INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Genetic Constraints on the Evolution of Sexual Size Dimorphism

Jeff Reeve

A Thesis

in

The Department

of

Biology

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

March 2000

Abstract

Genetic Constraints on the Evolution of Sexual Size Dimorphism

Jeff Reeve, Ph.D.
Concordia University, 2000.

In this thesis, I use an artificial selection experiment and multivariate simulation modelling to study some basic questions concerning the evolution of sexual size dimorphism (SSD). Fecundity selection is often suggested as the main causal factor underlying the prevalence of female-biased SSD, but this assumption has not been empirically tested. I selected female *Drosophila melanogaster* for increased or decreased fecundity for 20 generations, and measured the effect on SSD in three morphological traits. SSD generally increased with selection for increased fecundity, but showed no consistent trend with selection for decreased fecundity. These results support the general hypothesis that SSD can evolve rapidly in response to selection for increased fecundity.

SSD can evolve for a number of reasons. The two main causes are thought to be sexual selection on males, and natural selection favouring different trait optima in the two sexes. Lande (1980a,b) has produced analytical models that can be used for predicting the change in SSD through either of these mechanisms. Although these models are often cited, they have never been adequately tested, either empirically or through simulation modelling. They rely on a large number of simplifying assumptions, and their robustness to violations of these assumptions is largely unknown. In this thesis, I present results from stochastic simulation models designed to test the effects of mutational variance assumptions, finite populations, and finite numbers of loci on the robustness of the

analytical models' predictions. The quality of the predictions depends on the nature of allelic distributions in the original population. If allelic effects are approximately normally distributed, the predictions can be very accurate. If, as is likely, allelic effects have a leptokurtic distribution, Lande's equations underestimate the rate of response and correlated response, and overestimate the time required for the trait means to reach their equilibrium values. Predictions for the magnitude of SSD at equilibrium can be very accurate for weak sexual selection. However, with stronger sexual selection the total response is greater than predicted. The results suggest that genetic correlations constrain both the short-term and long-term evolution of SSD less than predicted by the Lande model.

Acknowledgements

I would like to thank my supervisor Dr. Daphne Fairbairn for her guidance, support, and patience throughout all aspects of this project. Her editorial comments and suggestions were very helpful in reorganizing the doggerel of my early drafts. Dr. Derek Roff provided several useful suggestions regarding the simulations, and commented on an earlier version of the thesis. Dr. Jim Grant also provided comments and suggestions on an earlier version of the thesis. Thanks also to the two external examiners for my thesis defence, Dr. George Gilchrist, and Dr. Paul Joyce.

I have greatly appreciated working with the students and technicians in the Fairbairn lab, both past and present. In particular, I would like to thank Ehab Abouheif and Ian Ferguson for many discussions about various aspects of evolutionary theory that have enhanced my understanding of the topic.

Finally, I would like to thank my parents and Gina Macintyre for their support throughout the thesis.

Table of Contents

List of Tables

Chapter 1.

Chapter 3.

List of Figures

Chapter 1.

Chapter 2.

Chapter 3.

Contributions of Authors

The three chapters in this thesis were all prepared for submission as journal articles. My supervisor, Daphne Fairbairn, is the co-author of the papers from chapters one and three. She contributed to the material for all three papers in the normal supervisory role, and was instrumental in the planning and methodology of chapters one and three. The subject matter of these papers is within the scope of Dr. Fairbairn's NSERC funded research.

Chapter one has been published as:

Reeve, J. P., and D. J. Fairbairn. 1999. Change in sexual size dimorphism as a correlated response to selection on fecundity. *Heredity* 83: 697-706.

Chapter two has been published as:

Reeve, J. P. 2000. Predicting long-term response to selection. *Genetical Research* 75: 83-94.

Chapter three has been submitted to *Evolution* as:

Reeve, J. P., and D. J. Fairbairn. Predicting the evolution of sexual size dimorphism.

General Introduction

A difference in body size between males and females, known as sexual size dimorphism (SSD), is a common characteristic of many animal species. Females are usually the larger sex, but males tend to be larger in some taxa, including most birds and mammals (Andersson, 1994; Fairbairn, 1997). Darwin (1871) suggested that the direction of dimorphism was a result of the different reproductive roles of the two sexes. Selection for increased fecundity favours larger females, since they tend to produce more offspring. Sexual selection, either through female choice or male-male competition, often favours larger males. The net difference between fecundity selection on females and sexual selection on males plays a large part in determining the equilibrium size expected in each sex, and therefore the equilibrium SSD (Arak, 1988). However, reproductive success is not the only component of an organism's life-history that is selectively important. Phenotypes that are expected to have high reproductive success may not have the highest lifetime fitness, because of trade-offs with other components of fitness such as viability (Hedrick and Temeles, 1989; Wikelski and Trillmich, 1997). Viability selection itself may favour different optimal trait values in the two sexes through ecological or physiological mechanisms (Selander, 1972; Ralls, 1977; Shine, 1989). Ultimately, differences between the two sexes in terms of lifetime fitness determine the level of SSD expected at equilibrium.

Most ecological studies assume that the two sexes are at or near their equilibrium (optimal) trait values, and that therefore the degree of SSD observed in a population is usually adaptive. However, SSD may sometimes be maladaptive, at least transiently. The major cause of transiently maladaptive SSD, if it exists, is likely to be genetic correlations

between homologous traits in the two sexes (Lande, 1980b; Roff, 1997, table 7.4). For instance, if the optimal value for a male trait increases, that trait will respond to selection, as will its homologue in females. When the female trait has increased to the point at which selection upwards on males is equal to selection downwards on females (towards their own, unchanged, optimum value), further response in the population will be slow because it must proceed against the sign of the genetic correlation. This may lead to a long period with suboptimal trait values in each sex, but quantitative genetic theory assumes that the optimal value in each sex will eventually be reached, unless the correlation is 1.0 or -1.0 (Lande, 1980b). In Lande's model, if there is only one trait under selection, permanently maladaptive SSD can only occur if the genetic correlations between the sexes are 1.0 or -1.0. At present, there are no empirical data available that can assess how common such large correlations are likely to be. Large body size dimorphisms occur in many taxa, and in both directions. For instance, males of some mammalian orders (e.g. Artiodactyla, Pinnipedia, Primates) can weigh from two to eight times as much as females, while dwarf males may be less than one percent of the female's weight in some marine invertebrates (Andersson, 1994; Fairbairn, 1997). The existence of such extreme SSD would argue that a great deal of divergence between the sexes is possible in at least some species.

The role of constraints in evolution has been the subject of several reviews (e.g. Maynard-Smith et.al., 1985; Loeschcke, 1987; Arnold, 1992; Schlichting and Pigliucci, 1998). Constraints caused by genetic correlations between the sexes can be studied in a manner conceptually similar to the study of constraints caused by genetic correlations amongst traits within a sex. Between-sex constraints are different mainly in that the characters under consideration are not present simultaneously in the same body. In

quantitative genetic studies sex differences are often analyzed as if they were the same traits expressed in two different environments (male and female bodies) (Falconer, 1989).

There are several methods available for studying the evolution of SSD, and each is used to address different types of questions. Artificial selection experiments (Alicchio and Palenzona, 1971; Reeve and Fairbairn, 1996) exert relatively strong selection on laboratory populations, in order to compress evolutionary response into a period of time over which it can easily be observed. The correlated response in one sex to selection for the homologous trait in the other sex can then be measured. Some experiments have imposed evolutionarily unrealistic selection (in form as well as intensity) in order to establish that genetic variation for SSD exists (Bird and Schaffer, 1972; Eisen and Hanrahan, 1972). Comparative studies (e.g. Cheverud et. al., 1985; Webster, 1992; Head, 1995; Abouheif and Fairbairn, 1997) can be used to look for patterns across taxa that may be used to suggest the causal factors responsible for extant distributions of SSD. Fecundity selection and sexual selection have been measured on numerous occasions (e.g. Price, 1984; Moore, 1990; Herrera, 1993), primarily using the multiple regression techniques of Lande and Arnold (1983). Although these studies estimate how selection is acting during a particular episode of the life-cycle, evolutionarily relevant conclusions concerning the adaptive significance of current levels of SSD require estimates of lifetime selection for both sexes (Conner et. al., 1996a,b; Ferguson and Fairbairn, 2000; Preziosi and Fairbairn, *in press*). Theoretical modelling has been used to examine potential causes of SSD and to predict its equilibrium value (Slatkin, 1984; Reiss, 1989; Parker 1992). The trajectories of male and female size as they move from their initial values to those at equilibrium under a constant intensity of sexual selection (Lande, 1980b; Lande and

Arnold, 1985) have also been modelled. The major conclusion from Lande's model is that SSD may take a very long time to reach equilibrium, possibly on the order of millions of generations if the genetic correlation between the sexes is high. This conclusion has been very influential in the SSD literature. It has lead to the assumption that SSD should in general be poorly adapted to environments that are changing at even moderate rates over time (e.g. Rogers and Mukherjee, 1992), and that populations will generally spend long periods of time in Lande's "second phase" of SSD evolution (see Chapter 3), wherein both sexes are at suboptimal body size (e.g. Price and Burley, 1993; Wright, 1993; Price, 1996).

This thesis uses two of these approaches to study the evolution of SSD - an artificial selection experiment, and genetic simulation modelling. Chapter one describes a selection experiment on fecundity in *Drosophila melanogaster*. Its goal is to see if fecundity selection can produce measurable changes in SSD over a small number of generations. Despite the long history of the fecundity selection hypothesis for the evolution of SSD, it has never been tested experimentally. The evidence for this hypothesis rests on the correlation usually found between female body size and fecundity, at least among invertebrates and poikilothermic vertebrates. However, the existence of such a correlation is no guarantee that SSD will change in the expected direction when fecundity is directly selected. An early selection experiment on body size in *Drosophila melanogaster* found the expected change in SSD when directional selection was applied to females, but no change occurred when selection was on males (Alicchio and Palenzona, 1971). A later experiment (Reeve and Fairbairn, 1996) found that the unselected sex tended to have a larger change in body size than the selected sex, so that SSD decreased

with all directions of selection. Such selection experiments are in effect mimicking sexual selection. It is possible that direct selection on fecundity produces greater and more consistent changes in SSD than direct selection on body size. This might occur if the genetic correlation between female size and fecundity was higher than that between male size and female fecundity. In that case, selection on fecundity would be able to more easily target body size genes that are expressed only (or to a greater extent) in females, compared with single-sex selection on body size itself.

The lack of fit between the predicted and observed change in SSD in Reeve and Fairbairn (1996) is likely a result of the highly complex interelationship between the many traits that must combine and interact to produce changes in body size. Traits such as development time, critical weight, growth rate, and post-critical growth period are all important in determining adult size in *Drosophila* (Partridge and Fowler, 1993; de Moed et. al., 1999; Partridge et. al., 1999). Unfortunately, even if all the selectively important traits involved in the regulation of body size were known and measurable, it is unlikely that the parameters required for applying the predictive equations could be measured with enough accuracy to produce reliable estimates of SSD evolution. For this reason, I do not use the response equations to predict changes in SSD in the *Drosophila* experiment. In addition to the problem of missing and unreliable parameter estimates, there is no guarantee that the many assumptions required by the predictive equations will be met in real populations. Of major concern is the fact that the genetic and phenotypic distributions of small populations will vary over time to at least some extent. The effect of such fluctuations on the predictive ability of the equations, which assume constant genetic and phenotypic distributions, is unknown.

In the second part of this thesis (Chapters 2 and 3), I use artificial populations created in stochastic simulation models to eliminate some of these problems. In the simulations, all traits under selection are known, and the parameters can be estimated with great accuracy. In addition, since the distributional properties of the simulated populations are allowed to change over time to reflect changes in the underlying allele frequencies caused by selection, the effect of changes in these distributions can be observed directly. Quantitative genetic models are used in preference to optimality models, since I am interested not only in equilibrium values in each sex, but in the dynamics of how these equilibria are reached.

Chapters two and three each use genetic simulation modelling to test a standard quantitative genetic equation for predicting response to selection. Two types of selection are considered. Viability selection is modelled as frequency-independent selection for an optimal trait value. The fitness of a phenotype declines approximately with the square of the distance from this optimal value. When the population mean trait value is at the optimum, selection is purely stabilizing. With increasing distance from the optimum, selection becomes increasingly directional. Sexual selection is modelled as frequency-dependent selection for increased trait values in males. In frequency-dependent selection, the expected fitness of a male is a function of its location relative to the distribution of male phenotypes present in the population, rather than of its phenotypic value per se, and selection is always purely directional.

Chapter two deals with the general case of selection acting through frequency-independent stabilizing selection. I consider the effect of a shift in the optimal value of one of a set of three genetically correlated traits, where the optimal value of the other two

traits has not changed. This model does not consider sex differences. Therefore, in the simulations, the two sexes are assumed to be genetically identical and experiencing the same type and intensity of selection. These simulations test the effect of constraints caused by selection acting on genetically correlated traits within a sex.

The single-sex, frequency-independent model of chapter two forms the basis of the more complicated model examined in chapter three. In real populations, selection will be acting simultaneously on both sexes. Because of differences in reproductive biology, selection will often favour different trait optima, and have a different intensity in each sex. In chapter 3, I model the evolution of SSD through frequency-dependent sexual selection on a single male trait. This trait, in both males and females, remains under frequency-independent stabilizing selection for its initial value. Therefore, equilibrium size in males is determined by the balance between viability and sexual selection, whereas female equilibrium size is determined solely by viability selection. Chapter three also considers the evolution of SSD through frequency-independent selection, the niche dimorphism hypothesis (Slatkin, 1984; Shine, 1989). Here, there is a change in the optimal (viability selection) value of the single male trait, with the optimal value of the homologous female trait remaining unchanged.

To date, most theoretical quantitative genetic studies of genetic constraints (e.g. Cheverud, 1984; Zeng, 1988; Björklund, 1996; Schluter, 1996) have not dealt with the complications brought about by sex differences. Of those that do, most involve models of sexual selection and the correlation between female choice and the male trait, rather than the evolution of SSD itself. Chapters two and three represent the first use of stochastic simulation modelling to study the evolution of SSD.

Chapter 1. Change in sexual size dimorphism as a correlated response to selection on fecundity.

Female-biased sexual size dimorphism has usually been explained as being due to fecundity selection. Several studies in the laboratory and in the wild have shown that size is often positively correlated with fecundity, and that selection favours increased size in adult females. Despite the widespread belief in fecundity selection as a driving force behind female-biased dimorphism, changes in SSD have never been shown to occur as a result of direct selection on fecundity in the laboratory. In this chapter, I select directly for fecundity in female *Drosophila melanogaster*, and measure the effect on SSD in three morphological characters.

This chapter has been published as:

Reeve, J. P., and D. J. Fairbairn. 1999. Change in sexual size dimorphism as a correlated response to selection on fecundity. Heredity 83: 697-706.

ABSTRACT

Fecundity selection is often suggested as the main causal factor underlying the prevalence of female-biased sexual size dimorphism (SSD), but this assumption has not been empirically tested. We selected female *Drosophila melanogaster* for increased or decreased fecundity (eggs laid over a single 18 h period, between days 5 and 7 posteclosion) for 20 generations, to see what effect this would have on SSD in three morphological traits (thorax width, abdomen width and thorax length). A direct response to fecundity selection was found in the downward direction (16.6%), whereas the response to upward selection (5.7%) was not statistically significant. Significant sex by selection interaction terms in the ANOVAs for thorax width and for abdomen width indicate that the two sexes responded differently. Females usually showed a greater correlated response than males. In lines selected for increased fecundity, the correlated response in females for thorax and abdomen width was greater than the direct response in standard deviation units. SSD generally increased with selection for increased fecundity, but showed no consistent trend with selection for decreased fecundity. These results support the general hypothesis that SSD can evolve rapidly in response to fecundity selection. Selection on fecundity also produced correlated responses in life history traits. Downward selection resulted in flies that had lower viability and longevity, and both directions of selection were associated with an increase in development time.

INTRODUCTION

Fecundity selection is often suggested as the main causal factor underlying the prevalence of female-biased sexual size dimorphism (SSD, defined here as female size / male size) in the animal kingdom (e.g. Darwin, 1874; Williams, 1966; Head, 1995; see

Shine, 1988 for a critique). As there is often a positive correlation between body size and fecundity in ectotherms (Roff, 1992, p. 126), larger females tend to be more fecund, and therefore "favoured by selection". Strictly speaking, this is not selection for large females (*sensu* Lande and Arnold's [1983] multivariate selection gradient method), because any evolutionary increase in body size, in this simplified scenario, is just a correlated response to direct selection acting on fecundity. Making this distinction between the two different targets of selection (body size or fecundity) is potentially important in terms of understanding the evolution of SSD. If selection acts directly on female size, the typically high correlations between sexes for body size (Lande, 1980; Roff, 1997, p. 247; Preziosi and Roff, 1998) would be expected to cause a strong correlated response in the size of males, and very slow evolution of sexual size dimorphism (Lande, 1980). This effect has been demonstrated in a previous experiment using artificial selection on body size in *Drosophila melanogaster* (Reeve and Fairbairn, 1996). Selection acting directly on fecundity would produce SSD through a correlated response in body size and might produce a more rapid divergence in size between the sexes, and therefore more rapid evolution of SSD.

