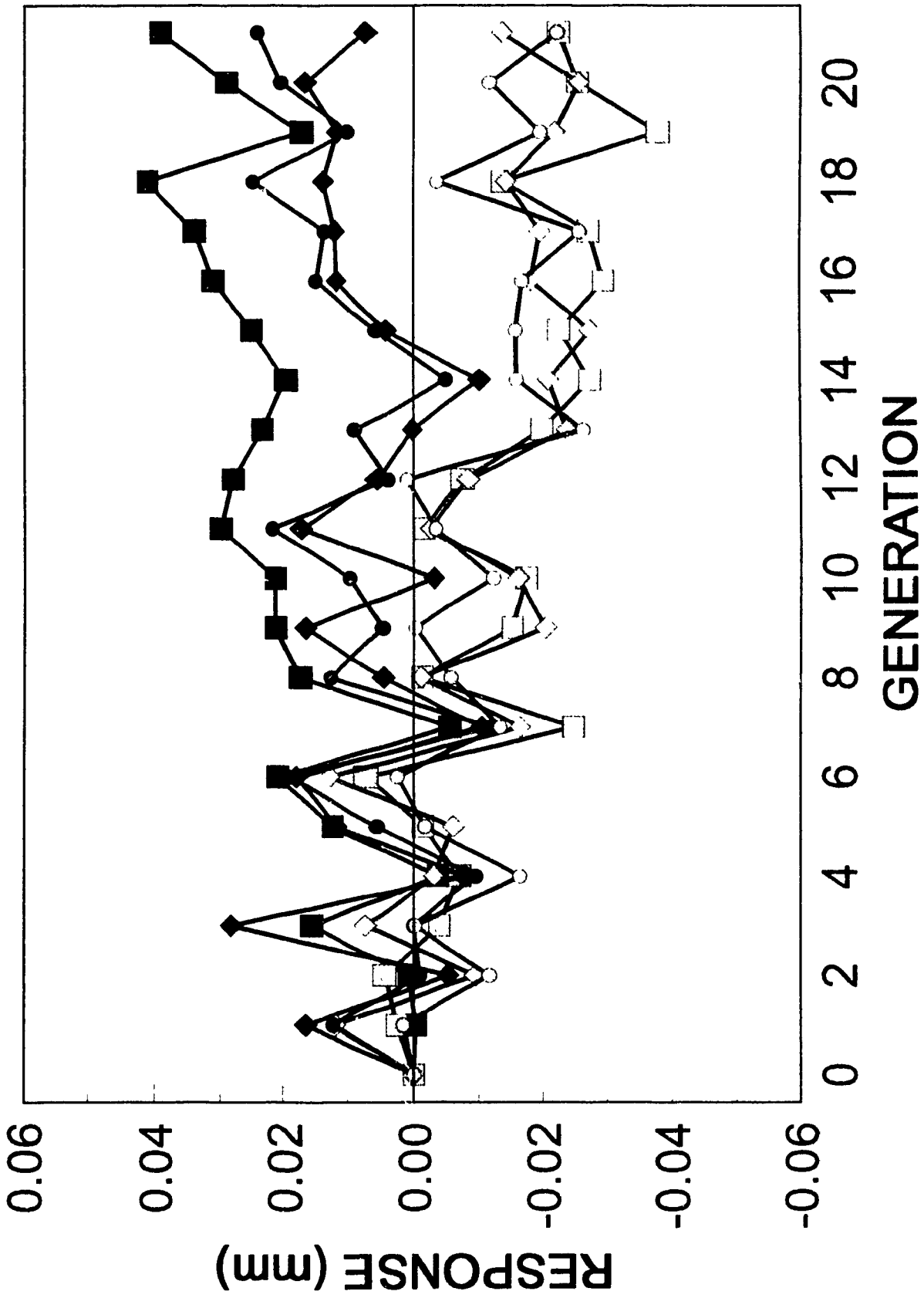

RESPONSE (mm)
0.06
0.04
0.02
0.00
-0.02
-0.04
-0.06
0
2
4
6
8
10
12
14
16
18
20
GENERATION

Figure 4 Response of *Drosophila melanogaster* males to selection on thorax width. Response is the mean of the 2 control replicates subtracted from the mean of the 2 replicates for each selection regime. Both sexes selected up (■), both sexes selected down (▨), females selected up (◆), females selected down (◈), males selected up (●), males selected down (◑).


RESPONSE (mm)
0.06
0.04
0.02
0.00
-0.02
-0.04
-0.06
GENERATION
0
2
4
6
8
10
12
14
16
18
20

grouping variable (Table 2) The data for this table were adjusted by subtracting the control means in each generation for each selection line, since there was a significant decline in body size in both control lines The response to selection, indicated by generation effects in the adjusted data, is significant in all lines except female up #1 and male down #2

Realized heritabilities and their standard errors (Table 3) were calculated using the methods given by Hill (1972). Heritability is estimated by the slope of the linear regression of cumulative response on cumulative selection differential. The response in each line in each generation is determined by taking the observed mean and subtracting the mean of the 2 control lines. The selection differential is the difference between the mean of the selected individuals and the mean of all those measured, with the selection differential of the unselected sex assumed to be zero. The regression is calculated without a constant (the intercept is set to 0), since the control and selection lines come from the same base population, and the difference between lines is assumed to be 0 in generation 0 (Hill, 1972) In Table 3, the numbers in parentheses are calculated from selection on the opposite sex, and are actually the correlated responses multiplied by 1/r, rather than true heritability estimates The value of r (0 93) was taken from the half-sib mating experiment The use of this adjusted correlated response as an estimate of heritability depends on the assumption of equal variances in each sex (from Falconer, 1989, eqn. 19 6)

The usual estimate of the standard error (SE) of the slope in a linear regression assumes only measurement variance, while that of the heritability also includes drift

Table 2 Probabilities derived from covariance analysis of thorax width as a function of generation, with sex as the grouping variable Raw data adjusted by subtracting control means from each line in each generation

Line	Replicate	Sex	Gen	Sex x Gen
Both Up	1	1.2×10^{-7}	3.3×10^{-12}	0.68
	2	5.7×10^{-4}	2.9×10^{-4}	0.89
Both Down	1	1.2×10^{-5}	3.8×10^{-5}	0.37
	2	7.6×10^{-5}	2.1×10^{-10}	0.81
Females Up	1	5.7×10^{-5}	0.95	0.86
	2	3.1×10^{-4}	0.048	0.88
Females Down	1	6.6×10^{-6}	4.7×10^{-10}	0.63
	2	4.3×10^{-6}	8.9×10^{-9}	0.47
Males Up	1	5.5×10^{-6}	5.4×10^{-4}	0.42
	2	6.6×10^{-5}	0.032	0.81
Males Down	1	2.3×10^{-5}	1.7×10^{-5}	0.25
	2	2.8×10^{-4}	0.062	0.89

Table 3 Realized heritabilities and correlated responses of thorax width.

			Females		Males	
Line	Replicate	S	h^2	SE	h^2	SE
Both Up	1	0 0130	0.136	0.025	0.133	0.025
	2	0.0127	0.158	0.027	0.148	0.027
	Avg.		0 147	0.011	0.141	0.008
Both Down	1	-0.0115	0.087	0.023	0.063	0.020
	2	-0.0117	0.174	0.031	0.167	0.030
	Avg.		0.131	0.044	0.115	0.052
Females Up	1	0.0060	0.107	0.048	(0.075)	(0.042)
	2	0.0064	0.129	0.049	(0.118)	(0 049)
	Avg		0.118	0.011	(0.097)	(0.022)
Females Down	1	-0.0066	0.212	0.060	(0.209)	(0.062)
	2	-0.0072	0.163	0.049	(0.143)	(0.047)
	Avg.		0.188	0 025	(0.176)	(0.033)
Males Up	1	0.0064	(0.182)	(0.060)	0.162	0.055
	2	0.0055	(0.127)	(0 059)	0.105	0.052
	Avg.		(0.155)	(0.028)	0.134	0.029
Males Down	1	-0.0065	(0.205)	(0.062)	0.145	0.051
	2	-0.0065	(0.142)	(0.053)	0.133	0.049
	Avg.		(0.174)	(0.032)	0.139	0.006

S = average selection differential
= total selection differential / #generations

h^2 = realized heritability = slope of linear regression of cumulative response vs cumulative selection differential

Avg is the average of the two slopes per selection type

SE = standard error of the slope using the method of Hill (1971). The SE of the average is the SE of the two slope estimates

Numbers in parentheses are correlated responses and can only be used as heritability estimates if additional assumptions are made (see text)

variance (Falconer, 1989). Because of this, the usual SE will underestimate the SE of the heritability. The SE estimates in Table 3 have been corrected by applying the following formulae (Hill, 1972):

$$V(b_t) = (1/S^2t^2)(t\sigma_d^2 + \sigma_e^2 + h^2\sigma_e^2) \quad (4)$$

where $(V(b_t))^{1/2}$ is the SE, S is the average selection differential per generation (which is assumed to be constant through the course of the experiment), t the number of generations of selection, h^2 the narrow-sense heritability, σ_e^2 the error variance, and σ_d^2 the drift variance

$$\sigma_d^2 = 2\sigma_P^2(h^2(1-h^2)/N_{EF} + h^4/M_{EF}) \quad (5a)$$

$$\sigma_e^2 = 2\sigma_P^2(1-h^2)/M_{EF} \quad (5b)$$

$$M_{EF} = (1/4M + 1/4M)^{-1} \quad (6a)$$

$$N_{EF} = (1/4N + 1/4N)^{-1} \quad (6b)$$

where M is the number of animals measured, N the number selected, with subscripts P for phenotypic, and EF for effective number.

The heritability estimates are greater for females than for males in all 12 of the selection lines, and also in the half-sib experiment, indicating that thorax width has a higher heritability in females than in males. The assumption of equal variances for estimating the heritabilities from the correlated responses in Table 3 does not affect this trend. Even without multiplying the responses by 1/r, females still have a higher h^2 than males in 11 of 12 lines, and in the remaining line, they are equal All of the heritability estimates are significantly greater than 0, with the exception of males in replicate #1 of FU. These estimates are much lower than those obtained from the half-sib experiment

(Table 1), and the ratio of female to male heritability was also reduced to 1.13 from 1.53 in the half-sib estimate.

Realized genetic correlations (r) are calculated as:

$r = ((C_m/R_m)(C_f/R_f))^{0.5}$ (Falconer, 1989) (7)

where R is response, C is correlated response, m is males, and f is females. The responses are given by the heritabilities in Table 3, since there is no constant. The correlated responses are the values in parentheses from Table 3 divided by 1/r (i.e. the actual slopes before adjusting). Since there are four single-sex selection lines in each direction, four combinations of response/correlated response estimates can be made for both up and down selection (e.g. the first female up replicate can be compared to each of the two male up replicates, and likewise for the second female up replicate). The four estimates of r for upward selection have a mean of 0.90 (SE = 0.03), while the four downward estimates have a mean of 1.00 (0.04). The two estimates are not significantly different, but it would not be unusual for them to be asymmetric (Bohren et al., 1966; Falconer, 1989). They also agree closely with the r (0.93) from the half-sib experiment. The high phenotypic correlation between the sexes can be seen by comparing Figures 1 and 2, where changes in one sex are closely mirrored by changes in the other.

3) Correlated response of dimorphism to selection on thorax width

The sex by generation interaction term, derived from the covariance analysis of thorax width (Table 2) was not significant in any of the selection lines. Lack of significant interactions mean that there were no significant differences between the responses of the two sexes in any given line.