The impact of fecundity selection on SSD will depend critically on the pattern of variances and covariances for fecundity and body size in the two sexes (i.e. the genetic architecture of these two traits). Although males do not express the phenotype fecundity (defined as egg production, not male fertility), they still possess genes for the trait. Sex differences may exist in the genetic correlation between fecundity and body size, both within and between sexes. In addition, the genetic correlation between the sexes for fecundity itself might be lower than the high values commonly found for morphological

traits. Selection acting directly on fecundity in females would produce a more rapid change in SSD than selection acting directly on body size if, for example, fecundity was strongly correlated with body size within females, but weakly correlated with body size in males (i.e. intersexual correlations were weak).

Unfortunately, there are no estimates of the genetic variance-covariance matrix (**G**) for body size and fecundity in males, or of the covariances between sexes for these two traits (the **B** matrix of Lande, 1980). Even if such estimates were available, empirical and theoretical results show that predictions of correlated responses are unreliable (e.g. Bohren et al., 1966; Gromko, 1995; Lascoux, 1997), and there is evidence that reciprocal selection may not always result in symmetrical responses among a set of traits (e.g. Shiotsugu et al., 1997). The hypothesis that SSD can evolve in response to fecundity selection must therefore be tested empirically.

In this paper, we apply artificial selection on fecundity in *D. melanogaster* to see if it is possible, in a small number of generations, to produce a relatively greater change in body size in females than in males. Despite the long history of the fecundity selection hypothesis, we are unaware of any other experiments that select on fecundity and measure relative changes in body size in the two sexes. As there has been much interest recently in the correlated responses seen amongst various life history traits (e.g. Zwann et al., 1992; Nunney, 1996; Shiotsugu et al., 1997), we also look at correlated responses in development time, preadult viability, and adult longevity.

MATERIALS AND METHODS

Selection experiment and direct response

The base population was derived from a cross between four stock strains of *D. melanogaster*. This population was kept in a random mating mass culture for four years at 25 °C, with new generations initiated every two to three weeks from approximately 1000 adults.

Selection on fecundity was maintained for 20 generations, at 25 °C. Two replicate lines were used for upward (U1, U2), and downward (D1, D2) selection, along with two controls (C1, C2) for six lines in total. Each generation, 90 females (five to seven days old - see below) were individually placed in small plastic vials (20 x 50 mm) containing a single randomly chosen male of the same age from the same replicate. Males were included in these vials because their presence is known to enhance egg laying (Bell et al., 1957). Each vial contained a mixture of agar and grape juice, which was coated with a thin film of yeast extract to enhance egg laying. Our measure of fecundity was the number of eggs laid over a single 18 h period, at an age between five and seven days posteclosion. This is the peak time for *Drosophila* at 25 °C. The number of eggs produced during the peak phase early in life is known to be highly correlated with lifetime fecundity (Gowen and Johnson, 1946).

For each line, 30 of the 90 sets of parents were selected. In the C lines, parents were selected randomly. In the D and C lines, vials with 10 or fewer eggs were excluded, to minimize selection for flies that were unhealthy, rather than just at the low end of the "normal" distribution of fecundities. The 30 pairs of selected parents were put in separate 30 x 100 mm plastic vials with commercial *Drosophila* media plus yeast for 24 h, and then

removed. Adult flies for the next generation were collected three days after the start of eclosion. Where possible, five flies of each sex were collected from each of the 30 vials, to minimize the effects of natural selection. These 300 flies were mixed together, and redistributed into 10 vials. From generation three onwards, each vial of flies was transferred to a new vial with fresh yeast two days after the first transfer, which greatly reduced the environmental variation between vials and improved the consistency of the selection response. The following day, nine randomly selected pairs of flies from each of the 10 vials were placed in the grape juice-agar vials, for egg counting the next day.

All manipulation was carried out using ether anaesthesia. Lines were spaced at one-day intervals, in the order D1, C1, U1, D2, C2, U2. All statistical analyses, Figures (except Fig. 1) and Tables refer to counts and measurements for flies at generation 20. Towards the end of the selection experiment, it became apparent that the average number of flies emerging was not equal in all of the replicates. In particular, D1 had many fewer adults per vial. Because many traits in *Drosophila* are correlated with larval density, all generation 20 flies were raised at a density of 25 eggs per vial.

Correlated responses – morphological traits

Thorax width was measured as the distance between the posterior sternopleural bristles on the ventral surface of the thorax. Thorax length was the distance from the tip of the scutellum to the front of the thorax, measured from above, and abdomen width was the width of the third abdominal segment in females and the fourth in males. Each of these was usually the widest point in the respective sexes. All measurements were made using digitizing software attached to a 40 X microscope.

Correlated responses – life history traits

Development time was measured as the time from the midpoint of egg-laying to the midpoint of the observation interval during which adults eclosed. Starting on the afternoon of the eighth day after egg-laying, at intervals of 6, 12 and 6 h, adults from each of 30 vials per replicate were removed, sexed and counted.

Preadult viability (egg to eclosion survivorship) was assessed during the experiment described above. Pupae were counted on day eight in each vial. Adults were collected at regular intervals (see above) until eclosion was complete. For this stage (pupa to adult survivorship), which required anaesthetizing and sexing the flies, flies from all vials within a replicate were collected together, so among-vial effects were not available for calculating standard errors.

Estimates of adult longevity were obtained for each replicate by placing eight newly eclosed flies (four of each sex) in each of 13 vials. Flies were transferred to new vials at three-day intervals, at which time surviving flies were counted. As the experiment progressed and flies died, vials were consolidated, in an attempt to maintain 6-10 flies per vial. Males died at a faster rate than females, so this consolidation included ensuring that at least one male was present in each vial.

Statistical analyses

The fecundity distributions were generally negatively skewed, so were rank-transformed before statistical analysis. For the statistical analyses presented in the Tables, body size data were not transformed before analysis, as there was no indication of non-normality or heteroscedasticity. Outliers were removed before analysis, using the method

given in Sokal and Rohlf (1995, p. 407). Analysis of variance was performed with sex and selection (control, down and up) as fixed factors, and replicate as a random factor nested within selection. Because one of the replicates (D1) was very different from the others in terms of preadult viability, longevity, fecundity, and development time, it is probable that its response resulted from a genetic mechanism quite different from the other replicates, including D2. For this reason, multiple comparisons were carried out using two different data grouping methods: (i) differences among selection types with replicate nested within selection; and (ii) differences among all six replicates. Two-tailed tests were used for all traits except fecundity, where there was an *a priori* expectation of direction of response in relation to the controls. We did not use one-tailed tests on morphological traits, as there is the possibility that trade-offs in resource allocation between growth and fecundity might lead to negative correlated responses, in contrast to the expected positive response.

RESULTS

Direct response

Figure 1 shows the direct response to selection on fecundity. A change in experimental protocol after generation three (see Methods) resulted in a noticeable reduction in variation between generations within replicates. The difference in fecundity between the selected and control lines at generation 20 (Tables 1 and 2) corresponds, on average, to an increase of 5.7% in the U lines, and a decrease of 16.6% in the D lines. Analysis of variance showed a significant effect of selection ($F_{2,3} = 90.5$, $P = 0.002$), but not of replicate ($F_{3,534} = 0.56$, $P = 0.64$). Nested posthoc one-tailed P-values, corrected for multiple comparisons, were: C vs. U (0.152), C vs. D (0.024) and D vs. U (< 0.0005)

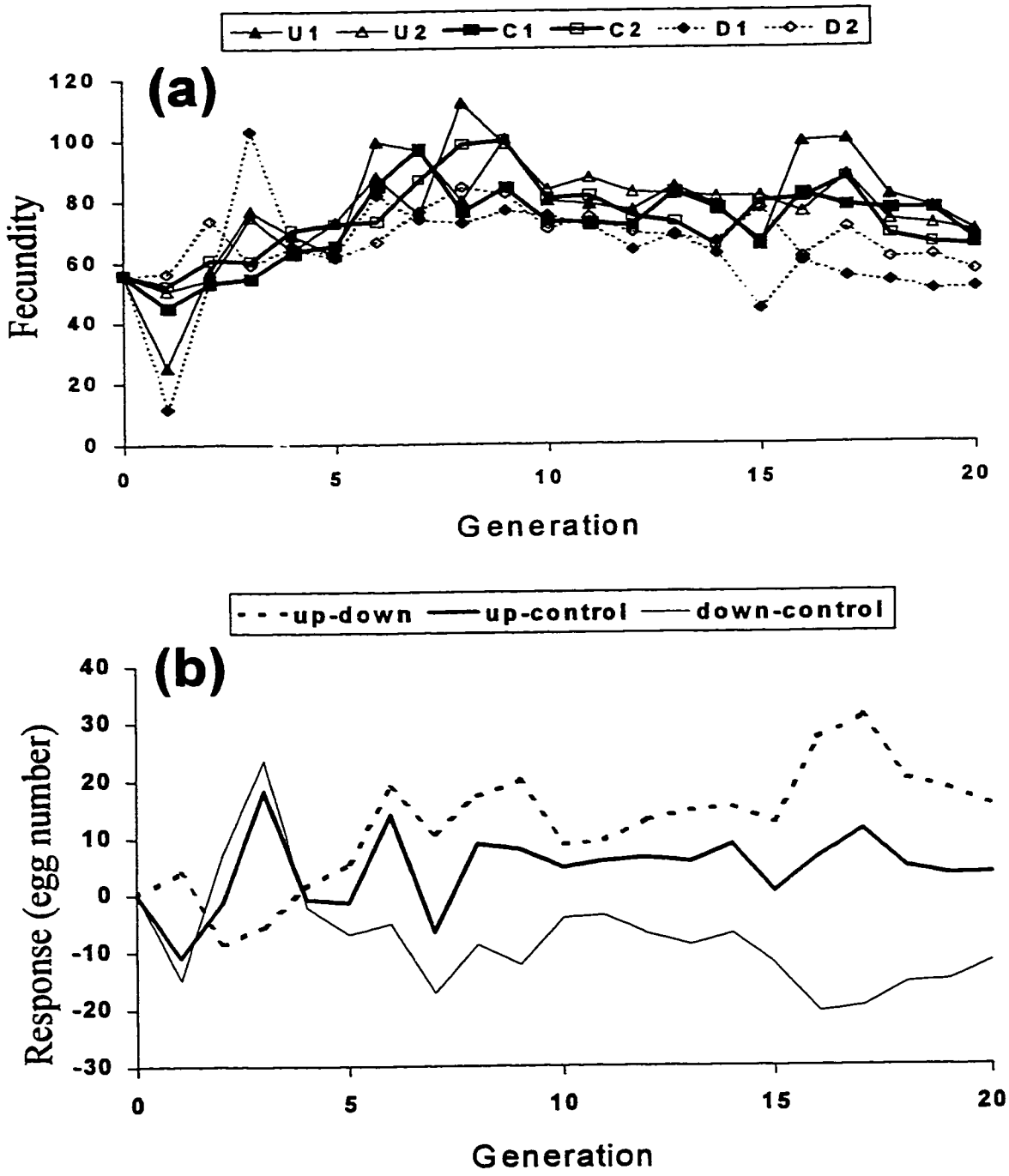

Fig. 1. Direct response to selection for increased or decreased fecundity in *Drosophila melanogaster*. (a) Response in each replicate. (b) Relative responses in the mean value of the two replicates per selection regime.

Table 1. Generation 20 means (and SE) for fecundity and three morphological traits in *Drosophila melanogaster*, with data from the two replicates pooled for each selection type. Superscripts indicate significant differences between selection types, after correcting for the appropriate number of multiple comparisons, using ANOVA with replicate nested within line. Sexes analysed separately.

Sex	Line	Fecundity	Thorax width	Abdomen width	Thorax length
Females	C	66.3 ^{a} (1.01)	0.698 ^{a} (0.001)	1.068 ^{a} (0.004)	1.044 ^{a} (0.002)
	D	55.3 ^{b} (1.32)	0.697 ^{a} (0.001)	1.080 ^{a} (0.004)	1.036 ^{a} (0.002)
	U	70.1 ^{a} (1.00)	0.712 ^{b} (0.001)	1.115 ^{b} (0.004)	1.054 ^{a} (0.002)
Males	C	—	0.630 ^{a} (0.001)	0.762 ^{a} (0.002)	0.901 ^{a} (0.002)
	D	—	0.631 ^{a} (0.001)	0.777 ^{a} (0.002)	0.893 ^{a} (0.002)
	U	—	0.639 ^{a} (0.001)	0.785 ^{a} (0.002)	0.905 ^{a} (0.002)

Table 2. Generation 20 means (and SE) for fecundity and three morphological traits in *Drosophila melanogaster*, for the six replicates. Superscripts indicate homogenous subsets using Tukey's HSD test. Sexes analysed separately.

Sex	Line	Fecundity	Thorax width	Abdomen width	Thorax length
Females	C1	67.3 ^{a} (1.49)	0.695 ^{a} (0.002)	1.070 ^{a} (0.006)	1.041 ^{a} (0.002)
	C2	65.2 ^{a} (1.37)	0.701 ^{a} (0.002)	1.066 ^{a} (0.006)	1.047 ^{a,b} (0.002)
	D1	52.6 ^{b} (2.28)	0.698 ^{a} (0.002)	1.078 ^{a} (0.006)	1.043 ^{a} (0.003)
	D2	58.1 ^{b} (1.30)	0.697 ^{a} (0.002)	1.081 ^{a} (0.005)	1.029 ^{c} (0.003)
	U1	71.0 ^{a} (1.08)	0.713 ^{b} (0.002)	1.117 ^{b} (0.005)	1.055 ^{b} (0.002)
	U2	69.2 ^{a} (1.61)	0.711 ^{b} (0.002)	1.112 ^{b} (0.005)	1.053 ^{b} (0.002)
Males	C1	—	0.628 ^{a} (0.001)	0.758 ^{a} (0.003)	0.900 ^{a,b} (0.002)
	C2	—	0.631 ^{a} (0.002)	0.766 ^{a,b} (0.003)	0.902 ^{a,b} (0.002)
	D1	—	0.634 ^{a,b} (0.002)	0.773 ^{b,c} (0.003)	0.897 ^{a,c} (0.003)
	D2	—	0.629 ^{a} (0.001)	0.781 ^{c,d} (0.003)	0.888 ^{c} (0.002)
	U1	—	0.638 ^{b,c} (0.002)	0.786 ^{d} (0.003)	0.907 ^{b} (0.002)
	U2	—	0.641 ^{c} (0.002)	0.783 ^{c,d} (0.003)	0.904 ^{a,b} (0.002)

(Table 1). Although the response in the U lines was not statistically significant, the fact that the U lines produced, on average, more eggs than the C lines in every generation from the eighth to the twentieth would indicate that a small but real response did take place. Realized heritabilities (total response / total selection differential), calculated from data in generations four through twenty, were 0.055, 0.036, 0.018 and 0.017 for lines D1, D2, U1 and U2, respectively.

Correlated responses of morphological traits

Correlated responses of the three morphological traits are also shown in Tables 1 and 2. In general, the U lines increased in size, whereas the response in the D lines was inconsistent. For thorax width, there was a significant effect of selection ($F_{2,3}$=14.9, P = 0.028), sex ($F_{1,1030}$ = 4866, $P < 0.0005$), replicate ($F_{3,1030}$ = 4.1, P = 0.006), and sex by selection interaction ($F_{2,1030}$ = 3.75, P = 0.024). The interaction is caused by an increase in SSD in the U lines, and a slight decrease in the D lines (Table 1 and Fig. 2). The up-selected lines had wider thoraxes than did the controls and down lines, but this difference was only significant in females. Although the lines were not significantly different in males, it can be seen that there is a trend for the U line males to be bigger than the others when the six replicates are considered separately (Table 2).

For abdomen width, there was a significant effect of selection ($F_{2,3}$=74.9, P = 0.003), sex ($F_{1,1034}$ = 15,911, $P < 0.0005$), and sex by selection interaction ($F_{2,1034}$ = 12.1, $P < 0.0005$), but not replicate ($F_{3,1034}$ = 0.87, P = 0.46). As with thorax width, the interaction term can be seen as an increase in SSD in the U lines, and a decrease in the D lines (Table 1 and Fig. 2), but here the effect was more pronounced. In females, the up-

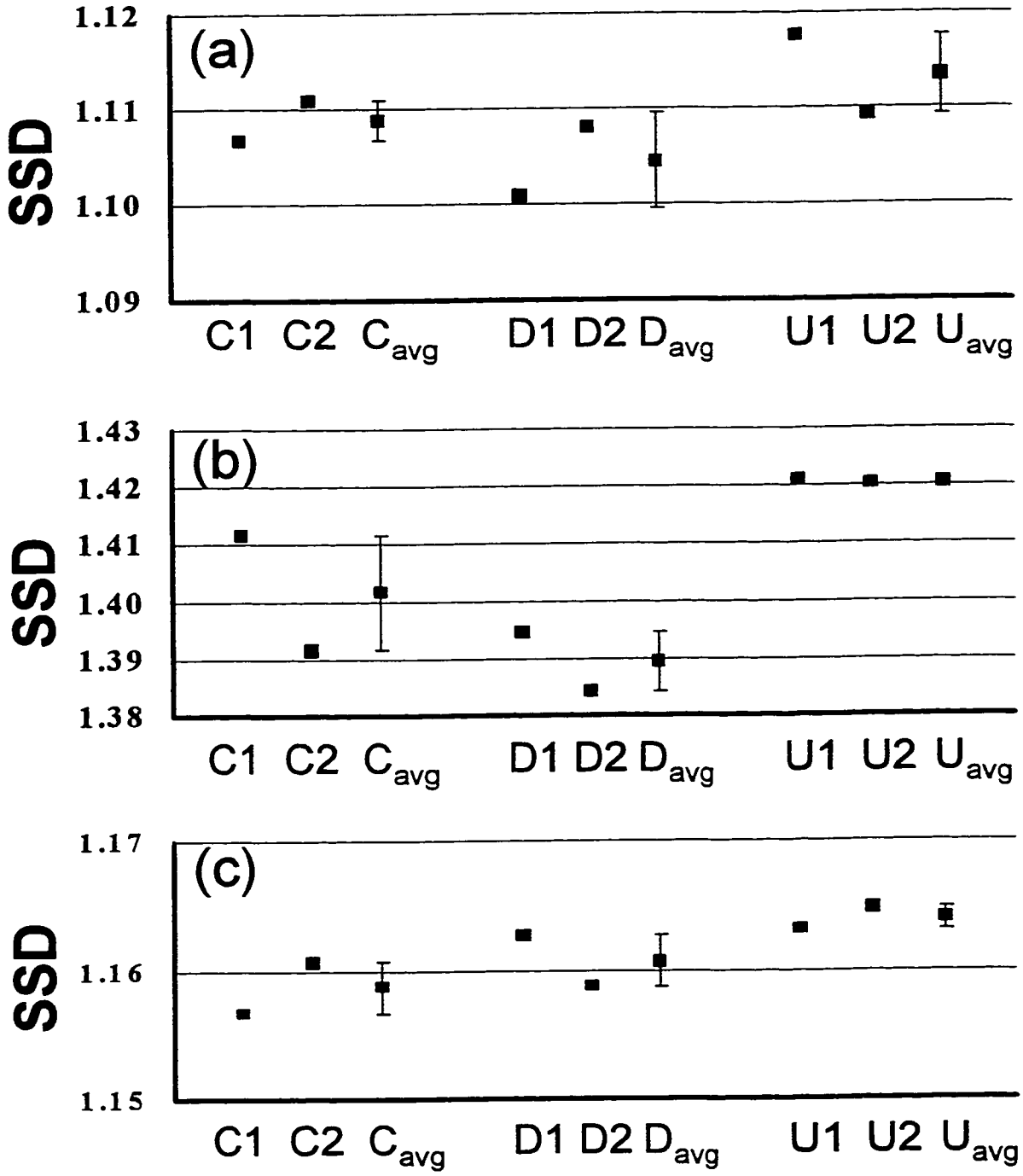

Fig. 2. Sexual size dimorphism (female size / male size) in each replicate of *Drosophila melanogaster*, and the mean ± 1 SE. (a) Thorax width. (b) Abdomen width. (c) Thorax length.

selected lines had significantly wider abdomens than did the controls and down lines, which were not significantly different from each other. In males, none of the lines was significantly different, but as with thorax length, a trend for larger abdomens in the U lines can be seen when the six lines are analysed separately (Table 2).

Thorax length showed a significant effect of sex ($F_{1,1033}$=11,333, $P < 0.0005$) and replicate ($F_{3,1033}$=9.39, $P < 0.0005$), but not for selection ($F_{2,3} = 4.64$, $P = 0.121$), or sex by selection interaction ($F_{2,1033} = 2.18$, $P = 0.114$). Although the effect of selection was not significant, the trends were similar to those seen in thorax width. In both sexes, the U lines had longer thoraxes than the C lines, which in turn were longer than the D lines. SSD increased not only in the U lines (Tables 1 and 2, and Fig. 2), but also in the D lines, as opposed to the decrease in SSD seen in the D lines for both thorax and abdomen width.

The difference between the selected and control lines can be summarized in terms of generation 20 control standard deviation units as follows. For the up lines, these differences were: fecundity (0.25), thorax width (0.81 in females, 0.75 in males), abdomen width (0.87, 0.82) and thorax length (0.50, 0.20). Thus, selection for increased fecundity produced larger correlated responses in females than in males for all three morphological traits. Furthermore, the size increase in females was greater than the direct response of fecundity. The correlated responses to selection for decreased fecundity were not as strong, and were larger in males: fecundity (-0.79), thorax width (-0.08, 0.12), abdomen width (0.22, 0.54) and thorax length (-0.37, -0.42).

SSD can be measured as either a ratio or a difference. In the absence of detailed knowledge about the genetic basis for variation in body size, there is no *a priori* reason for preferring one measure over the other. For this reason, all of the morphological data were

re-analysed after log-transformation. A significant sex by selection interaction in the ANOVA when using raw data indicates that the difference in size between the two sexes has changed. With log-transformed data, a significant interaction indicates that a change in size ratio has occurred. In all but one case, the statistical conclusions were the same as when raw data were used. The only difference was in the analysis of thorax width, when the interaction term was no longer significant ($F_{2,1030} = 2.58$, $P = 0.077$).

Correlated responses of life history traits

Table 3 shows the corrected significance levels for all pairwise comparisons between development times for all replicates, with sexes analysed separately. There was a tendency for development time to increase in both the up- and down-selected lines (Table 3 and Fig. 3). As the distributions were not normal, and differed markedly between the two D replicates, comparisons were made using a Kolmogorov-Smirnov two-sample test for differences between cumulative frequency distributions.

Egg-pupa survivorship (Fig. 4) showed no effect of selection ($F_{2,3}$=3.0, $P = 0.19$), apparently because of the large effect of replicate within selection ($F_{3,174}$=12.9, $P <$ 0.0005). If the data are analysed as six replicates, the D1 and D2 lines are significantly different from all other replicates (including each other) using a Tukey post-hoc test. The method of data collection does not allow statistical analysis of differences in pupa-adult survivorship. However, from Fig. 4, it would appear that the reduction in egg-adult viability in the D lines is mostly caused by increased larval, rather than pupal, mortality.

Gehan's generalized Wilcoxon's test (SPSS 7.5 Advanced Statistics Manual, p.269) was used to compare the adult survivorship curves of the six replicates, separately for each

Table 3. Significance levels for amongst-replicate multiple-comparisons of development time at generation 20 in *Drosophila melanogaster*. Female values are above the diagonal, male values below.

	C1	C2	D1	D2	U1	U2
C1	---		**		*	
C2		---	**		**	**
D1	**	**	---	**	**	**
D2	**		**	---	**	
U1	**		**		---	
U2	**		**	*		---

*$P \leq 0.05$, **$P \leq 0.01$.

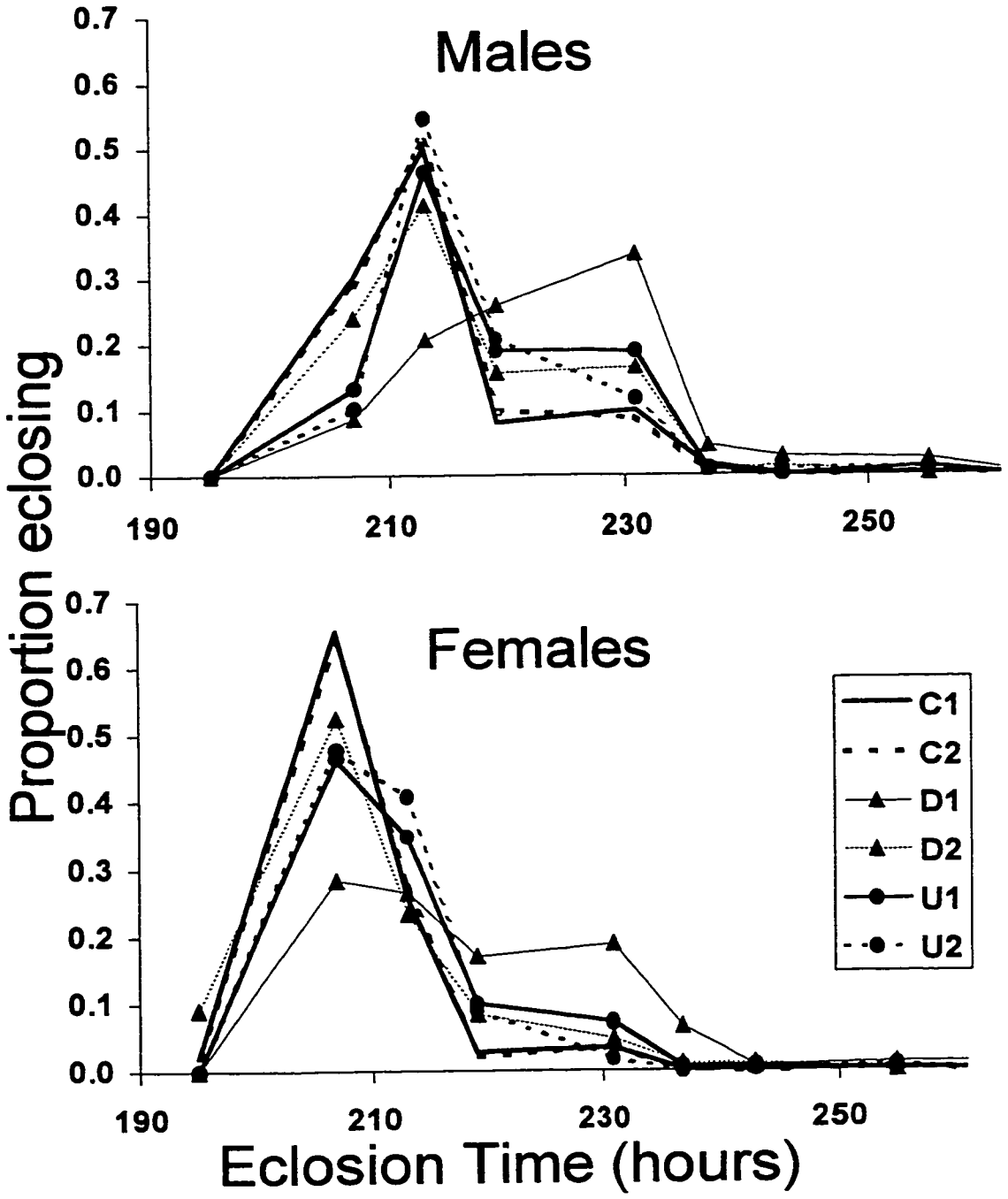

Fig. 3. Distribution of development times (hours from egg to eclosion) in the six replicates of *Drosophila melanogaster*. The proportion of the total flies within a replicate that eclosed over the observation interval is shown on the y-axis.

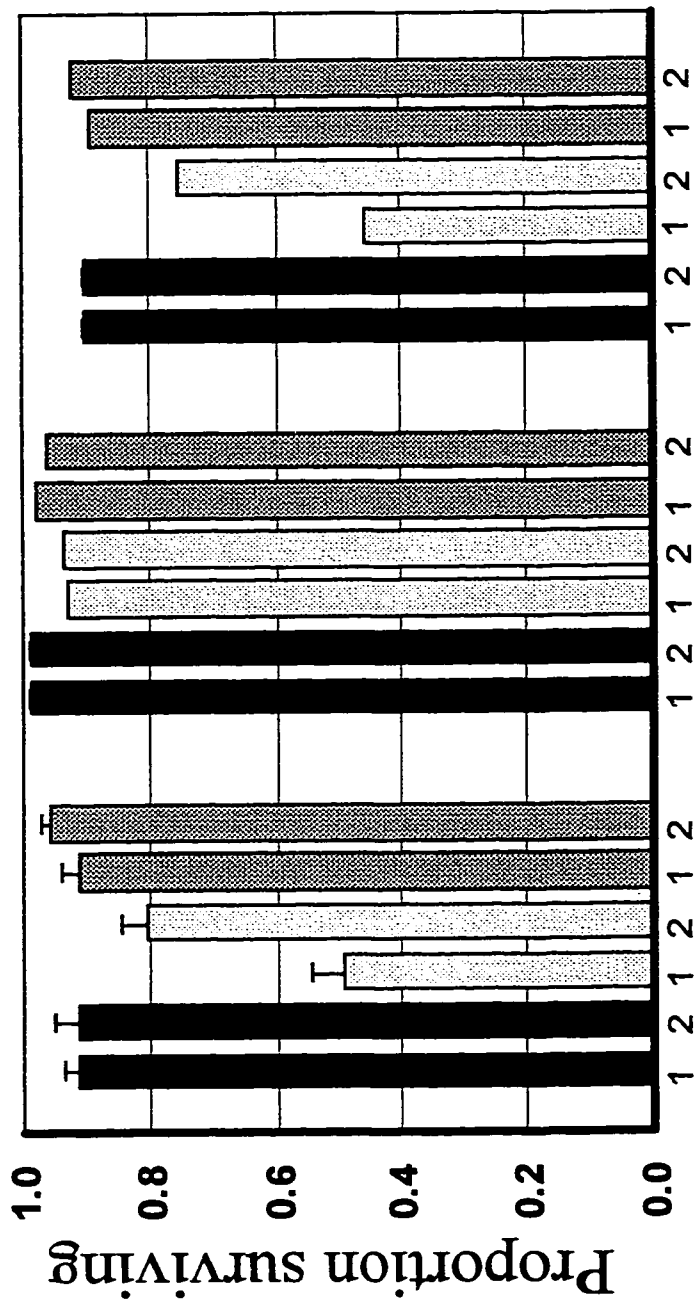

Fig. 4. Proportion of individuals surviving the egg-pupa and pupa-adult stages, and the combined survival proportion over both stages. Error bars indicate + 1 SE, estimated only for egg-pupal survival (see text).

sex (Fig. 5). In females, the D1 line was significantly different from all other replicates (corrected $P < 0.003$ for all pairwise comparisons), which were not different from each other. For males, there were no differences among the six replicates.

DISCUSSION

Selection on fecundity produced a significant decrease in the D lines, and an increase in the U lines that was not significant. Although the response in the U lines was small, it appears that selection did produce genetic changes, because some of the correlated responses were significant. These correlated responses (C = D < U for thorax width and abdomen width) were only significant in females, and were larger than the direct response of fecundity, in terms of standard deviation units. Despite this, there was a tendency in both sexes for an increase in all three morphological traits in response to selection for increased fecundity (Tables 1 and 2). Downward selection produced smaller and directionally less consistent changes in body size measurements. Significant sex by selection interaction terms in the ANOVAs for thorax width and abdomen width were produced mostly by greater increases in female size in the U lines. These results indicate that fecundity selection is potentially capable of producing fairly rapid changes in SSD.