Change in dimorphism was also analyzed by regressing SSD_d (difference) and SSD_r (ratio) on generation (Table 4). Considering first the SSD_d results, 4 of the 12 selection lines (BD1, FD2, MU1, and MD1) and 3 of the 6 averaged lines (BD, FD, and MD) had significant slopes over 21 generations All of these slopes were negative If significance levels are disregarded, 8 of the 12 selection lines and 5 of the 6 averages are negative For the SSD_r slopes, 4 of 12 lines (BU1, BD1, MU1, MD1) and 1 of 6 averages (BU) are significant, and again, they are all negative. Both replicates had positive slopes only in one selection type (FU). This was true for both SSD_d and SSD_r In order to get a better indication of which sex, if either, is primarily responsible for the changes in dimorphism, Table 4 also shows the regressions of adjusted response (control subtracted) vs generation for both sexes. The 4 significant negative SSD_d slopes can be seen to be due to faster decrease in females in the down lines, and faster increase in males in MU1 The same explanation accounts for the negative SSD_r slopes in BD1, MD1, and MU1 The negative slope in BU1 is best explained by a slightly faster upward response in males, coupled with a decreased ratio due to increased average thorax width.

4) <u>The models</u>

In his 1980 paper, Lande gave recursion equations for the evolution of the SSD of multiple correlated traits, which necessitated the use of matrix notation. Since only one trait, thorax length, is being considered here, the equations simplify to:

$$Z_n = Z_{n-1} + 0.5(\sigma^2_{Am} D_{m(n-1)} + D_{f(n-1)} B) \qquad (8a)$$

$$Y_n = Y_{n-1} + 0.5(\sigma^2_{Af} D_{f(n-1)} + D_{m(n-1)} B) \qquad (8b)$$

Table 4 Slopes[a] of linear regressions of SSD_d and SSD_r on generation, and corrected response for thorax width on generation.

Line	Replicate	SSD_d	SSD_r	Response	
				Females	Males
C	1	0.07	2.1		
	2	-1.1	-2.5		
	Avg	-0 50	-0.2		
BU	1	-1 7	-11**	19***	20***
	2	-0.93	-5 8	13*	14**
	Avg.	-1.3[+]	-8.2**		
BD	1	-4.5**	-11***	-14***	-9.3**
	2	-1.3	6.1	-23***	-22***
	Avg	-2.9**	-2.57		
FU	1	0.97	5.6	0.32	-0.65
	2	0.90	3.0	6.5	5.6
	Avg.	0.94	4 3		
FD	1	-2.3	.08	-20***	-18***
	2	-3.3*	-8.0	-9.9.**	-6.6
	Avg.	-2.8*	-4.0		
MU	1	-4.5*	-18*	8.1[+]	13**
	2	1.2	4.4	6.2[+]	5[+]
	Avg.	-1.6	-6.6		
MD	1	-6.7**	-19*	-17**	-11**
	2	0.79	6.9	-4.9	-5.7
	Avg.	-3.0*	-5.9		

[a] All slopes in table should be multiplied by 10^{-4}

2-tailed p values

H_0: slope = 0 + $p \leq 0.10$; * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$

C = Control

BU = Both Up

FU = Female Up

MU = Male Up

BD = Both Down

FD = Female Down

MD = Male Down

$$D_{mn} = (\theta_m - Z_n)/(\omega_m + \sigma_{Pm}) + Q_m \quad (9a)$$

$$D_{fn} = (\theta_f - Y_n)/(\omega_f + \sigma_{Pf}) + Q_f \quad (9b)$$

$$B = r((\sigma_{Af}\sigma_{Am})^{1/2}) \quad (10)$$

where Z_n and Y_n are the trait values at generation n in males and females respectively, σ^2_A is the additive genetic variance, Q is a measure of (sexual) selection, θ is the optimal (natural selection) trait value, B is the genetic covariance between the sexes, D is equivalent to the selection differential divided by the phenotypic variance, and ω is the width of the Gaussian fitness function for stabilizing selection

Parameter estimates used to test the models were obtained from the half-sib experiment (heritabilities and genetic correlations) and from a large sample of flies (n 400) measured at the time of the initial generation of the selection experiment (phenotypic standard deviations and initial thorax width). This initial generation of flies was grown under conditions identical to those under which the selection lines were kept The values used were $Z_0 = 0.2943$, $Y_0 = 0.3262$, $\sigma_{Pm} = 0.009406$, $\sigma_{Pf} = 0.01088$, h^2_f 0.66, h^2_m 0 43, and $r = 0.93$. The phenotypic standard deviations are not from the total variance, but from the average within vial variances. This is because 25% selection is applied within each vial, so selection acts on the within-vial variance. In the models, Q is a measure of the intensity of sexual selection, but it is in units of (selection differential)/(phenotypic variance), instead of the usual (selection differential)/(phenotypic standard deviation) Since the theoretical expectation of selection intensity when selecting 15 from 60 measured animals is known to be 1.25 (Appendix Table B, Falconer [1989]), this number was divided by the average σ_P, giving a Q rounded to 125, which was used to

approximate the level of artificial selection. For the unselected sex, Q is equal to zero. Values for ω_m and ω_f were estimated by trial and error, by choosing values which produced a response over 21 generations of selection that was within the range of that found in the experiment (e.g. about 0.02-0.05 mm), but that also put a reasonable limit on the theoretical equilibrium values reached after, in some cases, several thousand generations (e.g. no more than about 0.47 mm for females - the largest fly measured in the experiment was close to 0.4 mm). The value used for ω was 0.0003.

Since the control lines decreased in size over the course of the experiment, it is probable that natural selection was favouring a smaller body size than that initially seen. The model predictions assume that the natural selection optimum for each sex was equal to the average size of the control lines measured over the last three generations of selection: 0.3060 for females and 0.2774 for males. These are 0.0202 and 0.0169 mm smaller than the initial sizes for females and males respectively. The models were also tested assuming that the initial body sizes for each sex were already at their natural selection optima (results not shown), but the predictions were consistently less accurate than those with natural selection for reduced size.

The Lande model assumes equal variances and heritabilities in each sex, so the averages of the observed values were used to calculate the predicted thorax width values. The Cheverud model assumes that there are heritability and variance differences between the sexes, so the observed values for each sex were used, rather than their averages.

These models are based on the assumption of additive gene action, where a trait is coded for by a large number of genes, each of small and approximately equal effect

This assumption requires that the input data be provided using a measurement scale on which gene action is truly additive Our understanding of the genetics of quantitative traits is rarely sufficient to allow us to know what the most suitable scale is, but in practice, only the arithmetic and geometric are commonly used One of the assumptions of the models is that the variances remain constant through time, and are therefore independent of body size Since phenotypic variance is often correlated with body size, logarithmic transformation was suggested by Lande (1980) as a means of making this assumption more realistic. Log-transforming the original data to obtain variance estimates reduced the female to male variance ratio from 1 16 to 1 04 For the predictions, each model (Lande and Cheverud) was run using both arithmetic and logarithmic input The 4 models tested were therefore· Model I (Lande, arithmetic input), Model II (Cheverud, arithmetic input), Model III (Lande, logarithmic input), and Model IV (Cheverud, logarithmic input) For Models III and IV, the phenotypic variances and means are those of the log-transformed raw data, heritability remains the same, and all other parameters are calculated in terms of numbers of phenotypic standard deviations The significance of the transformation, other than keeping variances more stable, is that size differences between the sexes can be caused by the differential expression of the same genes in the two sexes If genes in females are expressed as some multiple of their expression in males, then the effects are only additive if the measurement scale is geometric There was no *a priori* expectation of better results from the log models, since the experimental data showed no correlation between mean body size and phenotypic variance (each averaged over the last three generations) in either sex. Size predictions from the log models were converted back to

arithmetic units for the comparison with experimental results.

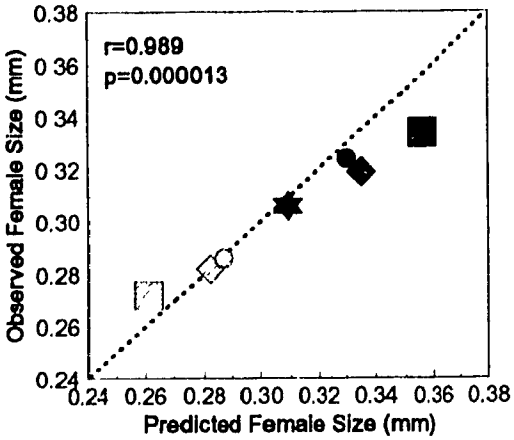
5) Comparing the model predictions with the experimental results

The trends in SSD from the experimental data are difficult to interpret in terms of slopes of SSD on generation, since the replicate lines often have opposite slope directions (Table 4). In addition, there are no selection types in which both replicates produce significant slopes. The general lack of significance is not surprising, due to the large amount of variability between generations within lines, not only of SSD but of body size itself. If the models are used to generate SSD on generation slopes, it isn't clear whether their predictions should be compared to the average slope of the 2 experimental replicates per selection type, only those selection lines which produced significant slopes, only those selection types which had the same direction of slope in both replicates, or some combination of the above. Although it is tempting to use just the 4 significant slopes (for both SSD_d and SSD_r), this ignores all the information gathered through the other selection lines. A further difficulty is that the models generally predict non-linear slopes, which would be very difficult to detect at a significant level in the observed data since only 21 data points are available for each regression. For these reasons, comparisons between the models and the experimental data were made from the overall mean of each sex from the last 3 generations of selection (19 - 21) in each selection type. This has the additional benefit of considering the differences between lines only at a point when selection has had sufficient time to produce substantial changes in body size.

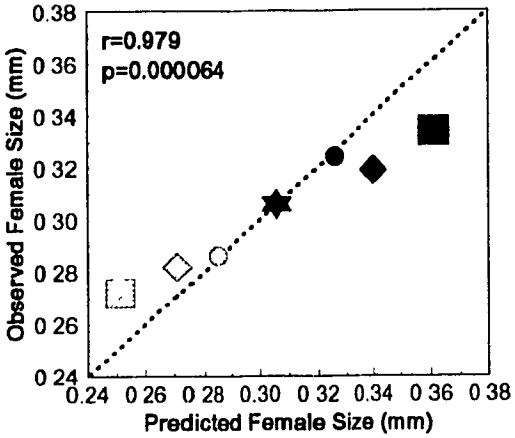
Figures 5 to 8 show the correlation between observed and predicted values for

Figure 5. Correlation (r) and 1-tailed p values for relationship between observed and predicted female thorax width in *Drosophila melanogaster*, using 4 models for the evolution of sexual size dimorphism. Model I = Lande, arithmetic input parameters Model II = Cheverud, arithmetic input parameters. Model III = Lande, logarithmic input parameters. Model IV = Cheverud, logarithmic input parameters. Control ( ), both sexes selected up ( ), both sexes selected down ( ), females selected up ( ), females selected down ( ), males selected up ( ), males selected down ( ). Each observed value is the average from the last 3 generations of selection in both replicates. The predicted values are the average of the predictions for the last 3 generations

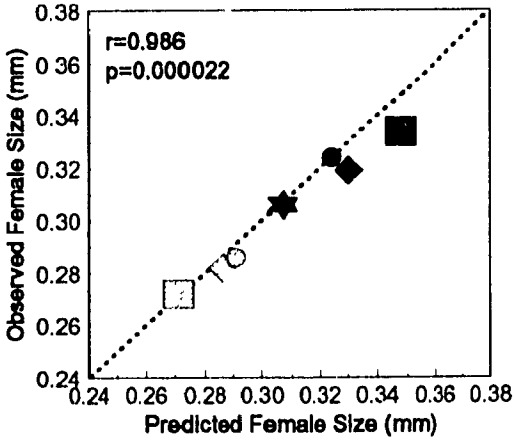
I


r=0.989
p=0.000013
Observed Female Size (mm)
Predicted Female Size (mm)
0.24 0.26 0.28 0.30 0.32 0.34 0.36 0.38