Because fecundity is highly correlated with fitness, it might be expected that past selection will have depleted most of the genetic variance for the trait. However, numerous studies have found significant heritability for fecundity, both in *Drosophila* (e.g. Tait and Prabhu, 1970), and in other invertebrates (see Roff, 1992, p. 360). There have been relatively few selection experiments on fecundity in invertebrates. Although responses are usually small (e.g. Narain et al., 1962; Richardson and Kojima, 1965, both with

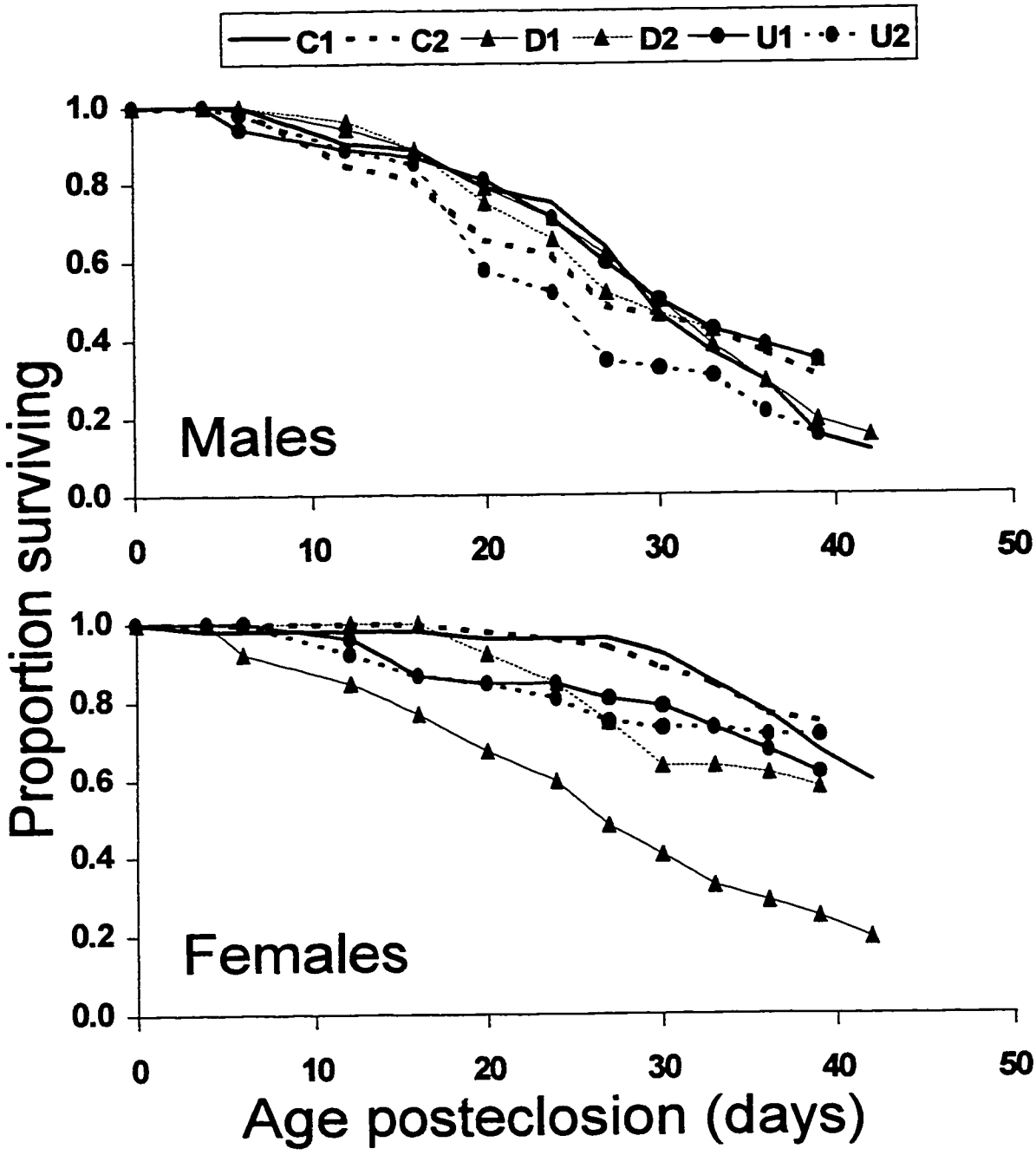

Fig. 5. Adult survivorship curves for male and female *Drosophila melanogaster*.

Drosophila), Orozco and Bell (1974), using *Tribolium castaneum*, were able to quadruple fecundity after just 20 generations. In addition, evidence from agriculture clearly shows that it is possible in domestic fowl to increase fecundity substantially (e.g. Gowe and Fairfull, 1990). Therefore, fecundity can be treated in the same way as other quantitative traits, and should respond to selection in either direction. In natural populations, fecundity is probably under balancing selection because of trade-offs with other fitness-related traits. In our study, the only evidence for such a trade-off was in the slight increase in development time found in the selection lines. The crowded conditions in the base population from which the lines were derived probably selected for rapid development time. Relaxation of this selection pressure in our lines allowed the increase in development time necessary to produce increases in body size and fecundity seen in the U lines. The U lines did not suffer from increased preadult or adult mortality, although our direct response may have been too small to produce significant detrimental correlated responses. A count of fecundity at day 28 (results not shown) in the six replicates gave no indication that the U lines had reduced late-life fecundity relative to the controls.

There is evidence that abdomen size is the component of body size that responds most strongly to selection for higher fecundity in insects (Wickmann and Karlsson, 1989; Preziosi et al., 1996). This is generally considered to result from space limitations, either for eggs or for storage of resources to produce eggs. In female *Drosophila*, the abdomen is thought to be space-limiting in both these respects (Robertson, 1957). The fact that a sex by line interaction term in the ANOVA was found for thorax width shows that the change in dimorphism was not solely caused by a secondary increase in size caused by swelling of the abdomen from extra eggs, as might be argued if the interaction had only

been found in the analysis of abdomen width.

The large direct response and unusual correlated responses in life-history traits in line D1 probably resulted from inadvertent selection for unhealthy flies, rather than for flies with low values for "fecundity genes". The patterns seen in development time for this line (increased development time, increased variance, flatter distribution) are similar to those seen in normal lines that have been grown in crowded conditions (Botella et al., 1985). However, D1 flies were unusual in that they were not smaller than the controls, as would be the case with flies grown in crowded environments. It is interesting that the increased adult mortality in line D1 was restricted to females. Sex differences in longevity are known to be very environment-specific (e.g. Zwann et al., 1992), and either sex may live longer under any particular set of experimental conditions. Chi-squared tests failed to show any significant reduction in the female:male sex-ratio of eclosing flies in this line, so the sex-specific detrimental effects seem to have been limited to the adult stage.

Many studies have examined the pattern of genetic and phenotypic relationships between body size and various life history traits in *Drosophila*. When development time is increased, either genetically or environmentally, body size increases, and this often results in increased fecundity. Depending on the base population used, it may be more difficult to select for decreased development time, but Nunney (1996), for instance, found that this resulted in a reduction of body size, and a sharp decrease in fecundity. SSD in body weight was not changed by this selection. Likewise, many studies have found correlated changes in development time when selecting on body size. Partridge and Fowler (1993) found that in lines selected for increased body size, development time increased. Reeve and Fairbairn (1996) found a positive correlation between body size and development time

in flies selected for small or large thorax width.

Our U lines are consistent with these patterns, but the D lines were unusual in that decreased fecundity was associated with an increase in development time and even a slight increase in abdomen width. Although it is possible that line D2 suffered some of the same problems in terms of general health as line D1, it is also possible that selection for decreased fecundity targetted genes that specifically broke down the correlation between fecundity and the other traits. A similar asymmetrical set of correlated responses has been found by Shiotsugu et al. (1997) in *D. melanogaster*. In lines selected either for longevity or crowding tolerance, urea resistance evolved, whereas selection for urea resistance did not produce a correlated response in either longevity or crowding tolerance. Although there is an underlying causal relationship between development time/body size and fecundity, it may be that there are also genes for fecundity that have sex-specific effects on body size. If selection is imposed on size or development time, the selection intensity on these sex-specific loci will not be as strong as during fecundity selection, and the general pattern found between the traits will be better maintained.

Both empirical results (Reeve and Fairbairn, 1996) and theoretical considerations (Lande, 1980) suggest that SSD may be difficult to alter through single-sex selection on body size. Our results show that significant changes in morphological traits may occur through fecundity selection, even when the direct response in fecundity is small. In general, correlated responses may be greater than the direct response when the heritability of the selected trait is low and the correlation between traits is high. A smaller correlated response in males than in females, as in this experiment, would indicate that the genetic correlation between fecundity and size in males is lower than the same correlation within

females. The present experiment supports the idea that fecundity selection may be a mechanism through which rapid changes in SSD are possible, by providing more direct access to genes with sex-specific effects on body size than is possible through single-sex selection on body size itself.

Chapter 2. Predicting long-term response to selection

Quantitative genetic formulae exist for predicting response to selection in a variety of different settings. The equations for predicting change in sexual size dimorphism are based on simpler equations that assume no sex differences. In this chapter, I use a stochastic genetic simulation model to test these simpler equations for predicting response to selection under a model of multivariate Gaussian selection. I examine a case where there has been a shift in the optimal value of one of a set of three genetically correlated traits.

This chapter has been published as:

Reeve, J. P. 2000. Predicting long-term response to selection. Genetical Research 75: 83-94.

ABSTRACT

Lande's (1980a) equation for predicting the response of trait means to a shift in optimal trait values is tested using a stochastic simulation model. The simulated population is finite, and each individual has a finite number of loci. Therefore, selection may cause allele frequencies and distributions to change over time. Since the equation assumes constant genetic parameters, the degree to which such allelic changes affect predictions can be examined. Predictions are based only on information available at generation zero of directional selection. The quality of the predictions depends on the nature of allelic distributions in the original population. If allelic effects are approximately normally distributed, as assumed in Lande's (1975) Gaussian approximation to the continuum-of-alleles model, the predictions are very accurate, despite small changes in the G matrix. If allelic effects have a leptokurtic distribution, as is likely in Turelli's (1984) "house-of-cards" approximation, the equation underestimates the rate of response and correlated response, and overestimates the time required for the trait means to reach their equilibrium values. Models with biallelic loci have limits as to the amount of trait divergence possible, since only two allelic values are available at each of a finite set of loci. If the new optimal trait values lie within these limits, predictions are good. If not, singularity in the G matrix results in suboptimal equilibria, despite the presence of genetic variance for each individual trait.

INTRODUCTION

Understanding the dynamics of phenotypic evolution is important, not only for predicting how traits should respond to selection, but for knowing how much can be

assumed about past selective forces, given present day trait distributions. Selection experiments have added greatly to our understanding of short-term response, and the results have been, for the most part, consistent with theoretical expectations (Falconer, 1989; Roff, 1997). Patterns of long-term evolutionary change must be studied primarily using non-experimental methods, given the difficulties associated with collecting suitable data.

There is a large body of theoretical work on long-term selection, but most of this concerns mutation-stabilizing selection balance (e.g. Lande, 1975; Turelli, 1984; Barton, 1986; Keightley and Hill, 1988; Burger et al., 1989). These models assume that the population's mean phenotype is already at or near the optimum, and are used primarily for predicting how much genetic variance can be maintained at equilibrium, given various assumptions concerning genetic details. Most directional selection theory is concerned with truncation selection, as used in laboratory experiments (e.g. Robertson, 1970; Bulmer, 1980; Hill, 1982; Keightley and Hill, 1987), and has therefore focussed mainly on short-term responses in small populations. Here, I use a stochastic model to simulate long-term response in finite populations undergoing directional selection to a new set of optimal trait values.

The basic equation for predicting response to a single generation of directional selection is

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

where R is the response in the trait mean, h^2 is the narrow-sense heritability, and S is the selection differential. The multivariate version of equation (1) is

$$\Delta \bar{\mathbf{Z}} = \mathbf{G}\beta \qquad (2)$$

(Lande, 1979), where $\Delta\bar{\mathbf{Z}}$ is a vector of changes in trait means, **G** is the genetic variance-covariance matrix, and β is the selection gradient, often written as the product of the inverse of the phenotypic variance-covariance matrix ($\mathbf{P}^{-1}$) and the vector of selection differentials (**s**).

Extending equation (2) to more than one generation of selection presents two distinct problems. The first is that **G** must be assumed to remain constant over time. How likely this is remains controversial, and empirical findings are equivocal (Shaw et al., 1995). The second problem is that directional selection (β) is unlikely to continue at constant intensity for long periods of time in natural populations. Even in experiments using truncation selection, the force of artificial directional selection is likely to be opposed by natural selection, acting either on the selected or on correlated traits (Lande and Arnold, 1983; Zeng and Hill, 1986; Hill and Keightley, 1988). There is strong evidence that stabilizing selection for intermediate trait values is common in nature (Endler, 1986). Therefore, it is of some interest to investigate the predictive ability of equations that model directional selection as a shift in the optimal values of a set of traits, under multivariate Gaussian selection. The standard equation for shifted optima (Lande 1980a) is

$$\Delta\bar{\mathbf{Z}} = \mathbf{G}(\mathbf{W}+\mathbf{P})^{-1}(\theta-\bar{\mathbf{Z}}) \qquad (3)$$

where **W** is a symmetrical matrix, the diagonal elements being the strength of stabilizing selection acting on each trait (large values = weak selection) and the off-diagonal elements a measure of the strength of correlational selection. The superscript $^{-1}$ indicates matrix inversion, and θ is a column vector of trait optima. In modelling evolution with this equation, it is assumed that an environmental change has brought about a change in θ,

causing directional selection until the traits have evolved to their joint optima. Therefore, the strength of directional selection decreases as $\bar{\mathbf{z}}$ approaches θ but the strength of stabilizing selection ($\mathbf{W}$, the curvature of the fitness surface) remains constant. Hereafter, I will refer to equation (3) as "peak-shift" selection, since the fitness optimum has been shifted to a new location. This should not be confused with the use of the term to describe the shift of a population's genotype from one fitness peak to another in speciation through genetic drift.

Equation (3) still requires a number of assumptions, the most important of which are multivariate normality of genotypic and phenotypic trait values in both current and descendant populations, and constant $\mathbf{G}$ and $\mathbf{P}$ matrices. These assumptions will be violated to some extent in finite populations with finite numbers of loci. The consequences of such violations are studied here using stochastic simulations.

In this study I use simulated populations, subject to the laws of Mendelian inheritance, to investigate the accuracy of equation (3), given various assumptions about the genetic details. The trajectories of the simulated populations' trait means are compared to predictions from equation (3) that are based solely on information available at generation zero of directional selection. Changes in the variance, skew, and kurtosis of the distribution of genotypic values are compared to those found or expected in previous models.

It is well established that the level of genetic variance that can be maintained by mutation-stabilizing selection balance (with or without genetic drift) depends on assumptions made about the distribution of mutational effects at each locus (Turelli, 1984). These assumptions have also been shown to be important in terms of the response

expected when an equilibrium population is subjected to exponential directional selection (Burger, 1993) of the form

$$w(z) = e^{(sz)} \tag{4}$$

where w(z) is the mean fitness of individuals with phenotype z, and s is the strength of directional selection. Therefore, it follows that the accuracy of equation (3) should also depend on the genetic details of the starting population. This section briefly describes the three genetic models that will be simulated in this paper. For a thorough review, see Bulmer (1989).

Models of mutation-selection balance can be classified into two groups – those that assume the mutational (and therefore allelic) effects are continuously distributed, and those that assume effects are discrete and finite. Models of the first type are generally based on the continuum-of-alleles model of Crow and Kimura (1964). This assumes an effectively infinite number of alleles at each locus, producing a continuous distribution of effects. Lande (1975) extended this model to multiple loci, and developed a formula for the equilibrium variance now known as the Gaussian approximation to the continuum-of-alleles model. This assumes that mutational effects, α, are normally distributed at each locus, and that the variance of these effects is small compared to the standing per locus allelic variance ($\alpha^2 << \sigma^2_a$). The Gaussian approximation requires (Burger et al., 1989)

$$\alpha^2 \leq 4\mu V_s \tag{5}$$

where μ is the haploid per locus mutation rate, and $V_s = \omega^2$ (the strength of stabilizing selection on each character, equivalent to the diagonal elements of **W** in equation 3) + σ^2_E (the environmental variance).

Lande argued that under these conditions, mutation-selection balance could

maintain levels of genetic variance consistent with those seen in natural populations. Turelli (1984), showed that maintaining observed heritabilities under Lande's assumptions would require per locus mutation rates far in excess of what is usually thought to be realistic. He proposed an alternative formula for the equilibrium genetic variance, called the "house-of-cards" (HC), or "rare allele" approximation. This applies when the variance of mutational effects at each locus is large compared to the standing allelic variance ($\alpha^2 >> \sigma^2_a$), and requires (Turelli, 1984)

$$\alpha^2 \geq 20\mu V_s \qquad (6).$$

This causes mutational effects to swamp the existing variance at each locus. The net effect is that most genetic variance is maintained by small numbers of mutant alleles at each locus, each of large effect. This tends to produce highly leptokurtic allelic distributions.

In the second type of model, it is assumed that only a small number of allelic values are possible at each locus, with mutational effects limited to moving from one value to another. The first such models assumed two possible alleles (Latter, 1960; Bulmer, 1972, 1980). These have since been extended to include three (Turelli, 1984; Houle, 1989), and five (Slatkin, 1987a) alleles. Since the allelic values are fixed, traits are restricted to a finite range of genotypic values if the number of loci is finite. When multiple traits are considered, there are also limits on the divergence between traits. With three or more traits, these limits are determined by the eigenstructure of the **G** matrix, rather than by the correlations between pairs of traits.

In this paper, three main types of initial population are simulated. The first two are continuum-of-alleles models that have either normal (= Gaussian) or leptokurtic (= HC)

allelic distributions at equilibrium. Both of these assume normally distributed mutational effects, and will be referred to as "continuous effects" populations. The third population type has two discrete values (-0.5 and 0.5) per allele, with equal forward and backward mutation rates, and will be referred to as "discrete effects" populations. The response in populations with continuous leptokurtic mutational effects is also compared to the main continuous effects results.

THE MODEL

The main simulations consist of 4000 diploid individuals, with three genetically correlated traits. Sexes are separate but identical, and all data are averaged over the two sexes. Mating is random, and generations are non-overlapping. Populations are given 20000 generations to reach stabilizing selection-mutation-drift equilibrium (hereafter simply equilibrium), before the start of directional selection. The trait means ($\overline{\mathbf{Z}}$) start and remain near their optimal values (θ) throughout this initial phase. Under most initial conditions, the genetic variances decline steadily for the first few thousand generations, before reaching their equilibrium levels (generally before generation 10000). To simulate directional selection, the optimal value for trait 3 is shifted upwards by 10 phenotypic standard deviation units. All other conditions are identical in both the equilibrating and directional phases, for a given population type. For the directional phase of a given population type, five replicates of 1500 generations are run.

Although all the graphs shown are from only three initial populations, many others with different parameter values were simulated, to check the generality of the results, in terms of the effect of population size (N = 4000 or 400), magnitude of peak-shift, and stabilizing selection intensity (**W**).

Creating the population

Each individual has L = 100 unlinked loci. Populations with continuous allelic effects are initialized by assigning a random normal variate with mean zero and standard deviation 1 to each allele in each individual. Discrete effects populations are randomly assigned a –0.5 or 0.5 at each allele. Each of the three traits is controlled by n = 50 loci, randomly assigned from the 100 available per individual. The pleiotropic relationship between traits was produced by randomly assigning 50 "1's" to each row of a 3 (traits) x 100 (loci) matrix **B** (equivalent to Wagner's (1989) **B** matrix). All other elements of **B** are assigned a "0". A "1" at element $\mathbf{B}_{ij}$ indicates that locus j contributes its allelic values to trait i. Columns with no "1's" represent loci that are not assigned to any trait, and are therefore selectively neutral. All individuals use the same **B**, which is assumed to be constant. The same matrix is used for all simulations discussed in this paper.

Assigning trait values

The genotypic value of each trait in an individual is defined as the sum of all allelic values at all loci that code (via **B**) for that trait, and is therefore additive between and within loci. The expected average genotypic value for all traits is zero before directional selection. Phenotypic values equal the genotypic values, plus a random normal deviate with a mean of zero and a standard deviation set so as to produce an initial heritability of 0.5 for all traits. This heritability is in general higher than that present after the population has equilibrated. The environmental variance for each trait remains constant throughout selection. The environmental deviates added to each genotypic value within an individual are independent, thus the expected environmental covariance between traits is zero.

Assigning survival probabilities and selecting parents

Each offspring is assigned a survival probability, according to

$$w(\mathbf{Z}) = \exp(-0.5(\mathbf{Z}-\theta)^{T}\,\mathbf{W}^{-1}(\mathbf{Z}-\theta)) \qquad (7)$$

(Lande, 1980a), where superscript ^{T} indicates matrix transposition. For directional selection, θ for trait 3 (θ_3) is set to $+10\ \sigma_P$ (phenotypic standard deviation units). θ for all other traits remains at zero throughout the simulation. Equation (7) gives values between 0.0 and 1.0, and can be interpreted as the probability of survival. Therefore, selection is frequency-independent, since the fitness of each individual, and the population as a whole, is determined solely by its proximity to the optimal vector of phenotypes. From those individuals that survive viability selection the previous generation, males and females are randomly assigned to monogamous pairs. Pairs are then randomly sampled with replacement, each time producing one offspring of each sex. Offspring consist of a random haploid compliment of genes from each parent. Offspring phenotypes and fitnesses are assigned as above. This procedure is repeated until there are enough surviving offspring to replenish the original population. The number of offspring that have to be sampled in order to re-establish the initial population size is therefore a measure of the mean fitness of the population. This method of "viability" selection (as used in Baatz and Wagner, 1997) produced results virtually identical to the alternative whereby parents were sampled (with replacement) each generation in proportion to their fitness (w), and produced offspring that automatically survived to stock the next generation (results not shown). All statistics and data are collected only from the surviving offspring.

Mutations are applied after selection, and do not affect that individual's phenotype.

Mutational effects are added to the value of pre-existing alleles. The formula for the house-of-cards approximation for the equilibrium genetic variance assumes that mutational effects are "essentially independent" (Turelli, 1984) of pre-existing allelic values. However, this assumption is required in order to simplify the mathematics, and is not intended as a statement concerning the actual effect of mutations in real populations. Therefore, "house-of-cards" populations in this paper have mutational effects of relatively high variance, and low mutation rates (compared to the Gaussian populations), but do not implement the simplication required for Turelli's approximation.

Constants, and parameter estimates

G and **P** are estimated at generation zero of directional selection, from the genotypic and phenotypic values of all individuals. The diagonal elements of **W** are set to 15 times the environmental variance of each corresponding trait. This is a value within the range of experimental estimates (Johnson, 1976; Turelli, 1984). The off-diagonal elements of **W** are set to zero. For the continuous effects models, genetic variance (V_G) = $2n\sigma_a^2 \cong$ 100, assuming global linkage equilibrium. Since the heritability of each trait is set at 0.5, V_E = the initial V_G. Mutational heritability (h^2_M), defined as the mutational variance V_M ($= 2n\mu\alpha^2$) / V_E, is set to 0.001, a value consistent with empirical findings (Lynch, 1988; Houle et al., 1996). Given V_E =100 and 2n=100, $\mu\alpha^2$ must equal 0.001 to produce this value (note that V_E is not set to the conventional 1.0). The Gaussian simulations use μ = 0.001 and α^2 = 1.0. While this violates $\alpha^2 << \sigma^2_a$, it does create Gaussian allelic distributions at equilibrium (confirmed by simulation). To simultaneously satisfy h^2_M = 0.001, $\alpha^2 << \sigma^2_a$, and n=50 would require mutation rates on the order of 10^{-2}. The HC

simulations use $\mu = 0.0001$ and $\alpha^2 = 10.0$. For populations with discrete allelic effects, a μ of 0.0001 is used as the rate at which each allele changes from 0.5 to –0.5, or vice versa. In all populations, the number of mutations per generation is drawn from a Poisson distribution with a mean of $2L\mu N$. For populations with leptokurtic mutational effects, the reflected gamma distribution is used, where the density function of mutational effects α (randomly assigned either a positive or negative sign) is given by

$$k^b e^{-k\alpha} \alpha^{b-1} / \Gamma(b) \qquad (8)$$

where Γ is the gamma function, b is a shape parameter and k is a scaling parameter adjusted so as to produce a mutational variance of 0.001 V_E as in the above simulations with normally distributed mutational effects. b is set to 0.5 to produce a highly leptokurtic distribution, as in several previous simulation studies (e.g. Keightley and Hill, 1989; Burger and Lande, 1994).

RESULTS

Figure 1 shows the observed and predicted trajectories of the trait means for the different models. The prediction for all three traits is very accurate for the Gaussian population (Figure 1A). The discrepancy between the average observed and predicted values is never greater than 0.3 σ_P for any generation. With a population size of 400 (results not shown), the predictions were nearly as good (maximum discrepancy = 0.6 σ_P). The predictions for the HC population (Figure 1B) are much less accurate, with discrepancies as large as 3.7 σ_P. In this population, equation (3) underestimates the rate of response and correlated response, while overestimating the time required to reach equilibrium. Predictions for N=400 HC populations underestimated the true response by 2.7 σ_P. When the starting population from Figure 1B was given a θ_3 shift of +5 σ_P instead

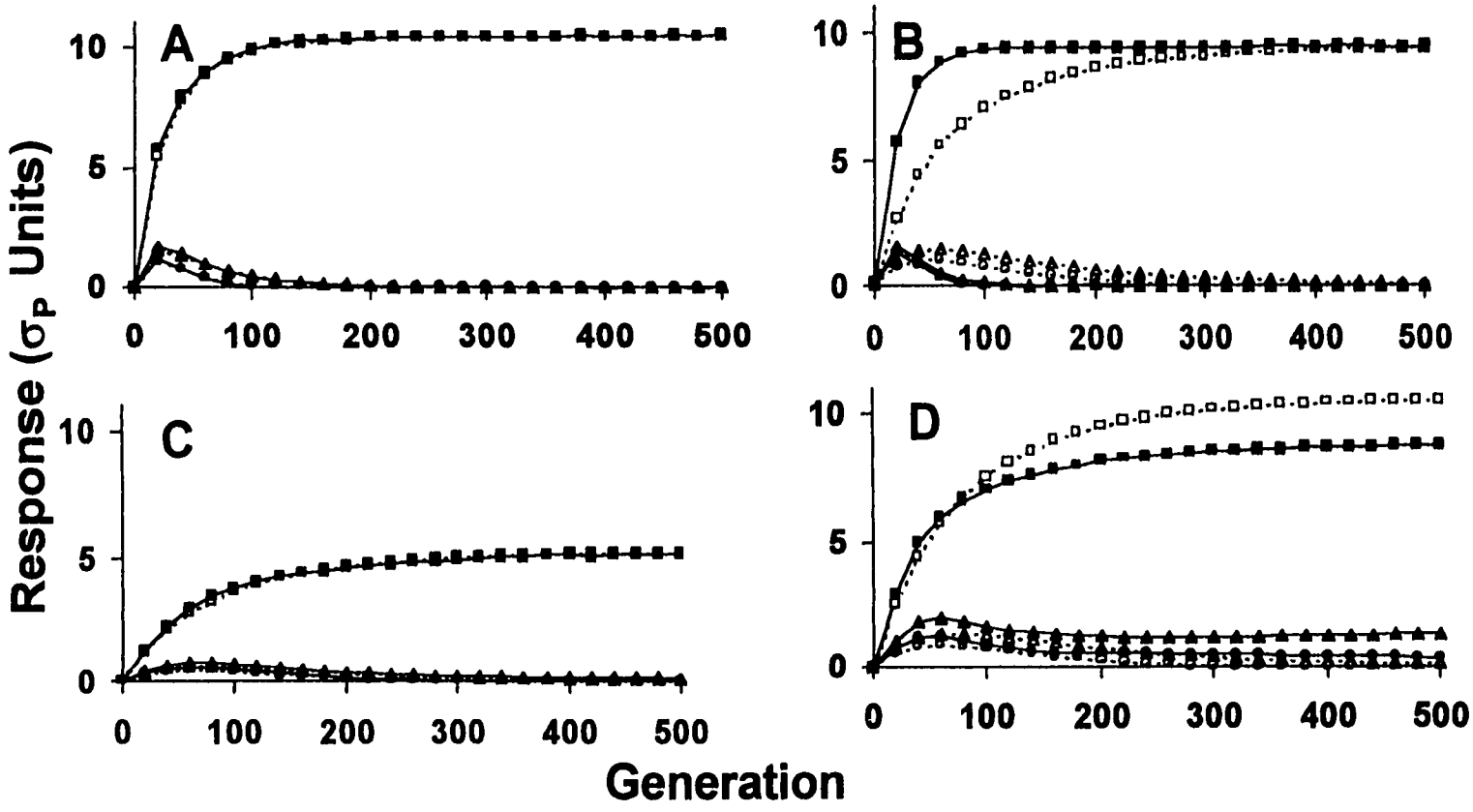

Fig. 1. Response of trait means to shifted-optima selection on trait 3. Units are phenotypic standard deviations. A) Gaussian conditions. B) House-of-cards conditions. C) and D) Biallelic loci. All peak-shifts are for +10 standard deviations, except C) which is for +5. Closed symbols, solid lines: simulation results. Open symbols, dashed lines: predictions from equation (3). Trait 1 = triangles, trait 2 = circles, trait 3 = squares. Heritabilities (averaged over traits) of the starting populations were 0.44, 0.14, 0.16 and 0.16 for A-D respectively. All graphs in this paper are based on the average of 5 replicate runs.

of +10, the results were qualitatively similar, with average discrepancies as large as 1.4 σ_P.

The quality of predictions from biallelic models depends on the relationship between the peak-shift and the selection limit (see Discussion). Figure 1C shows that when the peak-shift (5 σ_P in this case) is within the limit, predictions are good, and the average discrepancy was never larger than +0.2 σ_P. Average discrepancies in the 400 population size (results not shown) approached +0.3 σ_P. In Figure 1D, a peak-shift of 10 σ_P, as in Figures 1A and 1B, exceeds the limit, resulting in suboptimal evolutionary equilibria. Population size and mutation rate have no effect on this limit.

For all three types of population, running the directional phase of selection without mutation has virtually no effect on the trajectory of the means. Therefore, for the continuous effects populations, there is enough standing variance to easily move 10 σ_P. For the biallelic population, mutation rate has little effect on anything but the equilibrium variance (see Discussion).

Figure 2 shows the changes in several genetic parameters caused by directional selection, for the continuous effects models of Figures 1A and 1B. In the Gaussian population, genotypic variances increase by 15-25%, peaking at generations 80-90 (Figure 2A). The variances of traits 1 and 2 change more than that of trait 3, despite the means being displaced far less. In the HC population (Figure 2B), the variance peaks at generations 30-80, with an six-fold increase in trait 3, and a four-fold increase in traits 1 and 2. When the HC population was run at N = 400, variance still increased by up to four times. With N = 400 and $V_s = 60$ (rather than 16 as in the main simulations), variance increased by a factor of 1.8, although there was not a noticeable increase until about generation 20.

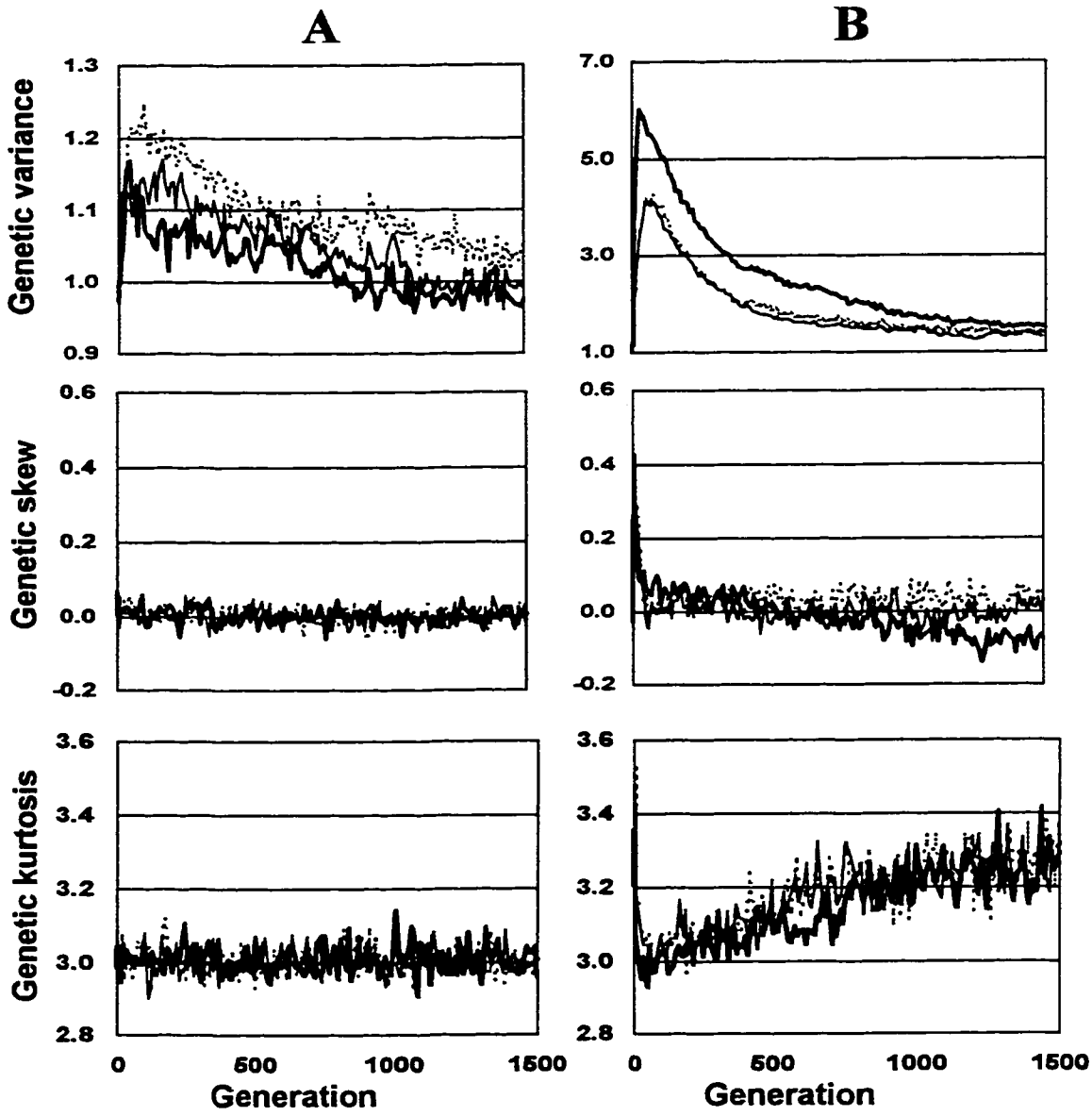

Fig. 2. Genetic variances, skews, and kurtoses for (A) Gaussian, and (B) House-of-cards populations corresponding to Figures 1A and 1B respectively. Genetic variances are standardized to the level in generation zero. Note that the scale is different for the two variance graphs. Trait 1 = thin dotted line; trait 2 = thin solid line; trait 3 = thick solid line.