II


r=0.979
p=0.000064
Observed Female Size (mm)
Predicted Female Size (mm)
0.24 0.26 0.28 0.30 0.32 0.34 0.36 0.38

III


r=0.986
p=0.000022
Observed Female Size (mm)
Predicted Female Size (mm)
0.24 0.26 0.28 0.30 0.32 0.34 0.36 0.38

IV

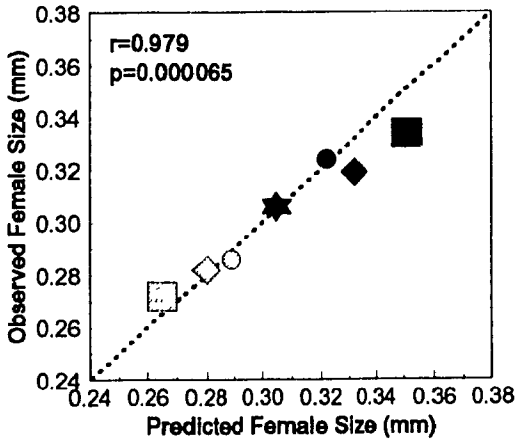
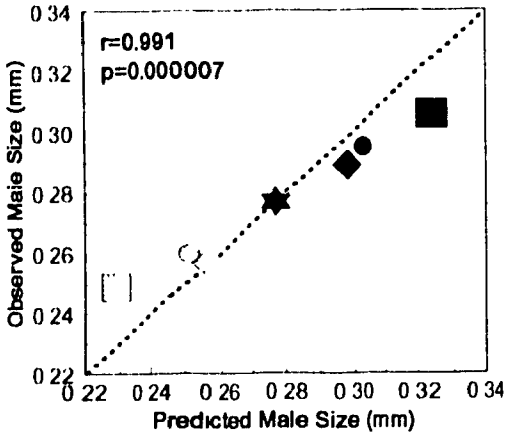
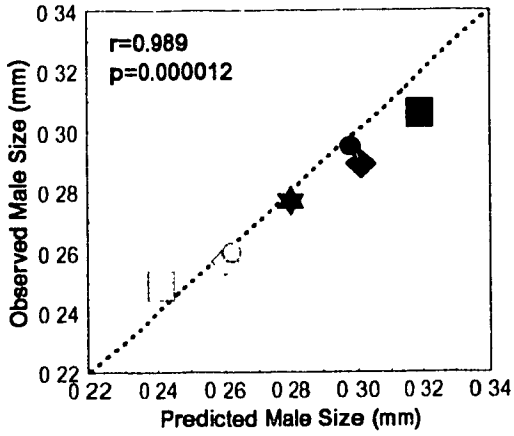

r=0.979
p=0.000065
Observed Female Size (mm)
Predicted Female Size (mm)
0.24 0.26 0.28 0.30 0.32 0.34 0.36 0.38

Figure 6. Correlation (r) and 1-tailed p values for relationship between observed and predicted male thorax width in *Drosophila melanogaster*, using 4 models for the evolution of sexual size dimorphism. Model I = Lande, arithmetic input parameters Model II = Cheverud, arithmetic input parameters. Model III = Lande, logarithmic input parameters Model IV = Cheverud, logarithmic input parameters Control (✸), both sexes selected up (■), both sexes selected down (▩), females selected up (◆), females selected down (◈), males selected up (●), males selected down (◍). Each observed value is the average from the last 3 generations of selection in both replicates. The predicted values are the average of the predictions for the last 3 generations.

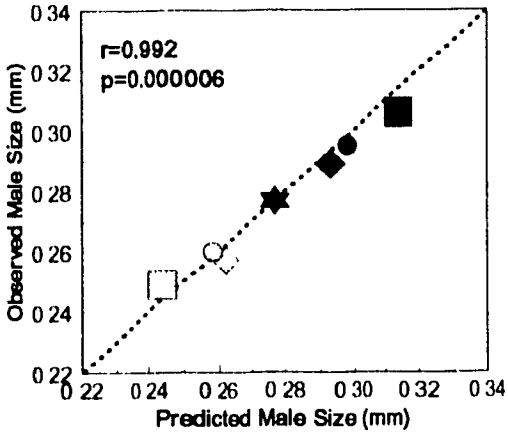
I


r=0.991
p=0.000007
Observed Male Size (mm)
Predicted Male Size (mm)
0 22
0 24
0 26
0 28
0 30
0 32
0 34

II


r=0.989
p=0.000012
Observed Male Size (mm)
Predicted Male Size (mm)
0 22
0 24
0 26
0 28
0 30
0 32
0 34

III


r=0.992
p=0.000006
Observed Male Size (mm)
Predicted Male Size (mm)
0 22
0 24
0 26
0 28
0 30
0 32
0 34

IV

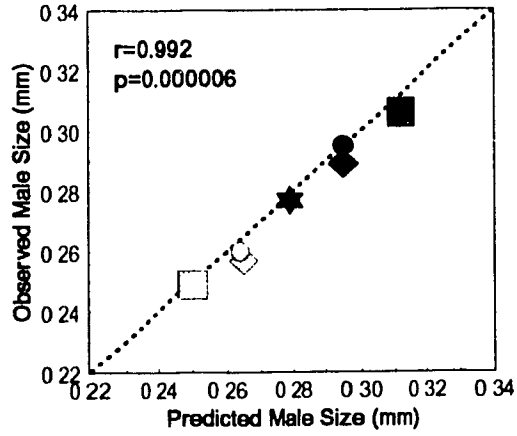
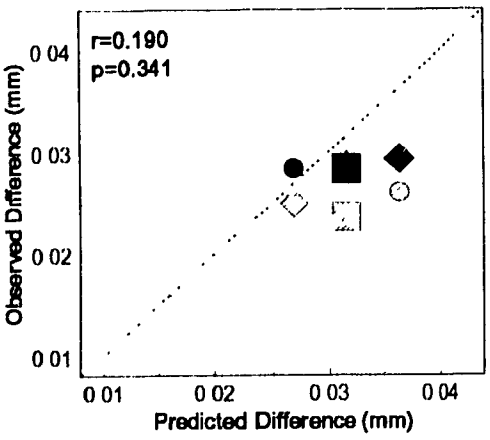
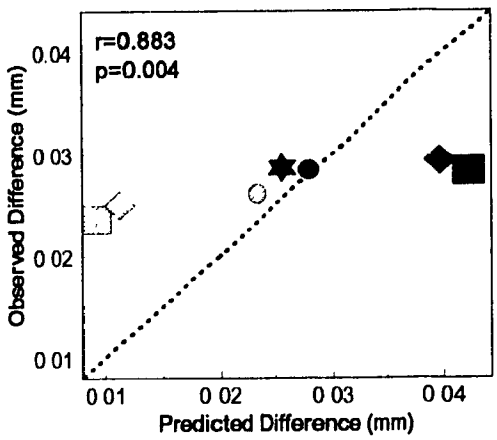

r=0.992
p=0.000006
Observed Male Size (mm)
Predicted Male Size (mm)
0 22
0 24
0 26
0 28
0 30
0 32
0 34

Figure 7 Correlation (r) and 1-tailed p values for relationship between observed and predicted SSD_d for thorax width in *Drosophila melanogaster*, using 4 models for the evolution of sexual size dimorphism Model I = Lande, arithmetic input parameters Model II = Cheverud, arithmetic input parameters Model III = Lande, logarithmic input parameters. Model IV = Cheverud, logarithmic input parameters Control (★), both sexes selected up (■), both sexes selected down (▩), females selected up (◆), females selected down (◈), males selected up (●), males selected down (◉) Each observed value is the average from the last 3 generations of selection in both replicates The predicted values are the average of the predictions for the last 3 generations

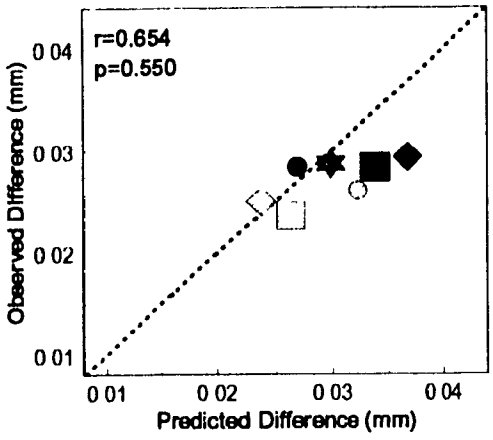
I


r=0.190
p=0.341
Observed Difference (mm)
0 04
0 03
0 02
0 01
0 01
0 02
0 03
0 04
Predicted Difference (mm)

II


r=0.883
p=0.004
Observed Difference (mm)
0.04
0 03
0 02
0 01
0 01
0 02
0 03
0 04
Predicted Difference (mm)

III


r=0.654
p=0.550
Observed Difference (mm)
0 04
0 03
0 02
0 01
0 01
0 02
0 03
0 04
Predicted Difference (mm)

IV

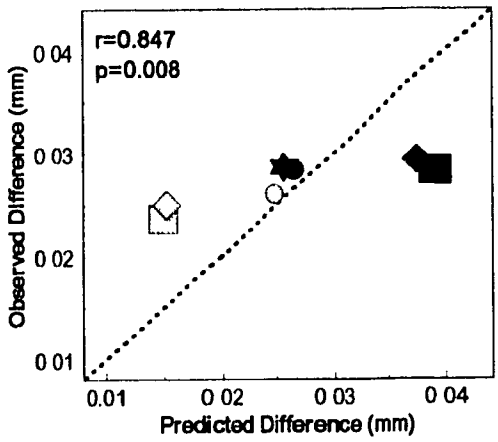
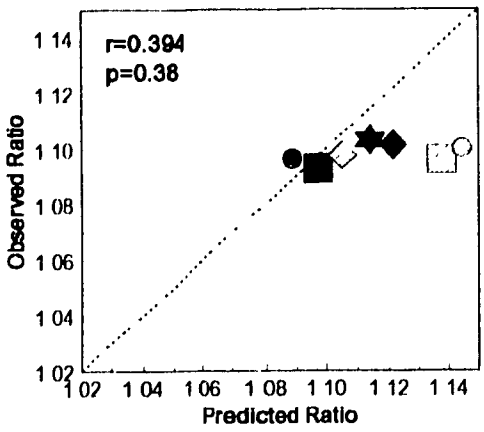
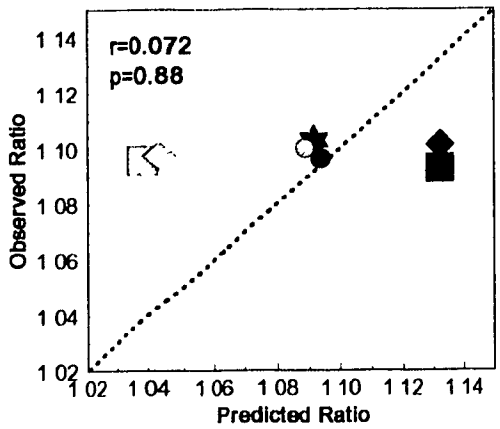

r=0.847
p=0.008
Observed Difference (mm)
0 04
0 03
0 02
0 01
0.01
0 02
0 03
0 04
Predicted Difference (mm)

Figure 8. Correlation (r) and 1-tailed p values for relationship between observed and predicted SSD_r for thorax width in *Drosophila melanogaster*, using 4 models for the evolution of sexual size dimorphism. Model I = Lande, arithmetic input parameters. Model II = Cheverud, arithmetic input parameters Model III = Lande, logarithmic input parameters. Model IV = Cheverud, logarithmic input parameters. Control (★), both sexes selected up (■), both sexes selected down (▩), females selected up (◆), females selected down (◈), males selected up (●), males selected down (◍) Each observed value is the average from the last 3 generations of selection in both replicates. The predicted values are the average of the predictions for the last 3 generations

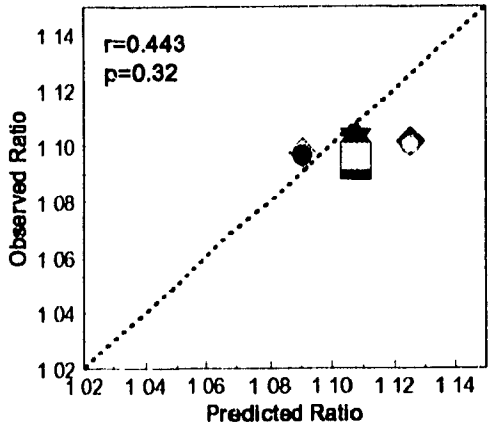
I


r=0.394
p=0.38
Observed Ratio
Predicted Ratio
1 02
1 04
1 06
1 08
1 10
1 12
1 14

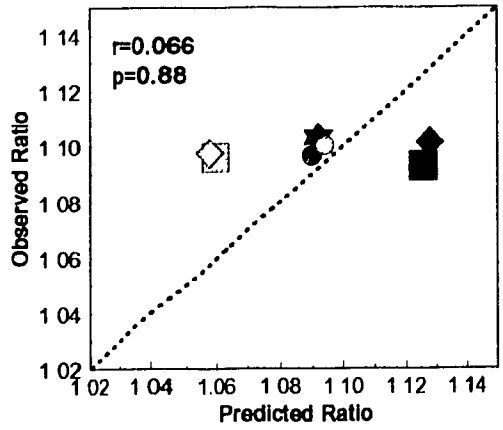
II


r=0.072
p=0.88
Observed Ratio
Predicted Ratio
1 02
1 04
1 06
1 08
1 10
1 12
1 14

III


r=0.443
p=0.32
Observed Ratio
Predicted Ratio
1 02
1 04
1 06
1 08
1 10
1 12
1 14

IV


r=0.066
p=0.88
Observed Ratio
Predicted Ratio
1 02
1 04
1.06
1 08
1 10
1 12
1 14

female size (Fig. 5), male size (Fig. 6), SSD_d (Fig. 7), and SSD_r (Fig. 8) for each of the 4 models. In each figure, the dotted line represents the 1:1 fit of observed and predicted values. A significant correlation can be taken as evidence that the model is at least partially correct (Roff, 1992) All 4 models had significant correlations for body size, with the average being 0.983 for females, and 0.991 for males The Lande models (I and III) were slightly better than the Cheverud models (II and IV) for female size (Figure 5), while all 4 models were equally good for male size (Figure 6). All 4 models overestimated the size of large females (Figure 5), while model II also underestimated the size of small females. The same tendency to overestimate the large and underestimate the small can be seen in the plots for males (Figure 6, I and II). This pattern of greater predicted than observed response is most likely due to the true heritabilities being lower, especially in the up direction, than the estimates obtained from the half-sib experiment at generation 0.

The correlation between observed and predicted SSD_d was significant for the Cheverud models (Figure 7, II and IV), but not for the Lande models. The effect of using log input is more marked in the Lande models (Figure 7, I and III), where the order of selection types is altered, mostly through a decrease in predicted dimorphism in the down selected lines. The opposite effect (increase in predicted dimorphism with log-transformation), can be seen in the Cheverud models, but it is not of sufficient magnitude to change the order between Figure 7 II and 7 IV. Ratio predictions (Figure 8) did not correlate significantly with observed values for any of the models, although the correlations were better for the Lande than the Cheverud models. The effect of taking logs

was the same as in Figure 7 - order was more noticeably changed in the Lande models.

The correlations between observed and predicted female size, male size, SSD_d, and SSD_r were also calculated for each selection type over the 21 generations of selection. The observed value each generation was taken to be the average of the two lines. These results differ from those shown in Figures 5 to 8 in that the correlations being calculated are between the changes through time of the variables, rather than their final values Table 5 shows the results for female and male body size All models performed well, and were comparable for predicting female size, while the log models were better than the linear models at predicting male size. Table 6 shows the results for SSD_d and SSD_r. For SSD_d, all models were again comparable, while for SSD_r, Model III (Lande log) was best. As in Figures 7 and 8, difference predictions were better than ratio predictions.

Another way of comparing the models is to calculate the sum of squares of the residuals with respect to the 1:1 line (Table 7). These values indicate the relative accuracy of each model in terms of predicting the observed results. The log models (III and IV) had a closer fit for all 4 traits, while the Lande models (I and III) were better than the Cheverud models (II and IV) for all traits except male size, where Model II was better than Model I. In every case, the Lande log model (III) gave the smallest residual variance.

The most striking feature of Figures 5 through 8 is that the models had much larger ranges of predicted values than those observed in the actual experiment This is especially noticeable in Figures 7 and 8, where there was very little observed change in either SSD_d or SSD_r The Cheverud models (II and IV) overestimated the observed range in SSD to a much greater extent than did the Lande models (I and III). One possible explanation for

Table 5. Correlations (1-tailed p values) between observed and predicted female and male thorax widths, over 21 generations of selection, for the four models of SSD evolution.

Model		I	II	III	IV
Female	C	0.371 (0.045)*	0.369 (0.045)*	0.363 (0.048)*	0.361 (0 050)*
	BU	0.655 (0.00047)***	0.653 (0.0005)***	0.625 (0.0095)**	0.625 (0 001)***
	BD	0.855 (2.0×10^{-7})***	0 852 (2.4×10^{-7})***	0.817 (1.5×10^{-6})***	0.813 (2.0×10^{-6})***
	FU	-0.878 (4.1×10^{-8})	-0.864 (1.1×10^{-7})	-0.900 (9.0×10^{-9})	-0.872 (6.0×10^{-8})
	FD	0.837 (5.0×10^{-7})***	0.833 (1.0×10^{-6})***	0.806 (3.0×10^{-6})***	0.799 (4.0×10^{-6})***
	MU	-0.200 (0.186)	-0.541 (0.0047)	-0.217 (0.166)	0.000 (0 999)
	MD	0 774 ($1\ 2\times10^{-5}$)***	0.773 (1.3×10^{-5})***	0.756 (2.4×10^{-5})***	0 754 (2.5×10^{-5})***
# significant		5	5	5	5
mean r		0 345	0.296	0 321	0 354
Male	C	0.369 (0.045)*	0.369 (0.046)*	0.366 (0.047)*	0 367 (0 046)*
	BU	0.708 (0.00012)***	0 706 (0.00012)***	0.674 (0.0003)***	0 675 ($2\ 9\times10^{-4}$)***
	BD	0.847 (3.4×10^{-7})***	0.847 (3.3×10^{-7})***	0.810 (2.5×10^{-6})***	0 813 (2.0×10^{-6})***
	FU	-0 776 ($1\ 1\times10^{-5}$)	-0 798 (4.5×10^{-6})	0 850 ($2\ 8\times10^{-7}$)***	0 033 (0 441)
	FD	0 832 (1.0×10^{-6})***	0 830 (1.0×10^{-6})***	0.792 ($5\ 5\times10^{-5}$)***	0 792 (5.5×10^{-5})***
	MU	0.185 (0.205)	0.242 (0.139)	0.246 (0 135)	0 452 (0.017)*
	MD	0 803 ($3\ 5\times10^{-6}$)***	0.805 (3.0×10^{-6})***	0 782 ($8\ 5\times10^{-6}$)***	0 787 ($7\ 0\times10^{-5}$)***
= significant		5	5	6	6
mean r		0 424	0.429	0 646	0 560

* $p \leq 0\ 05$, ** $p \leq 0\ 01$, *** $p \leq 0\ 001$

Table 6 Correlations (1-tailed p values) between observed and predicted SSD_d and SSD_r values, over 21 generations of selection, for the four models of SSD evolution

Model		I	II	III	IV
SSD_d	C	0 098 (0 333)	0 051 (0 411)	0 045 (0 422)	0 033 (0 442)
	BU	0 395 (0 035)*	-0 287 (0 098)	-0 254 (0 127)	-0 255 (0 126)
	BD	0 714 (0 0001)***	0 676 (0 0003)***	0 654 (0 0005)***	0 631 (0 001)***
	FU	0 298 (0 089)	0 333 (0 065)	0 300 (0 088)	0 317 (0 075)
	FD	0 470 (0 014)*	0 439 (0 021)*	0 453 (0 017)*	0 422 (0 025)*
	MU	0 303 (0 085)	0 285 (0 100)	0 302 (0 086)	0 321 (0 073)
	MD	-0 474 (0 013)	0 510 (0 008)**	-0 114 (0 306)	0 502 (0 009)**
# significant		3	3	2	3
mean r		0 258	0 287	0 198	0 282
SSD_r	C	0 054 (0 406)	-0 048 (0 417)	-0 040 (0 429)	-0 064 (0 388)
	BU	0 547 (0 005)**	-0 512 (0 008)	0 578 (0 003)**	-0 474 (0 013)
	BD	-0 264 (0 118)	0 265 (0 117)	0 400 (0 033)*	0 243 (0 138)
	FU	-0 280 (0 103)	-0 271 (0 111)	-0 275 (0 108)	-0 281 (0 103)
	FD	0 196 (0 192)	0 170 (0 225)	0 204 (0 181)	0 158 (0 242)
	MU	0 357 (0 052)	0 348 (0 057)	0 351 (0 055)	0 361 (0 050)*
	MD	-0 314 (0 078)	0 348 (0 056)	-0 290 (0 095)	0 346 (0 057)
# significant		1	0	2	1
mean r		0 042	0 043	0 133	0 041

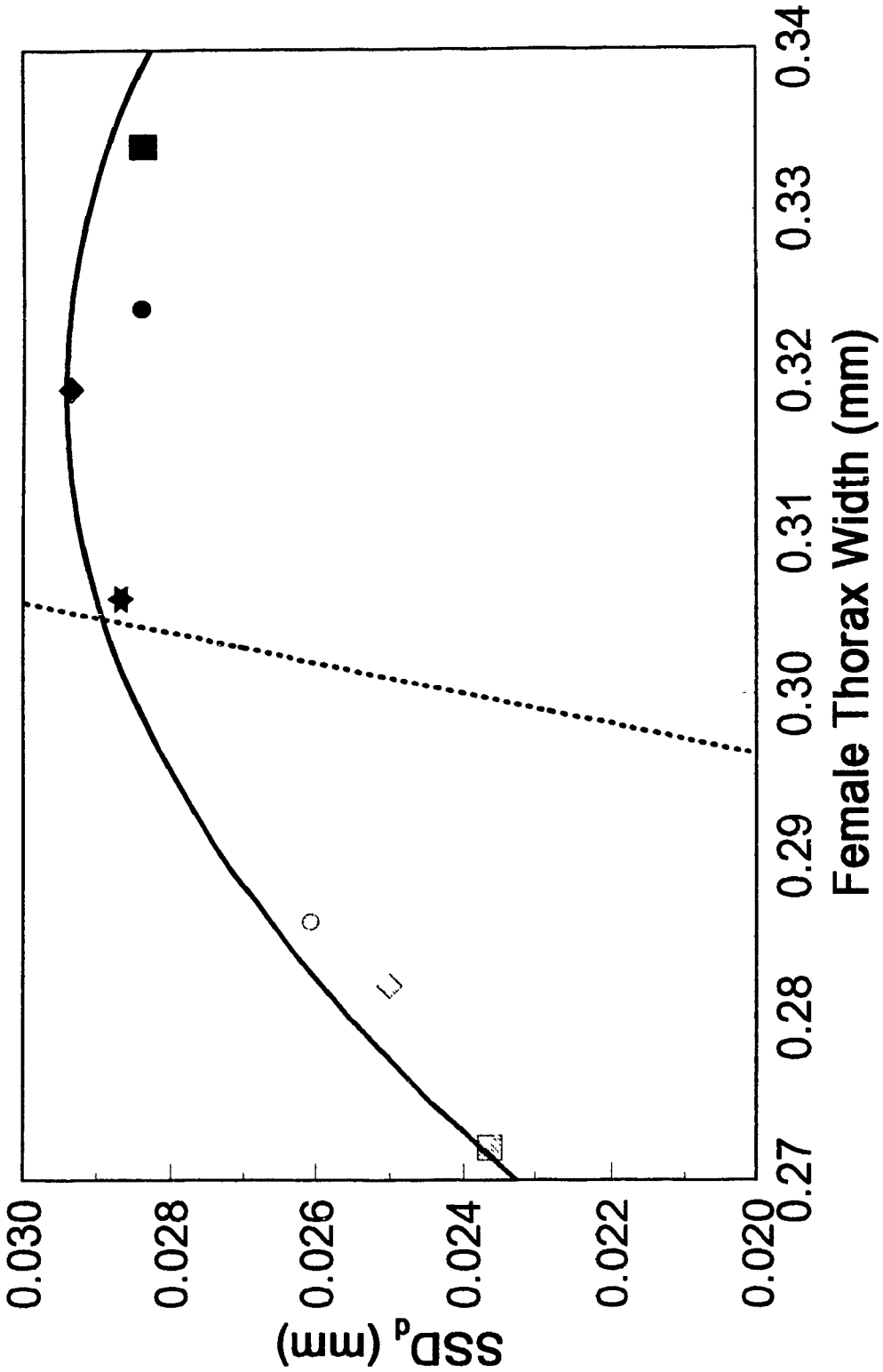
* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Table 7 Sums of squares of residuals about the 1 1 regression line of observed values on predicted values, for the average values in the last three generations of selection (sum of all 7 lines)