The skew and kurtosis of the Gaussian population's trait genotypic values remain near the values of normal distributions (0.0 and 3.0 respectively). In the HC population, the skew for all traits is initially near 0.0. The skew in traits 1 and 2 remains near 0 except for the first 15-25 generations, when there is a positive skew of up to ≈ 0.3. The skew for trait 3 reaches a higher peak (≈ 0.4), but continues to decline for hundreds of generations. Kurtosis was high in the starting population ($\approx$3.4) but was quickly driven to normal levels, and then steadily increased from generation 300 onward. It eventually returned to pre-directional selection levels. These figures show how the genetic characteristics of the populations continue to evolve long after the trait means have reached an apparent equilibrium.

Figure 3 shows the genetic changes in the biallelic populations from Figure 1C and 1D. As with the continuous effects models, both populations initially show an increase in genetic variance. Notice that in Figure 3B, the population maintains variance for all three traits, despite the trait means being at a suboptimal equilibrium. The normality of the starting population is a consequence of the symmetry of the allelic effects (-0.5 and 0.5) about the optimum (0.0). This guarantees that directional selection will produce skew and positive kurtosis proportional to the peak-shift, given the restrictions on mutational effects.

In Figure 4, the effect of directional selection on genetic correlations in the four populations is shown. Only in the Gaussian population do the correlations remain relatively constant (Figure 4A). In HC populations (Figure 4B), the patterns of change are

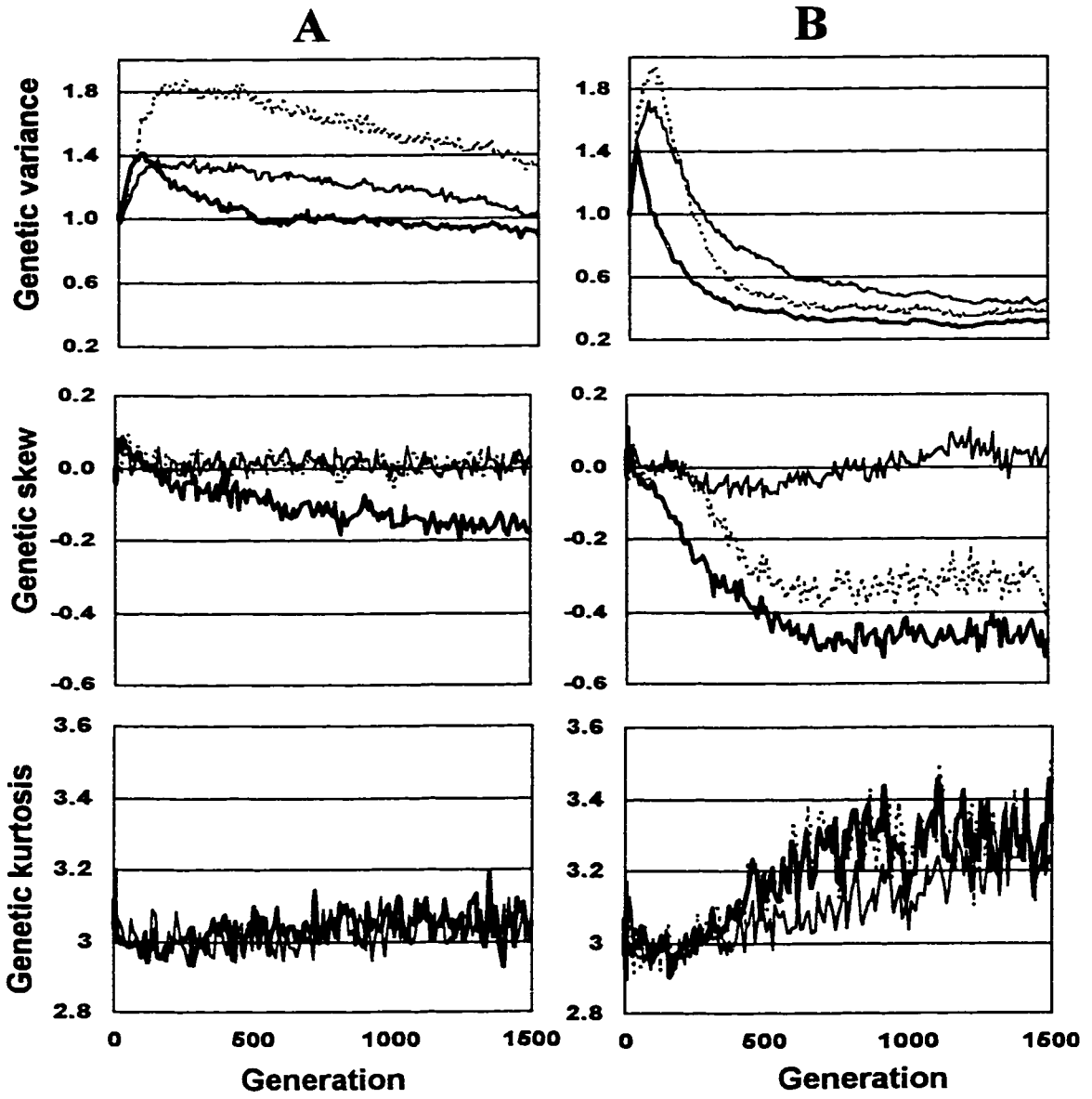

Fig. 3. Genetic variances, skews, and kurtoses for biallelic populations where the selective optima θ is either (A) within, or (B) beyond the selection limit. Genetic variances are standardized to the level in generation zero. A and B correspond to the populations from Fig. 1C and 1D respectively. Symbols as in Fig. 2.

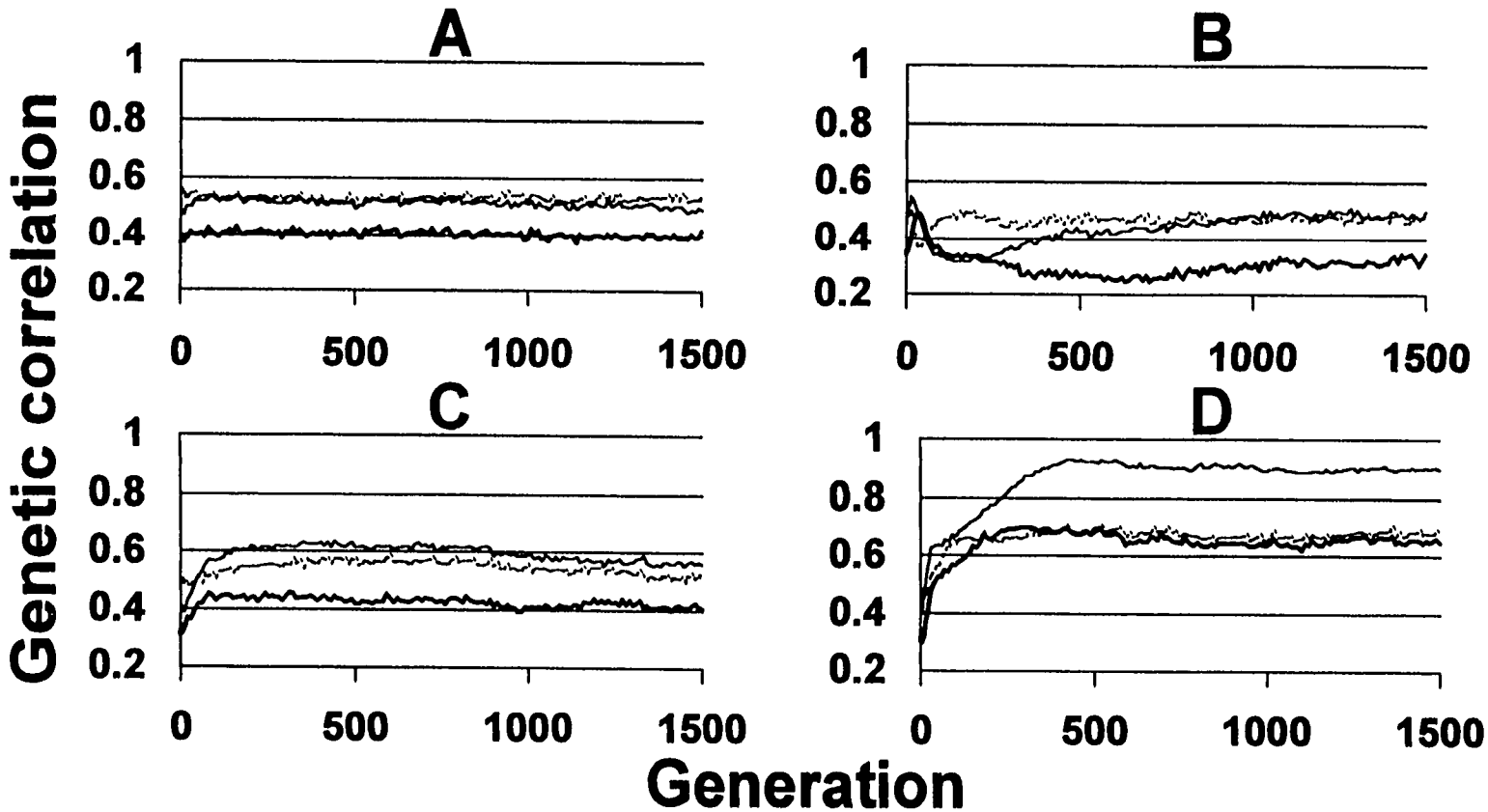

Fig. 4. Changes in genetic correlations caused by peak-shift selection A) Gaussian conditions. B) House-of-cards conditions. C) Biallelic loci. Peak-shift within limit. D) Biallelic loci. Peak-shift beyond the selection limit. A-D correspond to the trait responses with the same letter in Fig. 1. Correlation (r_{ij}) between traits i and j, r_{12} = thin dotted line; r_{13} = thin solid line; r_{23} = thick solid line.

irregular and are apparently very dependent on the exact starting conditions and the nature of the peak-shift. In the biallelic population (Figures 4C and 4D), correlations increase as loci unique to trait 3 increase the frequency of their high alleles. Figure 4D shows that pairwise correlations of less than 1.0 do not guarantee that traits will reach their optima in biallelic models if there are more than two traits in the system. This effect has previously been noted in algebraic models of multitrait systems (Slatkin, 1987b; Charlesworth, 1990).

Textbook descriptions of changes in the genetic correlation during directional selection are usually based on the biallelic model, as is appropriate for small, laboratory populations that are often derived from crosses between lines. The changes expected in large continuum-of-alleles populations (Figures 4A and 4B) are far less intuitive, due to the presence of continuously distributed allelic (mutational) effects.

The median allelic skew and kurtosis from the HC population are shown in Figure 5, where loci have been classified according to the combination of traits they control. There are eight classes (0,1,2,3,1+2,1+3,2+3,1+2+3), containing 19,7,11,14,13,10,6, and 20 loci respectively. The 19 neutral loci of class "0" are not shown in this figure, as they don't respond in any directed manner to the directional selection encountered here. As directional selection starts and rare alleles with large positive effect on trait 3 increase in frequency, the skew for the four classes that include trait 3 moves to a high level. (Note that the exact value of the skew in the equilibrium population is highly variable between generations, so these classes may start with almost any value.) The other non-neutral classes generate skew in the opposite direction at the point where the net selective forces start favouring smaller values of traits 1 and 2 (compare with Figure 1B). This effect of having both directions of skew in different subsets of genes will tend to produce genetic

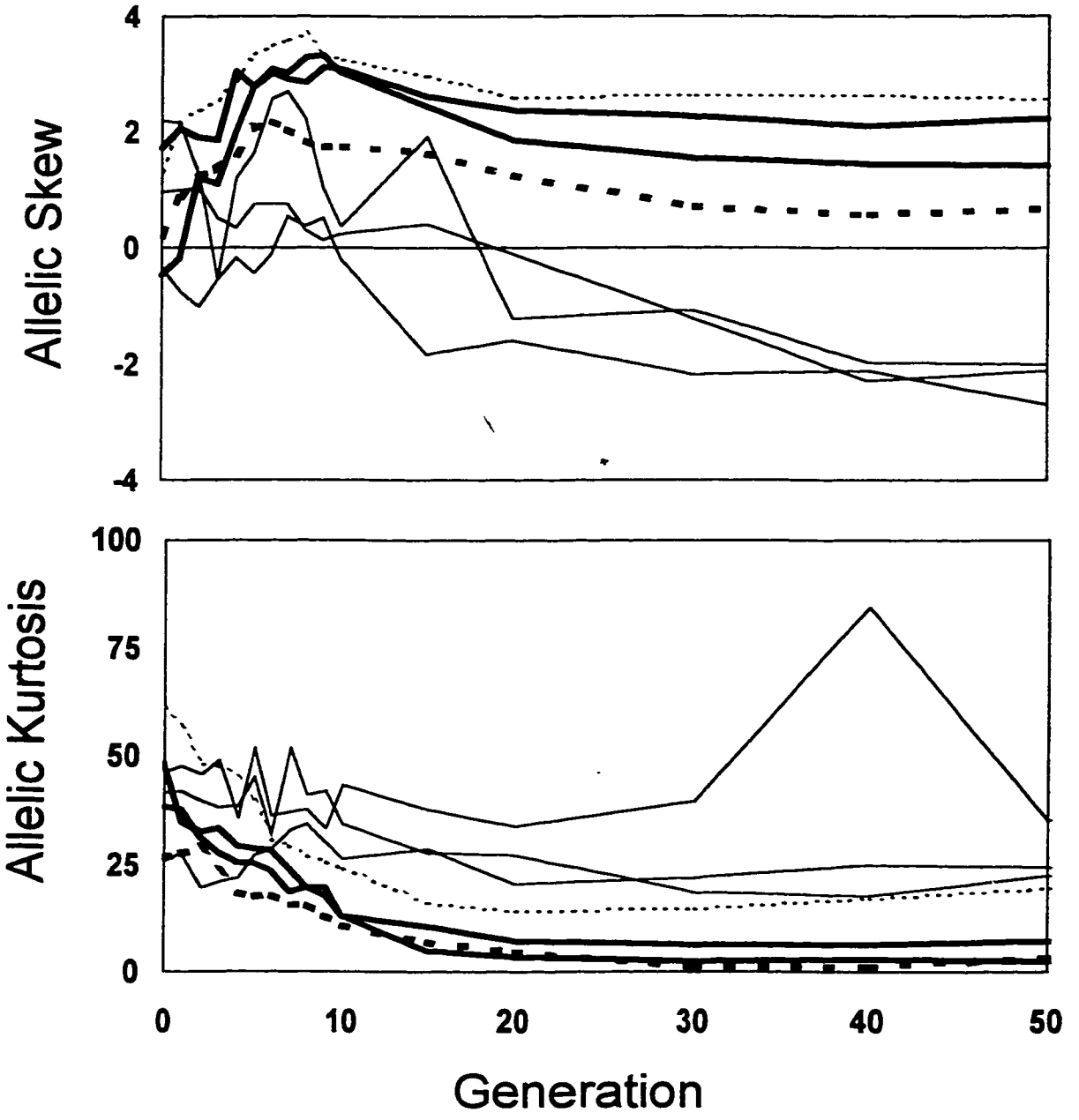

Fig. 5. Changes in median allelic skew and kurtosis for the HC population of Fig. 1B. Loci are classified according to the combinations of traits they control, as determined by the **B** matrix. Loci controlling traits: 1, 1+2, 2 = thin solid line; 1+2+3 = thin dotted line; 1+3,2+3 = thick solid line; 3 = thick dotted line.

distributions that are less skewed than their underlying allelic effects. This is important because standard Gaussian predictions like equation (3) assume that genetic variance remains constant on directional selection. Barton and Turelli (1987) have shown that this is a consequence of assuming that there is no allelic skew.

At the post-directional selection equilibrium, negative skew was highly significant in classes 3 (-1.6) and 1+3 (-1.03) (based on the median skew per class per generation, averaged over generations 2000 – 5000, measured every 50 generations to reduce autocorrelation (Keightley and Hill, 1988; Burger et al, 1989)). Thus, these two classes account for most of the negative skew seen in the genotypic distribution of trait 3 in the same simulations (Figure 2B). Although the variability amongst replicates is very small for the predicted and observed trait means, it increases rapidly with increasing moments of the genotypic distribution. As an example, two other sets of five replicates under the same conditions produced a) a long-term depression in the skew for trait 2, and b) no long-term depressions. It is likely that the populations can move between different equilibrium states (as in Barton, 1986), which can have a large influence on the behaviour of the higher moments, but less on the variance and, especially, the means. However, on the introduction of directional selection, all replicate sets behaved qualitatively as described above for Figure 5.

The kurtosis of the trait 3 classes declines rapidly on directional selection, again due to selection for rare alleles. The decline in kurtosis for the other classes is much weaker, and is generated by the same processes that produce negative skew.