Model	I	II	III	IV
Females	0 00091	0 00178	0 00037	0 00054
Males	0 00095	0 00038	0 00014	0 00015
SSD_d	0 00024	0 00071	0 00014	0 00037
SSD_r	0 00444	0 00887	0 00167	0 00480

the good correlation between observed and predicted body size and general lack thereof for SSD, is that small differences between observed and predicted body size, not large enough to greatly reduce the correlations in Figures 5 and 6, could produce large variations in SSD. The increased variance in observed SSD values would lower the probability of finding significant correlations between observed and predicted values. However, if this were true, it might be expected that the variance in observed SSD values would be large compared to the theoretical predictions. Instead, the observed variance in SSD is less than that predicted In addition, SSD values influenced largely by small but random discrepancies between observed and predicted body size measurements should not produce any consistent pattern when plotted against body size (given that no correlation was found between body size and phenotypic variance), other than that caused by the mathematical relationship between the data points plotted on each axis (see below). When the observed SSD_d data are plotted against female body size (Figure 9), it appears that the slope declines at high body sizes. Since female body size appears on both axes, the variables are not mathematically independent, and the null hypothesis for a linear regression is no longer a slope of zero. To estimate the null slope for these data, 1000 pairs of normally distributed random variables with means and standard deviations the same as those of the distributions of the experimental body size data (for the last 3 generations) were generated The average slope values for 5 sets of 1000 pairs of random variables were used The best-fit regression line through these simulated data is shown as the dotted line in Figure 9 The solid line is the regression through the observed data. The equation for the observed data is $y = -0.233 + 1.646x - 2.581x^2$, with both the linear

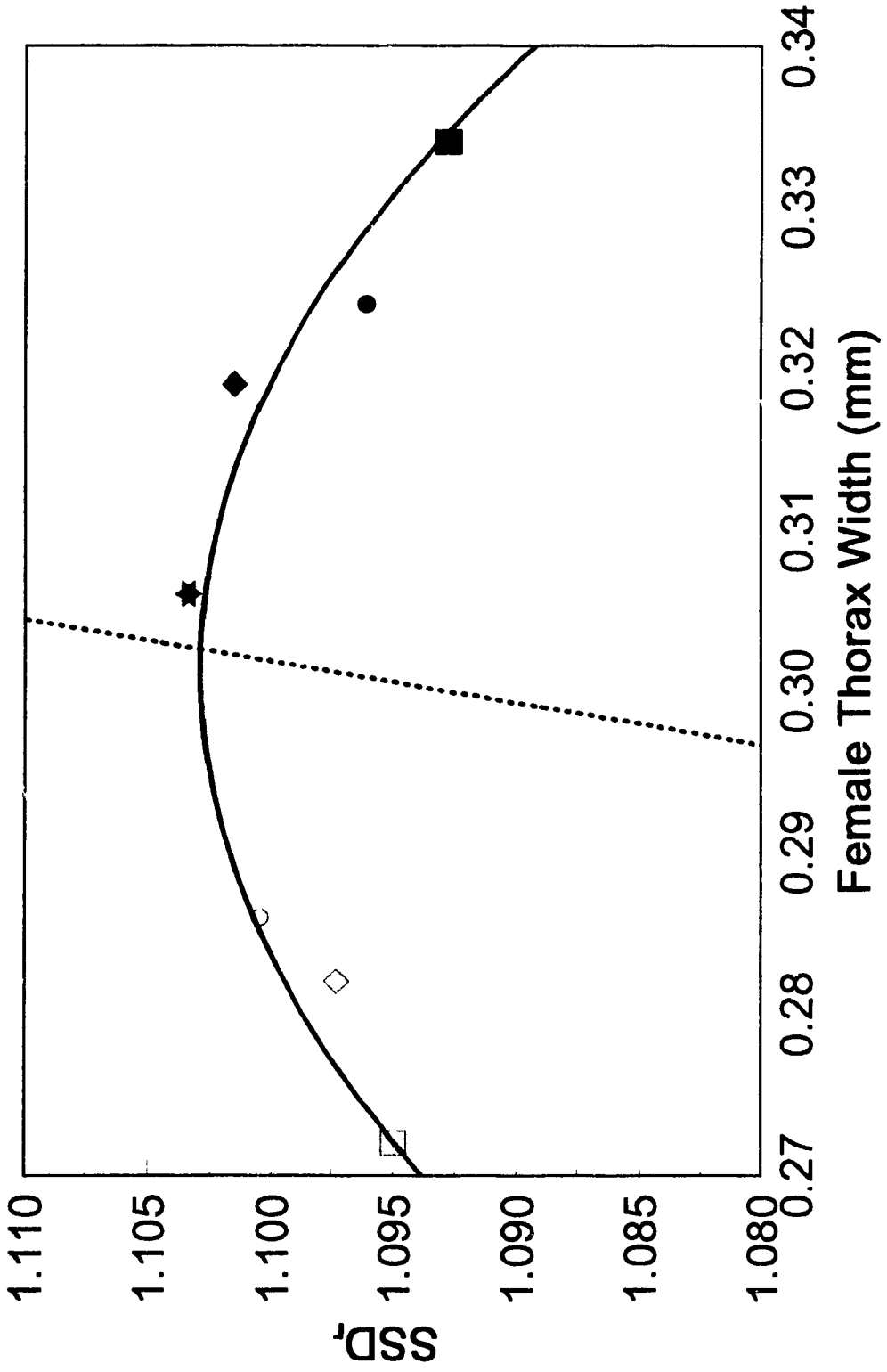
Figure 9. Plot of observed SSD_d vs female size in *Drosophila melanogaster*, each data point from the averages of the two replicates per selection regime over the last 3 generations of selection. Control (★), both sexes selected up (■), both sexes selected down (▨), females selected up (◆), females selected down (◈), males selected up (●), males selected down (◍). See text for an explanation of the regression lines.


0.030
0.028
0.026
0.024
0.022
0.020
SSD$_d$ (mm)
0.27
0.28
0.29
0.30
0.31
0.32
0.33
0.34
Female Thorax Width (mm)
O
U

(p=0.006) and quadratic (p=0.005) terms being significant (p = 0.0006, F = 80.7, df = 4, for the overall regression). The simulated data yield an equation of y = -0.291 + 1.049x ($p = 1 \times 10^{-15}$, F = 1220, df = 998 All values shown for the simulated data are averages from the 5 sets of random variables). This shows that the curvilinear relationship between SSD_d and female body size is not caused by the linear mathematical relationship between the variables on the x- and y-axes. Although SSD_d does decrease with decreasing female body size, it does not increase with increasing female size, as would be expected from the equation of the simulated data. The pattern of SSD_d change is different not only from that predicted by the Lande and Cheverud models, but also from the expectation from the mathematical dependence of the variables.

Figure 10 shows the relationship between SSD_r and female body size. From the plot, it appears that SSD_r decreases as body size moves in either direction from the centre. The equation of the observed data is $y = 0.269 + 5.533x - 9.178x^2$, with both the x (p 0.007) and x^2 (p = 0.006) terms being significant (p = 0.015, F = 14.2, df = 4, overall) The dotted line (calculated from the simulated data described above) has an equation of y = -0.06 + 3.841x, ($p = 1 \times 10^{-15}$, F = 971, df = 998) To see if these observed trends were true for a larger sample of data, SSD_r and SSD_d were plotted against female body size for all 21 generations of selection for all 7 selection types (average values of the two replicates per type) (plots not shown). In each case, the best fit curve through the data was a quadratic equation (SSD_d: $y = -0.091 + 0.665x - 0.902x^2$, with p values of 0.066 and 0.028 for x and x^2 respectively, $p = 1 \times 10^{-15}$, F = 52.2, df = 144, overall; SSD_r: $y = 0.723 + 2.37x - 3.72x^2$, with p values of 0.044 and 0.053 for x and x^2 respectively, p =

Figure 10 Plot of observed SSD_r vs female size in *Drosophila melanogaster*, each data point from the averages of the two replicates per selection regime over the last 3 generations of selection. Control (✸), both sexes selected up (■), both sexes selected down (▨), females selected up (◆), females selected down (◈), males selected up (●), males selected down (◍). See text for an explanation of the regression lines.


1.110
1.105
1.100
1.095
1.090
1.085
1.080
SSD
0.27
0.28
0.29
0.30
0.31
0.32
0.33
0.34
Female Thorax Width (mm)

0.017, F = 4.21, df = 144, overall). Both of these 21 generation curves, like those obtained from the data on the last 3 generations, is concave downwards.

Thus it appears that there is a relationship between female size and SSD, different from that expected simply through the mathematical relationship of the variables. The curvilinear relationship (for both SSD_d and SSD_r) is found even though the data points consist of SSD's from lines which have undergone very different selection regimes. This suggests that some biological factor related to body size itself may influence SSD to a greater extent than does differential selection intensity on the two sexes.

Another method of determining the relationship between body size and dimorphism, which does not have the drawback of using non-independent variables, is to test for SSD allometry, by plotting log(female thorax width) on log(male thorax width). A model 2 major-axis regression was calculated as in Sokal and Rohlf (1981), and tested against the null hypothesis of isometry (slope = 1.0). The data used were the mean female and male thorax widths over the last 3 generations in each selection type. If female size is related to male size by the general allometric equation

$$\text{female size} = a(\text{male size})^b \qquad (11)$$

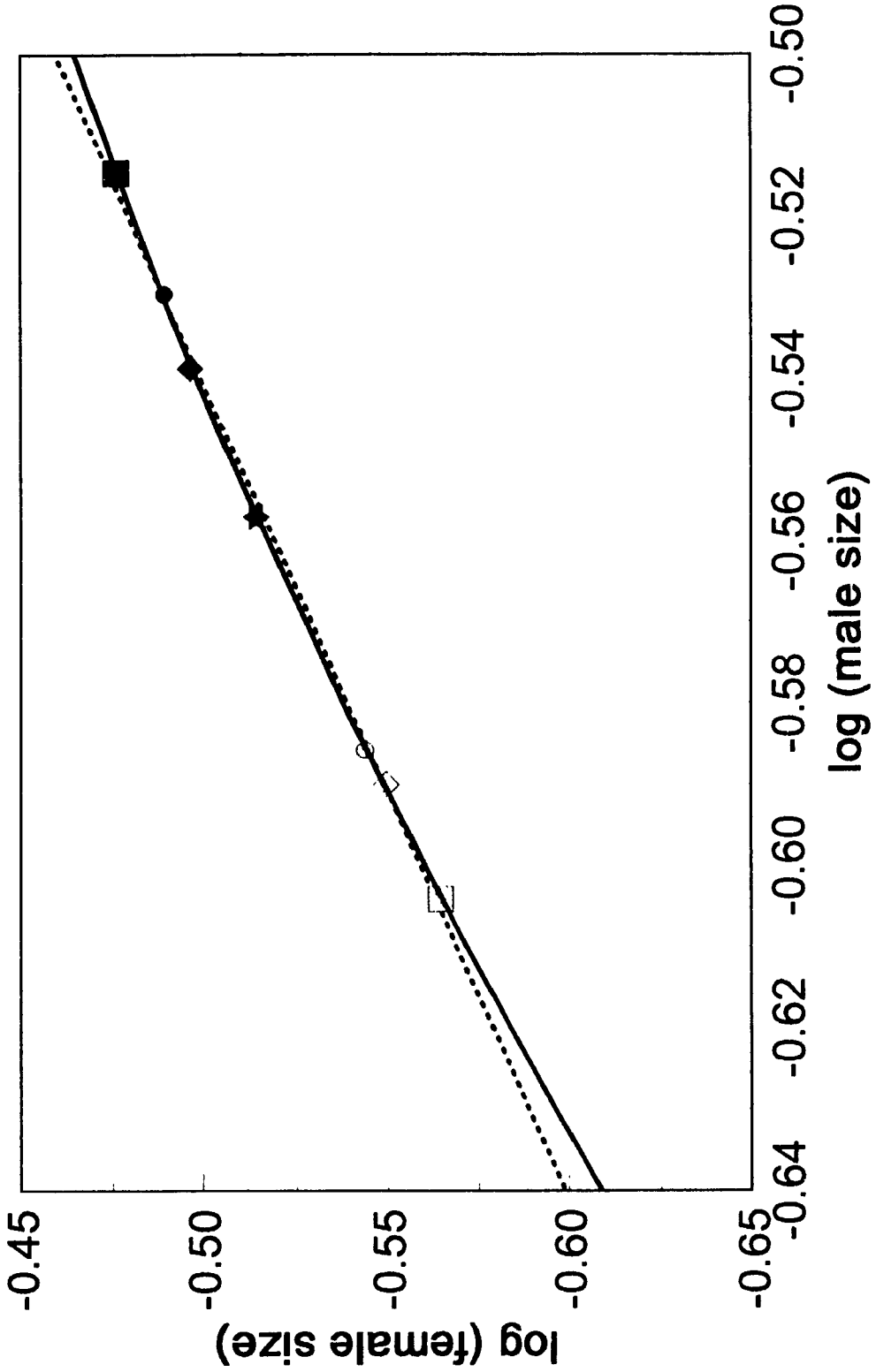
then

$$\log(\text{female size}) = \log(a) + b(\log(\text{male size})) \qquad (12)$$

When log(female size) is plotted against log(male size) a straight line will result. If the slope (b) of this line is 1.0, the relationship between female and male size is isometric (the ratio is constant and independent of size). A slope of greater than one (hyperallometry) indicates that females increase in size more rapidly than do males, while

a slope of less than one (hypoallometry) indicates the opposite. Figure 11 shows that this regression produced a slope not significantly different from 1 0 (95% confidence limits 0.93 - 1.06), so there is no evidence of allometry for SSD Since the plot is slightly concave downwards, the data were tested for non-linear allometry by calculating a polynomial regression including both a linear (log(male size)) and a quadratic (log(male size))^2 term. A model I polynomial regression was used since model 1 and model 2 regressions give very similar results if the r^2 (0.998 for these data) is higher than 0.9 (LaBarbera, 1989). This test yielded an equation of log(female size) = -0.548 -1.102 log(male size) -1.871(log(male size))^2, where both the linear ($p = 0.036$) and quadratic ($p = 0.004$) terms were significant ($p = 3\ 9 \times 10^{-8}$ overall) This indicates that there is a significant non-linear relationship between log(female size) and log(male size), with the slope of the regression line decreasing as body size increases. Males increase in size more rapidly in the upper part of the curve, while females decrease more rapidly in the lower part.

Figure 11. Major-axis regression of log (female thorax width) vs log (male thorax width) for the data of the last 3 generations of selection. Control (★), both sexes selected up (■), both sexes selected down (▩), females selected up (◆), females selected down (◈), males selected up (●), males selected down (⊛). The dashed line is the linear model: log(female size) = 0.037 + 0.994 log(male size), r^2 = 0.998, p = 5.6 x 10^{-8}. The solid line is the quadratic model: log(female size) = -0.548 - 1.102 log(male size) - 1.871 (log(male size))^2, r^2 = 1 000, p = 3.9 x 10^{-8}.


log (female size)
-0.45
-0.50
-0.55
-0.60
-0.65
log (male size)
-0.64
-0.62
-0.60
-0.58
-0.56
-0.54
-0.52
-0.50

DISCUSSION

Selection on thorax width for 21 generations produced strong responses that were consistent with the selection type applied, and with the high heritability estimates obtained from the half-sib experiment The estimate of the genetic correlation between the sexes was similar in the half-sib and the selection experiments. This value ($\geq$ 0.90) was high and in agreement with other studies, which usually find correlations of greater than 0.9 for morphological traits (Lande, 1980; Roff and Fairbairn, 1993)

The heritability values from the half-sib experiment are also high and consistent with those typically found for morphological traits in *Drosophila* (e.g. Roff and Mousseau, 1987; Cowley and Atchley, 1988; Wilkinson et al., 1990), and other animals in general (Mousseau and Roff, 1987) Female heritability for thorax width was higher than that of males, as estimated from the half-sib experiment and each of the selection lines. Female phenotypic variance was also higher than that of males. There do not appear to be any consistent differences between the sexes in heritability and phenotypic variance estimates reported in the literature. In most cases, estimates are not significantly different between the sexes, or are inconsistent between different studies (Maria, et al., 1993) Different traits within the same organism often show both directions of bias in terms of heritabilities (e g. Cowley and Atchley, 1988 in *D. melanogaster*; Maria et al., 1993 in sheep). Since little is known about the genetic basis of quantitative traits in general, and especially how these are differentially expressed in the two sexes, it would not be surprising if there were often differences between the sexes with regards to phenotypic variance and heritability. If this is a common phenomenon, then the implications of

Cheverud et al.'s theory are important in that SSD may be able to evolve more quickly than is generally believed possible, even if genetic correlations between the sexes are very high. Although Lande agrees that such variance dimorphism could be important (Lande, 1987), his 1980 paper is seldom interpreted by others in this way

Realized heritabilities were considerably lower than the half-sib estimates, which is not uncommon (Falconer, 1989). This is because half-sib estimates are only reliable for the first few generations, due to changing gene frequencies (through selection and drift) and increasing forces of natural selection opposing the artificial selection as the experiment progresses. Environmental conditions were also more variable in the selection experiment, due to competition between much larger numbers of larvae Thus, environmental variance probably resulted in a reduction in h^2

Dimorphism was slower to evolve than was thorax width, as predicted by all the models. Lande (1987) states that in the first phase of sexual selection, size change in the two sexes is nearly parallel. This is supported by the covariance analysis (Table 2) which failed to detect any significant differences between the slopes of the responses of the two sexes in any of the selection lines. Meagher (1992) suggests that genes with sex-limited effect may respond rapidly to single-sex selection (presumably through changing allele frequencies), so that a fast change in dimorphism is possible Once the sex-limited genes approach fixation, the sexes are less able to evolve independently This type of response cannot be supported by the results of the present study

Support for the Lande model of slow SSD evolution (as opposed to the fast response of sex-limited genes hypothesis) has previously been found in the fossil record,

where Wright (1993) concluded that skeletal elements of extinct peccaries followed the expectations of the Lande model in that sexually selected structures were found in both sexes before they were slowly reduced in females Also, single-sex artificial selection studies have often shown that the selected sex responds more than the unselected sex, but the difference between the two is seldom significant (e g. Frankham, 1968; Alicchio and Palenzona, 1971, Wilkinson, 1993)

Although SSD was slow to change compared to overall body size, there was a distinct trend in most lines for a decrease in SSD (Table 4), and this was significant in 4 lines for SSD_d and 4 lines for SSD_r In addition, the finding of non-linear allometry for SSD (Figure 11) indicates that SSD decreased with both directions of selection. This trend could be explained if there were some mechanism by which movement away from a point intermediate in size between that of the two sexes was more difficult than movement towards that point (with the limitation that both sexes cannot easily move simultaneously towards or away from this intermediate point due to the high genetic correlation). For instance, selection upwards on both sexes would move females away from this point, and males towards it, so male response would be faster and SSD would decrease Similarly, selection downwards on the two sexes would move females towards this point and males away from it, so that females would respond faster than males and SSD would again decrease The obvious candidate for such a mechanism is natural selection favouring an intermediate size To test this hypothesis, natural selection for an intermediate optimum was added to the models. However, to produce slope trends similar to the those found in the experiment, this natural selection had to be unrealistically strong (unrealistic in that

if stabilizing selection for intermediate values was really this strong, flies could only be maintained at their original starting sizes by some opposing force of selection which would be equivalent in strength to 80% truncation selection) Since there is no reason to believe such strong selection exists, the explanation via intermediate natural selection cannot be justified

Another possible explanation for the general reduction in SSD with selection is that heritability for thorax length is asymmetric For instance, if heritability for increased size in females was less than that for decreased size, the observed SSD trends might be produced. Asymmetric heritabilities are not uncommon (Falconer, 1989, Frankham, 1991), with lower heritabilities usually being found in the direction associated with increased fitness The most likely explanation for this is that selection has favoured evolution of the trait in a particular direction in the past, so that the gene frequencies are different from 0 5 A trait controlled by additive genes will normally only have equal heritabilities up and down when the average gene frequency is 0 5 This would explain not only the reduced SSD in the up selected lines, but also the only large asymmetry in the realized heritabilities (Table 3) - that between the female up and female down lines The difficulty with this explanation is that it requires a mechanism by which the asymmetry is only seen in females (or is in the opposite direction in males). Since most genes are shared by the two sexes, it is not clear how this sex-limited asymmetry could be achieved

A third possibility is that some component of growth is physically maximized in large flies, and that increasing the genetic value for this component has little or no phenotypic effect. This would allow upward selected males to "catch up" in body size if

they have not yet reached the size where the physical constraint occurs, and would explain the decrease in SSD with increasing body size Non-linear allometry has also been shown to occur among populations of the water strider *Aquarius remiges* (Fairbairn and Preziosi, 1994). In this species females are larger than males, and the regression curve was found to be concave downwards, as in the present study. Fairbairn and Preziosi interpreted this as indicating an upper boundary for size, which places a limit on the changes possible by altering growth trajectories within a species. A lower limit on size would also explain the slight hyperallometry seen at smaller body sizes in both the present study and the water strider data of Fairbairn and Preziosi.

Body size is a complex trait which is probably controlled by a large number of genetically correlated components If SSD is complicated by the differential responses of several of these components, then it might be too optimistic to expect the univariate model to accurately predict the response to selection of a complex trait like SSD for body size (or thorax width). Since Lande's original model is multivariate, it could be argued that the predictions would be more accurate if several important component traits, rather than a single measure such as thorax width, were included in the model.