The continuum-of-alleles simulations with leptokurtic mutational effects produced the expected results – greater allelic leptokurtosis, and therefore a greater increase in

variance in response to directional selection than populations with normally distributed mutational effects (results not shown). A mutation rate of 0.001 produced a maximum average deviation between simulation and equation that was approximately five times as large, and an increase in variance that was twice as large, as in the population of Figure 1A. Lowering the mutation rate to 0.0001 (at a fixed mutational heritability) resulted in four-fold increase in genetic variance over that of the population in Figure 1B, but little difference in the accuracy of trait mean predictions.

DISCUSSION

This paper attempts to answer a relatively straightforward question - if a change in environmental conditions causes selection for a new value of a single trait, can Lande's shifted optima equation be expected to predict accurately the trajectory by which this trait, and any others correlated with it, will evolve? The answer, like that to the question of how much variation can be maintained through mutation-selection balance, depends crucially on the nature of mutational effects. As almost nothing is known about the frequency, magnitude, or distribution of mutations at typical polygenic trait loci, it is not possible to assess the predictive accuracy of the equation. All that can be done, at least until more empirical data are available, is show what conditions are required for making accurate predictions and describe, qualitatively, the nature of the errors produced when these conditions are not met. Like other models, this simulation makes numerous simplifying and unrealistic assumptions. For instance, dominance, trait value epistasis, and physical linkage are absent, and fitness is determined completely by multivariate Gaussian selection on the trait values. However, the simulation is not primarily intended as a model of how evolution works. Rather, it makes assumptions consistent with those from standard

quantitative genetic theory, and asks how the unavoidable complications associated with finite populations and finite numbers of loci are likely to affect the predictions of a specific equation.

The main results from this paper can be summarized as follows:

1) If allelic distributions are approximately Gaussian, equation (3) can produce very accurate predictions, based only on information gathered at generation zero of directional selection. These predictions were accurate despite the fact that the genetic variances changed by more than 20% during the directional selection phase. Such populations start with little genetic skew or kurtosis, and this changes little with directional selection. Thus, under these conditions, evolutionary trajectories may be understood in terms of the simple parameters of equation (3).

2) Under HC conditions, the predictions can be very inaccurate. Equation (3) underestimates the rate of response and correlated response, and overestimates the time it takes for the traits to get to their equilibrium values. In most of the populations tested, including that in Figure 1B, correlated responses were also of greater magnitude than predicted, sometimes markedly so. Directional selection can result in very large increases in genetic variance as initially rare alleles increase in frequency.

3) If mutational effects in continuum-of-alleles populations are leptokurtically distributed, directional selection will cause larger increases in genetic variance for any given mutational heritability. This causes the trait means to respond more quickly than predicted, even in high mutation rate (= smaller mutational effect) populations.

4) For biallelic models, equation (3) makes good predictions as long as the peak-shift does not require trait mean equilibrium values that are more divergent than can be

accommodated by the genetic architecture of the "species" (see below). Predictions become progressively worse as the optimum exceeds the selection limit. In such cases, suboptimal equilibria will be reached, despite the presence of genetic variation in each individual trait.

5) For many of the parameter combinations used to test the generality of the main results, peak shifts of 10 σ_P resulted in genetic variances and covariances at the new equilibrium that were remarkably close to those seen in the population before directional selection. Therefore, interpretation of the role of drift versus selection in shaping the **G** matrix should be made with caution. It may be that the changes in **G** brought about by selection to new optima are often temporary, even in relatively small populations. If this is the case, the changes in **G** found in short-term laboratory selection experiments may be fundamentally different from those expected between populations or closely related species that have been experiencing different selection regimes for long periods of time.

Peak-shift models of the sort considered here have received relatively little theoretical attention, and most of this has dealt with very different types of questions. For instance, conservation biology issues have motivated research on the ability of a population to keep up with an optimum that is changing either gradually or randomly (Lynch and Lande, 1993; Burger and Lande, 1994; Burger and Lynch, 1995). Charlesworth (1993) considered an optimum that could also change cyclically, to study the effect of directional selection on the evolution of sex and recombination. Each of the above studies used a single trait, and did not use the univariate form of equation (3) to describe the evolution of the mean phenotype. Zeng (1988) used a modification of equation (3) to look at the effects of correlational selection on patterns long-term

correlated response in infinite populations. Other papers have considered the evolution of two traits, one under stabilizing and the other under exponential directional selection (e.g. Burger, 1986; Wagner, 1988; Baatz and Wagner, 1997). Barton and Turelli (1987) used allelic recursion equations to simulate peak-shift selection in a single trait system, using moment generating functions to make predictions for the mean and higher moments.

This is the first paper to test the predictions of equation (3) by simulation. Although more general predictive equations are available (Barton and Turelli, 1987; Burger, 1993), they require detailed information about higher genetic moments or cumulants that are generally unobtainable. In addition, none of these has been extended to multivariate systems. Equation (2) and its descendants, including equation (3), are popular because they attempt to predict, or at least explain, evolution in terms of a small number of relatively familiar parameters that can, in principle, be estimated.

Causes of prediction error

The rapid increase in genetic variance seen in the HC population (Figure 2B) causes the means to respond to selection far more rapidly than predicted by equation (3). This increase in variance as a result of directional selection has been shown previously in single trait simulations by Barton and Turelli (1987; peak-shift), Keightley and Hill (1989; pure directional [w = Z]), Burger (1993; exponential), and Burger and Lande (1994; shifting optimum). The increase is largely due to selection for rare alleles that initially contribute little to the variance. As these alleles increase in frequency, they contribute more to the variance, and soon cause the mean to increase at an accelerating rate (Barton and Turelli, 1987). For a fixed mutational variance, lowering the mutation rate will result in equilibrium populations that have less genetic variance, but higher allelic skew and

kurtosis. The accelerated response of the mean is not simply a consequence of the lower equilibrium heritability (see caption for Figure 1) in HC populations. Increasing this heritability to Gaussian population levels, either by decreasing the strength of stabilizing selection or increasing the total (L) and trait-specific (n) number of loci, did reduce the amount by which the variance increased during directional selection. However, the quality of the predictions was only slightly improved, with the discrepency between observed and predicted response remaining roughly an order of magnitude greater than in the Gaussian populations (results not shown).

In using equation (3), it is assumed that the distribution of genotypic values is, and will continue to be, multivariate normal, and therefore that the dynamics of trait mean evolution can be described completely in terms of the mean and variance. HC populations have far more evolutionary potential, in terms of rate of response, than Gaussian populations with the same variance. Therefore, heritability is not an accurate predictive statistic in such populations.

An increase in genetic variance is seldom seen in artificial selection experiments, which would seem to be evidence against the generality of HC conditions (discussed in Keightley and Hill (1989)). Burger (1993) concluded that a significant increase in variance is unlikely if N_e is less than about 500. This figure was based on typical parameter estimates for the HC model, combined with the fact that genetic variance in his model converged to the mutation-drift equilibrium level under a particular form of exponential selection. For the peak-shift selection used in the present simulation, HC populations with an N of 400 ($N_e \cong 300$) still had a fourfold increase in variance. When the intensity of stabilizing selection was decreased (N=400, $V_s = 60$), there was a 1.8 times

increase in variance. Therefore Burger's conclusion may not extend to all forms of directional selection. However, the N_e in selection experiments is typically much smaller than 300. In addition, we know nothing about how existing univariate estimates of stabilizing selection intensity should be adjusted when considering multivariate systems. Therefore, failure to detect increased variance in selection experiments probably cannot be taken as evidence against the HC model.

Selection limits in models with discrete allelic effects

Discrete effects models have been used extensively in quantitative genetics (e.g. Latter, 1960; Bulmer, 1972, 1980; Barton, 1986,1989; Turelli and Barton,1990). They may be interpreted either as a realistic representation of the allelic effects for at least some loci, or as a method of simplifying the analysis of continuous effects models (Houle, 1989). In the latter case, results such as those in Figure 1C may lead to unwarranted confidence in the predictive ability of standard theory (such as equation (3)) if HC conditions are the norm. The lack of rare alleles of large effect in the biallelic simulation has produced a result consistent with equation (3) because the behaviour of the variance is more similar to that seen in the Gaussian than in the HC population. Here, trying to extend the results of the "simplifying" model to the situation with continuous effects would bias the conclusions.

Alternatively, if discrete effects models are taken as a realistic representation of allelic effects, the limit problem deserves some consideration, at least when modeling the evolution of trait means. To see why the limit occurs, consider a system of 20 genes with free recombination, where loci 1-12 and 9-20 control traits 1 and 2 respectively. If the two traits are selected in opposite directions, the correlation between traits will approach

1.0 as the loci unique to each trait fix in the appropriate direction. Genetic variance for both traits remains, however, since loci 9-12 will still be segregating. The amount of divergence between traits is a function of the number of loci unique to each trait, and their allelic effects. The same situation exists for systems of three or more traits, except that pairwise correlations at the limit no longer have to be 1.0 (e.g. Figure 4D), since each trait will generally share segregating loci with more than one other trait. At a suboptimal limit, as in Figure 1D, the genotypic fitness of the population cannot increase, since mutations cannot improve upon the alleles that are already present. Therefore, mutation rate has no effect on the limit. In Figure 1D, the eigenvalue corresponding to an increase in trait 3 and a decrease in the other traits is near zero, so the **G** matrix is nearly singular for that direction of response. It is not completely singular because mutations continue to produce genotypes that are slightly less fit than those at the limit.

If the alleles at each locus are typically restricted to a finite number of values, the simulations suggest that the situation found in Figure 1D might be common, since all it requires is a large peak-shift (= prolonged directional selection). This situation would be characterized by a lack of response in a population, despite the presence of genetic variance for each trait, genetic correlations less than 1.0, and non-zero values for the coefficients of the phenotypic selection gradient (measured as in Lande and Arnold, 1983). These non-zero coefficients exist because environmental variance can produce phenotypes more fit than those at the genetic limit, but this fitness difference is not heritable. There are in fact several examples of such a lack of response in natural populations (e.g. Price et al., 1988; Alatalo et al., 1990; van Tienderen and de Jong, 1994; Weis, 1996) although in most cases, the authors have provided compelling evidence for simpler explanations.

These include the effect of missing traits on the analysis, and non-heritable traits influencing the focal trait(s) and fitness through different pathways (Price et al., 1988; Rausher, 1992).

It should be noted that in discrete effects populations with allele frequencies near fixation (as when directional selection has driven the population to a suboptimal selection limit), subsequent selection in the direction of the rare allele causes a pattern of response similar to that seen in HC populations (results not shown). Genetic variance increases as the rare alleles become more common, and the trait means respond to selection more rapidly than predicted by equation (3). However, the response is very slow compared to HC models, due to much lower initial heritabilities.

Although this paper has examined the predictive ability of only one equation, a large number of other theoretical models are based on the same underlying assumptions, stemming from the use of equation (1). These include models for the evolution of sexual size dimorphism (Lande, 1980b), phenotypic plasticity (Via and Lande, 1985), maternal effects (Kirkpatrick and Lande, 1989), and epigenetic effects (Atchley and Hall, 1991) to name only a few. If "house-of-cards" assumptions are more realistic than those of the Gaussian model, some of the conclusions of these models are likely to be at least quantitatively inaccurate. For instance, in Lande's (1990b) paper on the evolution of sexual size dimorphism, he models the situation where sexual selection for increased values of a trait in males causes a temporary, maladaptive increase in the homologous trait in females. Given typical genetic correlations between the sexes, he concludes that the time for the traits in each sex to reach their equilibrium values may be on the order of millions of generations. From the simulation results in this paper, HC conditions might be

expected to reduce that time substantially.

Given current estimates of mutation rates and mutational heritabilities, it is likely that allelic effects are leptokurtically distributed. Therefore, directional selection in moderate to large-sized populations is likely to cause an increase in genetic variance. Because of this, Gaussian-based quantitative genetic models will often underestimate the rate at which trait means respond to selection. In models involving stabilizing selection, this will result in overestimates of the time required for populations to reach equilibrium.

Chapter 3. Predicting the evolution of sexual size dimorphism

Sexual selection is presumed to be a major factor in the evolution of SSD. In the previous chapter, I explored the effect of frequency-independent stabilizing selection, with either biallelic or continuum-of-alleles simulation models. I now consider the complications related to predicting response to selection under frequency-dependent sexual selection. In addition, the single-sex model of chapter two is used to look at the evolution of SSD through frequency-independent selection.

This chapter has been submitted to *Evolution* as:

Reeve, J. P., and D. J. Fairbairn. Predicting the evolution of sexual size dimorphism.

ABSTRACT

Lande's (1980b) equations for predicting the evolution of sexual size dimorphism (SSD) through frequency-dependent sexual selection were tested against results obtained from a stochastic genetic simulation model. SSD evolved faster than predicted, due to temporary increases in the genetic variance brought about by directional selection. Predictions for the magnitude of SSD at equilibrium were very accurate for weak sexual selection. With stronger sexual selection the total response was greater than predicted. In simulations with continuous allelic effects, large changes in sexual size dimorphism can occur without significant long-term change in the genetic correlation between the sexes. Sex differences can evolve more quickly under stabilizing natural selection, caused by a shift in trait optima, than through sexual selection that is opposed by stabilizing natural selection. Our results suggest that genetic correlations constrain both the short-term and long-term evolution of SSD less than predicted by the Lande model.

INTRODUCTION

Sexual size dimorphism is generally attributed to differences in the net selection acting on the two sexes, so that the sexes have different optimal sizes. When each sex is at its optimum, the population will be at an evolutionary equilibrium. If selection changes, the optimal value for one or both sexes may shift. Since the genetic correlation between sexes is very high for most morphological traits (Lande, 1980b; Roff, 1997, table 7.4), changes in the mean size of one sex are expected to be associated with similar changes in the other sex. Short-term laboratory experiments have demonstrated that artificial selection on the size of one sex produces a nearly parallel response in the other (Alicchio and Palenzona, 1971; Reeve and Fairbairn, 1996). Thus it is possible that the SSD found in natural populations might often be maladaptive, if long periods of time are required to overcome genetic constraints. How long this maladaptive state persists is important in terms of understanding the adaptive significance of SSD. Quantitative genetics provides a useful framework for understanding and predicting the evolutionary trajectories of traits such as body size, and sex differences in these trajectories.

Lande has introduced a number of equations for predicting response to selection under a variety of settings. Most are variations of the "multivariate response equation" (Lande, 1979)

$$\Delta \mathbf{Z} = \mathbf{G}\boldsymbol{\beta} \qquad (1)$$

(symbols are defined in Table 1), which is the multivariate equivalent of the "breeder's equation"

Table 1. Symbol definitions (with dimensionality). n = number of traits.

B	Between-sexes genetic variance-covariance matrix (n x n) [†]
β	Selection gradient $= \mathbf{P}^{-1}\mathbf{S}$ (n x 1)
c	Intensity of sexual selection (n x 1) $= \mathbf{P}^{-1}\mathbf{S}^{*}$
G	Genetic variance-covariance matrix (n x n)
h^2	Heritability
k	Coefficient for mating success function
N	Population size [males + females]
P	Phenotypic variance-covariance matrix (n x n)
R	Response to selection
S	Selection differential (n x 1)
S*	Sexual selection differential (n x 1)
θ	Optimal trait value (n x 1)
w	Expected fitness
W	Stabilizing selection matrix (n x n) [Smaller values indicate stronger selection]
z	Individual phenotype (n x 1)
Z	Population phenotypic means (n x 1). $\Delta\mathbf{Z}$ = change in **Z** across generations

[†] Unlike **G**, **P**, and **W**, **B** is not (necessarily) symmetric.

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

Equations 1 and 2 are valid only for a single generation of selection. Extending the predictions over longer periods requires two assumptions, one controversial and the other generally unwarranted. First, the genetic parameters (**G** or h^2) must remain constant over time. How likely this is in natural populations has been the subject of considerable debate (e.g. Shaw et al., 1995). Second, selection (β or S) must remain constant, which will not generally be true. For example, under stabilizing selection, β will change as the trait means move relative to their optimal values. To deal with changes in β under a model of multivariate Gaussian (stabilizing) selection, Lande (1980a) introduced a "peak-shift" equation

$$\Delta\mathbf{Z} = \mathbf{G}(\mathbf{W}+\mathbf{P})^{-1}(\theta-\mathbf{Z}) \qquad (3)$$

where $^{-1}$ indicates matrix inversion, that models evolution as being driven by changes in the optimal value(s) of one or more traits in a population. From this equation, it is clear that selection in any generation depends in part on how far the trait means are from their optimal values, and therefore will not be constant unless the population means are at an equilibrium (i.e. there is no further response).