In previous reports of selection on both sexes simultaneously, some studies (using *Drosophila*) have found a decrease in SSD with both directions of selection (Zeleny, 1921 [eye facet number]; Robertson and Reeve, 1952 [thorax length and wing length]; Partridge and Fowler, 1993 [thorax length]). The only explanation for such a pattern was given by Zeleny (1921), who suggested that inbreeding caused a reduction in SSD in all lines, but this is not surprising given that the selection lines were maintained

from a single mating pair each generation. Alicchio and Palenzona (1971), selecting on wing length, found a large increase in SSD_d when selecting up, and a large decrease when selecting down. Higuet (1991) found the same pattern of response in four of four down lines, and in three of four up lines, when selecting on body weight. In a second line selected for wing length, Robertson and Reeve (1952) found an increase in SSD_l with both directions of selection. There appears to be no consistent pattern to these results, and many of them are suspect due to excessive inbreeding and lack of significance. There is at present no satisfactory explanation for the patterns in SSD change seen in these selection experiments.

The Lande model was generally as good as or better than the Cheverud model at predicting the observed responses to selection in the present experiment. This is clear from Figures 7 and 8, where the Cheverud models consistently overestimate the change in dimorphism. Since reported values for heritabilities and variances are usually not significantly different between the sexes, it could be that the assumption of equal variance and heritability is valid. Since variances can change rapidly from environment to environment, the true values could fluctuate between generations in each sex, but still be roughly equal on average, in which case there is no need to include such differences in the model. One specific prediction of the Cheverud model is that dimorphism will evolve most rapidly when selection is acting on the sex which has the higher genetic variance (Cheverud et al., 1985). For the present study, these are the lines selected for females only. It is clear from Table 4 that the female selection lines did not undergo the greatest changes in SSD.

Many quantitative genetic studies involving body size assume that there are no differences between the sexes in terms of heritabilities and variances, and often pool the data for the two sexes. Since differences in these values could be important, more attention to should be paid to them. Although a difference in heritabilities was found in this study, it could be that the estimated magnitude of this difference is inflated. The half-sib experimental design used here measures only autosomal variance in male offspring, but includes autosomal and X-linked variance in females. This is because paternal half-sisters share a common X chromosome from their father, while paternal half-brothers do not. If a high proportion of the genes for the trait in question lie on the X chromosome, half-brother estimates will underestimate the true heritability in males. The X chromosome makes up about 20% of the genome in *Drosophila*. Cowley et al. (1986), and Cowley and Atchley (1988) found X-linked variance greater than zero in 22 of the 28 morphological traits they examined in *D. melanogaster*. When using intraclass correlations (as in the present experiment), they found that females had higher heritabilities than males for 23 of the 28 traits. The experimental design they used allowed estimation of heritabilities using the method of iterated weighted least-squares. When calculated in this way, heritabilities were nearly equal between the sexes, with males being on average 0.02 higher than females. Nevertheless, in the present study, higher female heritability for thorax length is supported by the fact that realized heritabilities were higher for females in all lines It is likely that the male h^2 estimate in Table 1 is to some extent an underestimate of the true value, but that female h^2 is still greater than male h^2 for this trait. Although Cheverud's dimorphic variance model is not supported by the results of the

present experiment, it is possible that his model is better for traits which show greater variance dimorphism.

In the only other test of the Cheverud model, Rogers and Mukherjee (1992), analyzing data on the rate of change in dimorphism in 3 human traits, concluded that the Cheverud mechanism of stronger response in the sex with the higher heritability was not sufficient to account for the existing relationships seen between dimorphism and body size throughout the primate order. However, their study was based on incomplete data, and several assumptions had to be made to transform the data to a form suitable for analysis

Choice of the correct measurement scale for testing the models depends on knowledge of the underlying genetic mechanism controlling body size, which is at present only poorly understood. Although it might be thought that evidence from phenotypic plasticity studies could be used to determine the most appropriate scale, studies of reaction norms do not always produce consistent results. Sang (1949), found that flies grown in low food conditions tend to be about the same size in each sex, at about one-quarter the body weight of normal females. Bigger, well fed flies get progressively more dimorphic, at least up to a certain size, but the patterns at still larger sizes are less clear. David et al (1994), found that flies became more dimorphic with decreasing body size when the environmental variable being altered was temperature. The genes responsible for changes due to varying environmental conditions are not necessarily the same as those responsible when the animals compared are genetically different For instance, it is known that flies grown at higher temperatures are smaller mainly because of smaller cell size (Alpatov, 1930), whereas those which have been grown at high temperatures for many generations

(and are genetically smaller) are small mainly because of fewer cells (Partridge and Fowler, 1993). The log models were generally better than the arithmetic models for predicting the observed results from the experiment. Since there was no correlation between variance and mean body size, it must be concluded that this superiority in predictive ability is caused by allelic effects in females being expressed as some multiple (>1) of those in males, and that the geometric scale is better than the arithmetic for analyzing results pertaining to the differential response of the sexes to selection on thorax width.

In summary, SSD evolves much more slowly than body size, as predicted by the Lande and Cheverud models Since the log models were better than the linear models, the genetic basis for body size is probably additive only on the geometric scale The Lande models were usually better than the Cheverud models, so that for the trait measured in this experiment, variance and heritability differences between the sexes are not large enough to affect the predicted responses. The models were good at predicting body size changes, but slightly overestimated the responses of the up selected lines. They were not as successful for predicting changes in SSD The Cheverud models overestimated the SSD in large lines and underestimated it in small lines. The Lande models were able to predict the SSD of small lines fairly accurately, but there was a reduction in SSD at large body sizes which was not predicted. Predictions based on the log Lande model were the most accurate, but overestimated the response of body size to upward selection, and predicted an increase rather than a decrease in SSD at large body sizes Clearly, genetic processes not included in the model influence the response of both thorax width and SSD to selection.

REFERENCES.

Alexander, R. D. J., L. Hoogland, R D. Howard, K. M Noonan, and P W. Sherman 1979. Sexual dimorphisms and breeding systems in pinnipeds, ungulates, primates, and humans Pages 402-435 in N. A. Chagnon and W Irons eds. Evolutionary biology and human social behavior: an anthropological perspective. Duxbury, North Scituate, Mass

Alicchio, R. and D. L Palenzona 1971. Changes of sexual dimorphism values in *Drosophila melanogaster*. Bolletino de Zoologia 38: 75-84

Alpatov, W. W. 1930. Phenotypical variation in body and cell size of *Drosophila melanogaster*. Biological Bulletin 58: 85-103

Andersson, M 1994. Sexual Selection. Princeton University Press, Princeton

Ashburner, M. and J N. Thompson, 1978 The laboratory culture of *Drosophila*, in The Genetics and Biology of *Drosophila*, vol. 2a, M. Ashburner and T.R F Wright, eds, Academic Press, London.

Atchley, W., C Gaskins, and D. Anderson. 1976. Statistical properties of ratios. I Empirical results Systematic Zoology 25. 137-148

Becker, W A. 1984. Manual of Quantitative Genetics (4th ed) Academic Enterprises, Pullman, Wash., USA.

Berry, J F. and R Shine. 1980. Sexual dimorphism and sexual selection in turtles (order Testudines). Oecologia 44· 185-191

Bohren, B B , W. G. Hill, and A. Robertson 1966. Some observations on asymmetrical correlated responses to selection. Genetical Research 7 44-57

Bradbury, J. W. and M. B Andersson, eds 1987. Sexual Selection Testing the Alternatives Wiley, Chichester, U K.

Bull, J. J. 1983. Evolution of Sex Determining Mechanisms Menlo Park Benjamin/Cummings.

Cheverud, J. M , M. M. Dow, and W Leutenegger. 1985. The quantitative assessment of phylogenetic constraints in comparative analysis. sexual dimorphism in body weight among primates. Evolution 39: 1335-1341.

Cowley, D E and W. R. Atchley. 1988. Quantitative genetics of *Drosophila melanogaster* II. Heritabilities and genetic correlations between sexes for head and thorax traits Genetics 119. 421-433.

Cowley, D. E, W R. Atchley, and J. J Rutledge. 1986. Quantitative genetics of *Drosophila melanogaster* I. Sexual dimorphism in genetic parameters for wing traits. Genetics 114: 549-566.

Darwin, C 1874 The Descent of Man and Selection in Relation to Sex 2nd ed. John Murray, London

David, J. R., B. Moreteau, J P. Gauthier, G. Petavy, A. Stockel, and A G. Imasheva. 1994. Reaction norms of size characters in relation to growth temperature in *Drosophila melanogaster*: an isofemale lines analysis. Genetics, Selection, Evolution 26: 229-251.

Fairbairn, D. J. and R. F. Preziosi 1994 Sexual selection and the evolution of allometry for sexual size dimorphism in the water strider, *Aquarius remiges*. American Naturalist 144:101-118.

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

Fisher, R. A. 1958. The Genetical Theory of Natural Selection. 2nd ed. Dover, New York.

Frankham, R. 1968. Sex and selection for a quantitative character in *Drosophila*. I. Single-sex selection. Australian Journal of Biological Science 21: 1215-1223.

Frankham, R 1991 Are responses to artificial selection for reproductive fitness characters consistently asymmetrical? Genetical Research 56: 35-42

Ghiselin, M. T 1974 The Economy of Nature and the Evolution of Sex. University of California Press, Berkeley.

Gould, S J and R C Lewontin 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proceedings of the Royal Society of London B 205 581-598.

Higuet, D 1991 Directional selection on body weight and hybrid dysgenesis in *Drosophila melanogaster*. Genetics, Selection, Evolution. 23: 205-219.

Hill, W G 1972. Estimation of realised heritabilities from selection experiments. II. Selection in one direction Biometrics 28: 767-780.

Honek, A. 1993. Intraspecific variation in body size and fecundity in insects: a general relationship. Oikos 66: 483-492.

Kiltie, R. A. 1985. Evolution and function of horns and hornlike organs in ungulates. Biological Journal of the Linnaen Society 24: 299-320

LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution Annual Review of Ecology and Systematics 20 97-117.

Lande, R. 1980. Sexual dimorphism, sexual selection and adaptation in polygenic characters. Evolution 34: 292-307.

---. 1987. Genetic correlations between the sexes in the evolution of sexual dimorphism and mating preferences, in Sexual Selection: Testing the Alternatives, J.W.Bradbury and M.B. Andersson, eds John Wiley and Sons, Chichester, UK.

Leutenegger, W and J. M. Cheverud. 1985. Sexual dimorphism in primates the effects of size. in: Size and Scaling in Primate Biology. ed W L. Jungers, pp 33-50. New York: Plenum Press.

Maria, G. A , K. G. Boldman, and L D van Vleck 1993. A note on heritability estimates for growth traits in male and female Romanov sheep Animal Produce 57: 326-328

Meagher, T. A. 1992. The quantitative genetics of sexual size dimorphism in *Silene latifolia* (Caryophyllaceae). I Genetic variation. Evolution 46. 445-457

Mousseau, T. A. and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59: 181-197.

Partridge, L and K Fowler. 1993 Responses and correlated responses to artificial selection on thorax length in *Drosophila melanogaster*. Evolution 47: 213-226

Payne, R. B 1984. Sexual selection, lek and arena behaviour, and sexual size dimorphism in birds. Ornithological Monographs 33: 1-53

Ralls, K. 1976. Mammals in which females are larger than males Quarterly Review of Biology 51: 245-276.

. 1977. Sexual dimorphism in mammals: avian models and unanswered questions. American Naturalist 111: 917-938.

Rice, W. R 1984. Sex chromosomes and the evolution of sexual dimorphism Evolution 38: 735-742.

Robertson, F. W., and E. C. R. Reeve 1952 Studies in quantitative inheritance. I. The effects of selection of wing and thorax length in *Drosophila melanogaster*. Journal of Genetics 50: 414-448

Roff, D. A. 1992. The Evolution of Life Histories. Theory and Analysis. Chapman and Hall. New York.

Roff, D. A. and T. A. Mousseau. 1987. Quantitative genetics and fitness: lessons from *Drosophila* Heredity 58. 103-118.

Roff, D. A and D J. Fairbairn 1993. The evolution of alternate morphologies: fitness and wing morphology in male sand crickets. Evolution 47: 1572-1584.

Rogers, A. R , and A Mukherjee. 1992. Quantitative genetics of sexual dimorphism in human body size. Evolution 46: 226-234.

Sang, J H. 1949. The ecological determinants of population growth in a *Drosophila* culture. Physiological Zoology 22: 183-210.

Selander, R K. 1972 Sexual selection and dimorphism in birds. in Campbell, ed., Sexual Selection and the Descent of Man, 180-230. Aldine, Chicago.

Shine, R. 1979. Sexual selection and dimorphism in the Amphibia. Copeia 1979: 297-306.

. 1988. The evolution of large body size in females: a critique of Darwin's "fecundity advantage" model. American Naturalist 131: 124-131.

Shine, R. 1994. Sexual size dimorphism in snakes revisited. Copeia 1994: 326-346.

Slatkin, M. 1984. Ecological causes of sexual dimorphism. Evolution 38: 622-630.

Sokal, R. R and F. J. Rohlf. 1981. Biometry: the Principles and Practice of Statistics in Biological Research W. H. Freeman, San Francisco.

Turner, J. R G. 1978 Why male butterflies are non-mimetic: Natural selection, sexual selection, group selection, modification and sieving. Biological Journal of the Linnaen Society 10: 385-432.

Waddington, C H. 1962 New Patterns in Genetics and Development. Columbia University Press, N. Y.

Walker, E. M. and P S. Corbet. 1975. The Odonata of Canada and Alaska. Vol. 3, Anisoptera. University of Toronto Press, Toronto.

Wiklund, C., and B. Karlsson. 1988. Sexual size dimorphism in relation to fecundity in some Swedish satyrid butterflies. American Naturalist 1313: 132-138.

Wilkinson, G. S. 1993. Artificial sexual selection alters allometry in the stalk-eyed fly *Cyrtodiopsis dalmanni* (Diptera: Diopsidae). Genetical Research 62: 213-222.

Wilkinson, G. S., K. Fowler, and L. Partridge. 1990. Resistance of genetic correlation structure to directional selection in *Drosophila melanogaster*. Evolution 44: 1990-2003.

Wright, D. B. 1993. Evolution of sexually dimorphic characters in peccaries (Mammalia, Tayassuidae). Paleobiology, 19: 52-70.

Zeleny, C. 1921. Decrease in sexual dimorphism of bar eye *Drosophila* during the course of selection for low and high facet number. American Naturalist 55: 404-411.

APPENDIX

Table A1 Female generation means for thorax width, replicate #1.

Gen.	Mean thorax width (mm)						
	C	BU	BD	FU	FD	MU	MD
n=	15	60	60	60	60	15	15
0	0.3238	0 3208	0.3193	0.3200	0.3199	0.3363	0.3393
1	0.3173	0.3165	0.3208	0.3332	0.3283	0.3369	0.3326
2	0.3238	0 3330	0.3375	0.3234	0.3210	0.3279	0.3186
3	0.3082	0.3057	0.2965	0.3351	0.3198	0.3192	0.3134
4	0.3175	0.3271	0.3295	0.3138	0.3168	0.3058	0.2898
5	0.3110	0.3210	0.3025	0.3350	0.3215	0.3219	0.3109
6	0 3102	0.3102	0.3192	0.3224	0.3240	0.3128	0.3050
7	0.3337	0.3278	0 2976	0.3247	0.3072	0.3056	0.3229
8	0.3095	0.3106	0.3040	0.3076	0.3024	0.3234	0.2861
9	0.3021	0.3183	0.2959	0.3127	0.2852	0.3183	0.2966
10*	0.3073	0.3294	0.2975	0.3101	0.2918	0.3137	0.2983
11	0.3038	0.3217	0.2934	0.3222	0.2989	0.3169	0.2899
12	0.3088	0.3316	0.3033	0.3156	0.3008	0.3250	0.3190
13	0.3325	0.3463	0.3118	0.3315	0.2984	0.3524	0.2874
14	0.3301	0.3484	0.3063	0.3186	0.3101	0.3255	0.3125
15	0.3011	0.3236	0.2897	0.3104	0.2820	0.3164	0.3069
16	0.3051	0.3398	0.2895	0.3216	0.2751	0.3295	0.2764
17	0.3059	0.3463	0.2813	0.3201	0.2750	0.3176	0.2751
18	0.2992	0.3420	0.2821	0.3037	0.2686	0.3247	0.2815
19	0.3191	0.3302	0.2831	0.3140	0.2784	0.3303	0.2876
20*	0.3153	0.3361	0.2795	0.3253	0.2706	0.3267	0.2820
21	0.3013	0.3451	0.2902	0.3156	0.2886	0.3324	0.2652

* n = 60 for both sexes in both replicates for all lines in generations 10 and 20
Abbreviations as in Table 4

Table A2. Male generation means for thorax width, replicate #1.

Gen.	Mean thorax width (mm)						
	C	BU	BD	FU	FD	MU	MD
n=	15	60	60	15	15	60	60
0	0.2920	0.2866	0.2906	0.2871	0.2872	0.2894	0.2894
1	0.2817	0.2823	0.2866	0.3025	0.3000	0.3017	0.2939
2	0.2973	0.3005	0.3087	0.2888	0.2935	0.2964	0.2932
3	0.2874	0.2824	0.2695	0.3111	0.2953	0.2835	0.2826
4	0.2868	0.2956	0.2976	0.2898	0.2904	0.2744	0.2706
5	0.2890	0.2864	0.2746	0.3043	0.2816	0.2831	0.2851
6	0.2825	0.2819	0.2919	0.2904	0.2926	0.2833	0.2731
7	0.3006	0.2939	0.2688	0.2923	0.2857	0.2801	0.2958
8	0.2760	0.2839	0.2804	0.2744	0.2742	0.2936	0.2609
9	0.2734	0.2903	0.2694	0.2844	0.2613	0.2928	0.2731
10*	0.2802	0.3042	0.2690	0.2800	0.2628	0.2876	0.2709
11	0.2738	0.2901	0.2687	0.2956	0.2663	0.3034	0.2609
12	0.2813	0.3029	0.2793	0.2873	0.2717	0.2896	0.2949
13	0.3023	0.3168	0.2870	0.2940	0.2708	0.3165	0.2694
14	0.2975	0.3157	0.2812	0.2813	0.2766	0.2926	0.2820
15	0.2829	0.2987	0.2670	0.2793	0.2545	0.2883	0.2756
16	0.2749	0.3115	0.2641	0.2913	0.2546	0.2964	0.2542
17	0.2780	0.3176	0.2575	0.2877	0.2576	0.2878	0.2585
18	0.2661	0.3120	0.2571	0.2673	0.2469	0.2966	0.2611
19	0.2856	0.3010	0.2609	0.2859	0.2533	0.3020	0.2599
20*	0.2824	0.3074	0.2611	0.2946	0.2432	0.2987	0.2622
21	0.2785	0.3164	0.2685	0.2850	0.2582	0.3087	0.2453

Abbreviations as in Table 4

Table A3 Female generation means for thorax width, replicate #2

Gen	Mean thorax width (mm)						
	C	BU	BD	FU	FD	MU	MD
n	15	60	60	60	60	15	15
0	0 3154	0 3297	0 3297	0.3270	0 3270	0 3295	0.3300
1	0 3177	0 3136	0 3173	0 3185	0 3147	0 3169	0 3002
2	0 3164	0 3190	0 3220	0 3174	0 3049	0 3175	0.2993
3	0 3050	0 3427	0 3077	0 3352	0 3065	0 3096	0.3011
4	0 3185	0 3113	0 3049	0 3080	0 3145	0 3163	0 3099
5	0 3053	0 3320	0 3140	0 3136	0.3035	0 3178	0 3009
6	0 2976	0.3409	0 2939	0 3228	0 3154	0 3360	0 3102
7	0 3272	0 3206	0 2992	0 3101	0.3014	0.3135	0.3038
8	0 3069	0 3333	0 2955	0 3177	0 2950	0.3136	0 3042
9	0 3099	0 3382	0.2782	0 3336	0 2818	0 2981	0 3062
10*	0 3005	0 3214	0 2785	0 2940	0 2858	0 3126	0 2898
11	0 3026	0 3435	0 3005	0 3163	0 2930	0 3144	0 3092
12	0 3059	0 3483	0 3009	0 3246	0 3025	0 3127	0.3011
13	0 3256	0 3659	0 3018	0.3326	0 3028	0.3274	0 3113
14	0 3185	0 3492	0.2865	0 3204	0 3036	0 3224	0.3102
15	0 3116	0 3506	0 2797	0 3225	0 2814	0 3144	0 2774
16	0 3037	0 3331	0 2609	0 3224	0 2979	0 3208	0 2972
17	0 3132	0 3382	0 2723	0 3289	0.2857	0.3287	0 2695
18	0 2904	0 3332	0 2756	0 3227	0 2819	0 3167	0.2948
19	0 3089	0 3283	0 2527	0.3319	0 2869	0 3119	0 2942
20*	0 2895	0 3233	0 2619	0.3152	0 2749	0.3194	0.2883
21	0 3020	0 3419	0 2675	0 3100	0.2930	0 3227	0.2968

Abbreviations as in Table 4

Table A4 Male generation means for thorax width, replicate #2

Gen	Mean thorax width (mm)						
	C	BU	BD	FU	FD	MU	MD
n=	15	60	60	15	15	60	60
0	0 2939	0.3029	0.3029	0 2977	0 2935	0 3035	0.3035
1	0 2847	0.2835	0 2853	0 2972	0 2903	0 2896	0.2758
2	0 2937	0 2920	0 2914	0 2912	0 2784	0 2927	0.2746
3	0.2706*	0 3070	0 2805	0 3034	0 2774	0 2747	0.2749
4	0 2970	0.2806	0 2719	0 2780	0 2870	0 2905	0.2805
5	0 2735	0 3011	0 2851	0 2818	0 2684	0 2908	0.2739
6	0 2652	0 3077	0 2705	0 2936	0 2811	0 3025	0.2809
7	0.2918	0 2871	0 2743	0 2793	0 2739	0 2887	0.2703
8	0.2754	0.3024	0 2683	0.2863	0 2748	0 2833	0.2786
9	0 2820	0.3074	0 2558	0 3044	0 2534	0 2723	0.2818
10*	0.2740	0 2925	0 2504	0.2676	0 2592	0 2859	0.2589
11	0.2710	0 3143	0 2725	0.2839	0 2729	0 2845	0.2766
12	0 2823	0 3166	0.2690	0 2876	0 2748	0 2820	0.2707
13	0 2974	0 3296	0 2730	0.3061	0 2815	0 3011	0.2779
14	0 2934	0 3144	0 2551	0 2893	0 2727	0 2880	0.2772
15	0 2820	0 3164	0 2531	0.2945	0 2559	0 2883	0.2578
16	0.2845	0 3099	0 2365	0.2917	0 2686	0 2932	0.2717
17	0.2814	0 3096	0 2478	0.2959	0 2627	0 2992	0.2496
18	0 2667	0 3031	0 2484	0.2935	0 2576	0 2857	0.2641
19	0 2772	0 2963	0 2261	0 3003	0 2659	0 2808	0.2638
20*	0.2645	0 2974	0 2352	0 2858	0 2527	0 2888	0.2616
21	0 2760	0 3162	0 2412	0 2841	0 2692	0 2940	0.2648

Abbreviations as in Table 4