Each of the above equations assumes that the sexes are identical. To model the evolution of sexual size dimorphism, Lande (1980b) used a variant of equation 3 that requires an additional matrix (of genetic variances and covariances between the sexes),

separate input values for each sex, and a term for the intensity of sexual selection

$$\Delta \mathbf{Z}_m = 0.5\mathbf{G}_m((\mathbf{W}_m+\mathbf{P}_m)^{-1}(\theta_m-\mathbf{Z}_m)+\mathbf{c}_m) + 0.5\mathbf{B}((\mathbf{W}_f+\mathbf{P}_f)^{-1}(\theta_f-\mathbf{Z}_f)+\mathbf{c}_f) \quad (4a)$$

$$\Delta \mathbf{Z}_f = 0.5\mathbf{G}_f((\mathbf{W}_f+\mathbf{P}_f)^{-1}(\theta_f-\mathbf{Z}_f)+\mathbf{c}_f) + 0.5\mathbf{B}^T((\mathbf{W}_m+\mathbf{P}_m)^{-1}(\theta_m-\mathbf{Z}_m)+\mathbf{c}_m) \quad (4b),$$

where $_m$ and $_f$ refer to males and females respectively, and ^{T} denotes matrix transposition.

Lande considers the situation where both sexes are initially under stabilizing selection for their optimal trait values, and the population is at equilibrium. Sexual selection of a constant intensity is then added, favouring an increase in male size. Iterating equation 4 over many generations results in male and female trajectories that Lande describes as occurring in two phases. In the first, rapid phase, the two sexes evolve nearly in parallel, and the mean size quickly reaches a point between the male and female final equilibrium values. Although the optimal size for females does not change, mean size in females temporarily increases due to the genetic correlation with males. In the second phase, which takes much longer than the first if the correlation between sexes is high, the two sexes slowly diverge. Females return to their original optimum, and males move to a new optimum determined by a balance between viability selection favouring their original size, and sexual selection favouring increased size. The optimal values will always be reached if there is *any* genetic variation for the required direction of change present at generation zero. This is a consequence of the assumption that the genetic parameters (**G** and **B**) remain constant, so that even a very small amount of the appropriate variance in the original population will never be exhausted.

Reeve (2000) used a stochastic genetic simulation model to test predictions from equation 3 for a single-sex system of three genetically correlated traits. The accuracy of predictions depended critically on assumptions concerning the distribution of mutational effects in the simulation model. Predictions were very good under "Gaussian" assumptions (high mutation rates, low mutational variance). However, it is likely that such conditions are unrealistic (Turelli, 1984). Using more realistic "house-of-cards" conditions (lower mutation rates, higher mutational variance), predictions were less reliable. In particular, equilibrium values were reached much more quickly than predicted. This result was caused by an increase in genetic variance in the early generations of directional selection, which is a result of the leptokurtic distribution of allelic effects expected under house-of-cards conditions (Turelli, 1984).

In this paper, a similar genetic simulation model, with house-of-cards assumptions, is used to test the accuracy of equation 4 in predicting the evolution of SSD. Predictions are made based solely on information available at generation zero of sexual selection. Although the simulation is set up to match the basic conditions of Lande's analytical model in terms of modes of selection and life-cycle stages, it differs in one important respect. Since the simulation uses finite numbers of individuals, each with finite numbers of loci, changes in allele frequencies result in changes in phenotypic and genetic distributions. Therefore, several variables (**G**, **P**, **B**, and $\mathbf{c}_m$) which are assumed constant in equation 4, are free to vary in the simulation. Since all real populations violate the constancy assumptions to at least some extent, we are interested in testing the effect of such violations on the quality of predictions from the analytical model. We focus on two main questions: 1) Does equation 4 accurately predict the equilibrium SSD? and 2) Does it

accurately predict the rate of separation between the sexes in the initial generations of selection.

Most of our results rely on frequency-dependent (sexual) selection to shift male size, as in Lande's (1980b) paper. However, we compare these with results from simulations where SSD is caused by frequency-independent (natural) selection for a higher optimal size in males, as might occur under the dimorphic niche model for the evolution of SSD (Slatkin, 1984; Shine, 1989).

METHODS

Although Lande's SSD model can deal with multiple traits, our simulation focuses on the evolution of a single trait, for simplicity of presentation. However, we retain the matrix notation of equation 4 for referring to the parameters and variables, to be consistent with Lande's equations. For the single trait case examined here, **G** and **P** are equivalent to the genetic and phenotypic variances respectively, and **B** is equivalent to the genetic covariance between the sexes. The basic simulation model is described fully in Reeve (2000), but several details have been modified, in particular those required for implementing sexual selection. Each individual in the population is represented by 50 unlinked, autosomal, diploid loci. Allelic values are initially assigned by drawing numbers at random from a normal distribution with mean of zero and a variance of 1.0. Genotypic trait values are purely additive, with no dominance or epistasis. For each sex, three of the 50 loci do not contribute to the genotypic value (i.e. are not expressed). These three loci are different in males and females, so that sex differences are caused by differences between the sum of the three "unique-to-female" and the three "unique-to-male" loci. For each individual, the genotypic value of its single trait is calculated as the sum of the allelic

values at its 47 contributing loci. Phenotypic values are created by adding to each genotypic value a random normal deviate, with a mean of zero, and a variance chosen to give an initial (pre-equilibrated) heritability of 0.5. Populations consist of 2000 individuals of each sex, and generations are non-overlapping. All Figures are based on average values of ten replicates, unless otherwise stated. The simulations were initiated from base populations that were equilibrated for 20,000 to 50,000 generations of stabilizing selection (i.e. no sexual selection), where θ was set at 0.0 for both sexes. During equilibration, the heritabilities fell to levels determined by the strength of stabilizing selection, the mutation rate, the variance of mutational effects, the number of loci, and the effective population size. In equation 4, **Z** is the only variable, as all the other symbols are assumed to remain constant. In the simulation, θ, **W**, c_f, and k (a coefficient that determines the strength of the linear relationship between mating success and relative body size in males, and thus is largely responsible for the magnitude of c_m) are the only fixed parameters. The variables (**P**, **G**, **B**, and c_m) are free to change over time. The ten replicates are sampled from the appropriate base population at 3000 generation intervals, starting at generation 20,000. Therefore, they are replicating the population parameters but the variables vary amongst replicates. The term "generation 0" in the text and Figures refers to the populations at the start of either sexual selection or natural selection for a new optimum.

Life-cycle

Each generation consists of three stages:

1) Frequency-independent viability selection.

Each individual is assigned an "expected" fitness, according to

$$w_1 = \exp(-0.5(\mathbf{z}-\theta)^T \mathbf{W}^{-1}(\mathbf{z}-\theta)) \quad (5)$$

(Lande, 1980a). Equation (5) yields values between 0.0 and 1.0, and can be interpreted as the probability of survival. Fitness for this stage is at a maximum when the phenotype ($\mathbf{z}$) is at the natural selection optimum.

2) Frequency-dependent sexual selection for increased size in males.

Expected fitness for this stage is based on a linear function of each male's rank,

$$w_2 = 1 - (\text{rank} / (N / 2))(k) \quad (6)$$

where the rank of the largest = 1, and the smallest = N/2. Lifetime expected fitness in males is the product of the fitness in stages 1 and 2. Female lifetime expected fitness is based entirely on stage 1. There is no fecundity selection in females independent of that caused by differences in survivorship.

3) Offspring production. Parents are sampled (with replacement) with a probability proportional to their lifetime expected fitness, with each set of parents producing one offspring of each sex. Offspring consist of a random haploid compliment of genes from each parent. Offspring phenotypes are assigned as above. This procedure is repeated until there are enough offspring to replace the parental population.

Parameter estimates

All parameter estimates for equation 4 were made during generation zero. **P**, **G**,

and **B** are based on the population before selection is applied (i.e., they are not fitness weighted statistics). The intensity of sexual selection, $\mathbf{c_m}$, is defined (in single-trait equations) as the mean size of males after sexual selection minus the mean size after natural selection (= the sexual selection differential, **S***), divided by the phenotypic variance (Lande, 1990b). We calculate this from fitness-weighted means. Although the weighting after natural selection is determined from equation 5, that after sexual selection is based on actual lifetime fitness (the total number of offspring produced by each individual), and is therefore a probabilistic function of equations 5 and 6. $\mathbf{c_f}$ is set to 0.0. All Figures except Fig. 6B and 6C use values measured before stage one of the life-cycle (see above). Since the base populations have a mean near 0.0 in each sex, we define SSD as the difference between male and female size, rather than as their ratio.

Mutations

Our main results use a continuum-of-alleles "stepwise" mutation method. The mutation rate is set at 1×10^{-4} per haploid locus. The mutational variance is set so that the mutational heritability is approximately 0.001 times the environmental variance, a value consistent with empirical findings (Lynch, 1988; Houle et. al., 1996). Mutational effects have a mean value of zero, and are added to the allele's pre-mutational value. In our discrete (biallelic) simulations, all alleles are either "a" or "–a", where the value of "a" is adjusted to give initial genetic variances similar to the continuous alleles model. Here, mutations simply cause the allelic value to flip from one state to the other.

RESULTS

Figure 1 shows predicted and observed responses in the two sexes to either sexual (Fig. 1A, B) or natural (Fig. 1C) selection for increased male size. In the continuum-of-

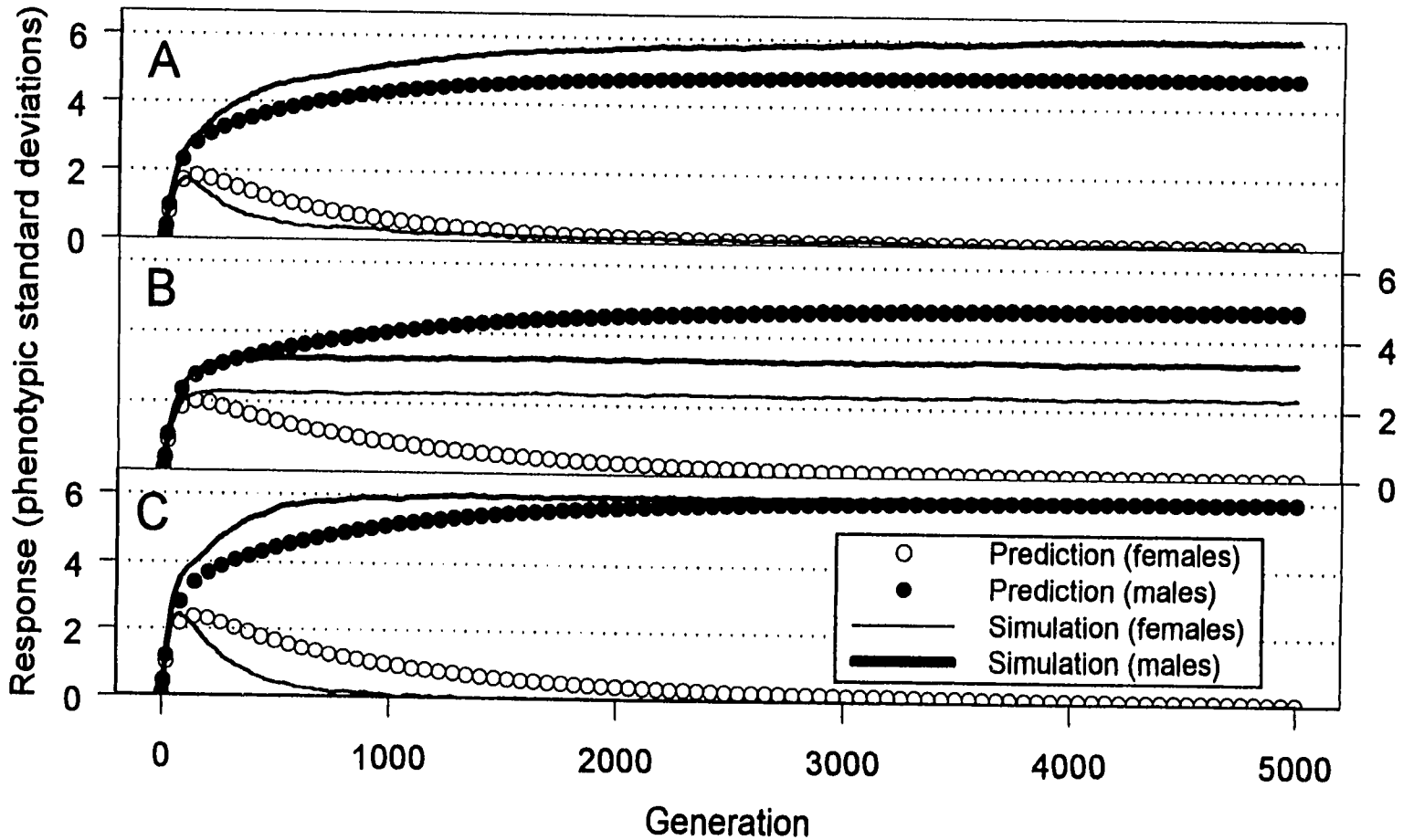

Fig. 1. Response to selection for increased male size. A) Sexual selection, continuous model. B) Sexual selection, discrete (biallelic) model. C) Natural selection, continuous model. The same coefficient of stabilizing selection (**W** = 19) is used for all three panels. The mating success coefficient (k) is 0.8 for A and B, and 0.0 for C. θ_m is 0.0 for A and B, and 6.0 for C.

alleles sexual selection simulations (Fig. 1A), the equilibrium male size, and therefore equilibrium SSD, are larger than predicted. From generations 500 to 5000, males and females are predicted to change size at approximately the same rate, but in opposite directions. However, in the simulated populations over the same time period, males increased in size by more than three times the amount by which females decreased. This occurs because, in the simulated populations, the intensity of sexual selection (c_m) increases as male mean size is driven further from the stabilizing selection optimum (see below for an explanation of this). The net result is that, contrary to expectations from equation 4, males tend to be further from their final equilibrium size than females for most of their evolutionary trajectories, and the equilibrium values for males and SSD are higher than predicted.

In the discrete model (Fig. 1B) equilibrium male size and SSD are smaller than predicted. This is because there is a limit to the separation possible between the sexes in populations with a finite number of allelic values at each locus. In our simulations, this limit is three times the difference between homozygotes, occurring when the diploid loci are fixed at "a" for the three male-specific loci, and "–a" for the three female-specific loci. This limit is responsible for the sub-optimal equilibrium reached in Fig. 1B. Generally, in finite-alleles models the agreement between observations and predictions depends largely on whether the optimal divergence exceeds this genotypic limit. See Reeve (2000) for a more thorough discussion of this point as it pertains to the evolution of correlated traits within a sex.

For comparison with the rate at which equilibria are reached under sexual selection, simulations were run where the shift in optimal male body size was achieved through purely natural selection (Fig. 1C). In this case, θ (equation 3) in males was changed to equal the same number of phenotypic standard deviations as the equilibrium deviation in Fig. 1A, and the same initial (generation 20,000) base population was used. It is clear that the equilibrium values are reached much more quickly with pure natural selection. This is due mainly to the fact that with sexual selection, the net selection intensity for increased size is reduced due to the conflict with stabilizing selection. With pure stabilizing selection, there are no conflicts and the realized selection intensity is therefore stronger. The equilibrium sizes of both sexes are predicted correctly since they depend only on θ_m and θ_f, which are parameters for both the equation and the simulation. Table 2 shows the significance of the difference between observed and predicted body sizes in the two sexes, for the three sets of simulations in Fig. 1.

As expected, the distribution of allelic effects in continuum-of-alleles base populations was highly leptokurtic (Fig. 2), usually characterized by having one or two common alleles, and several rare alleles with extreme values.

Fig. 3 shows changes in the genetic covariance and correlation between the sexes, from the simulations shown in Fig. 1. In the continuous models (Fig. 3A, C), movement of trait means is accompanied by large, but transient increases in the genetic covariance. In Fig. 3A, the covariance eventually equilibrates at a level considerably lower than that of the base populations. In Fig. 3C, it equilibrates at approximately the same level found in the base populations. The larger increase in covariance in 3C is due to the stronger net directional selection (β) under purely stabilizing selection. The genetic correlation

Table 2. Paired t-test for predicted vs. simulated body size. Values shown are the mean difference over the 10 replicates (observed – predicted). The three sets of results correspond to the three panels of Fig. 1.

	Sexual selection: continuous model		Sexual selection: biallelic model		Natural selection: continuous model	
Generation	Female	Male	Female	Male	Female	Male
10	-0.002	0.012	0.018	-0.003	0.015	0.043
100	-0.062	0.172 *	0.123 *	0.057 †	0.074	0.665 ***
250	-0.737 ***	0.526 ***	0.421 ***	0.067	-0.901 ***	0.851 ***
500	-0.872 **	0.921 ***	0.819 ***	-0.227 **	-1.233 ***	1.119 ***
1000	-0.535 *	1.013 ***	1.392 ***	-0.769 ***	-0.890 ***	0.778 ***
2000	-0.247 †	1.139 ***	1.957 ***	-1.259 ***	-0.365 *	0.305 *
3000	-0.112	1.142 ***	2.178 ***	-1.442 ***	-0.191 †	0.138
4000	-0.071	1.217 ***	2.236 ***	-1.531 ***	-0.083	0.086
5000	-0.048	1.230 ***	2.305 ***	-1.501 ***	-0.004	0.037

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

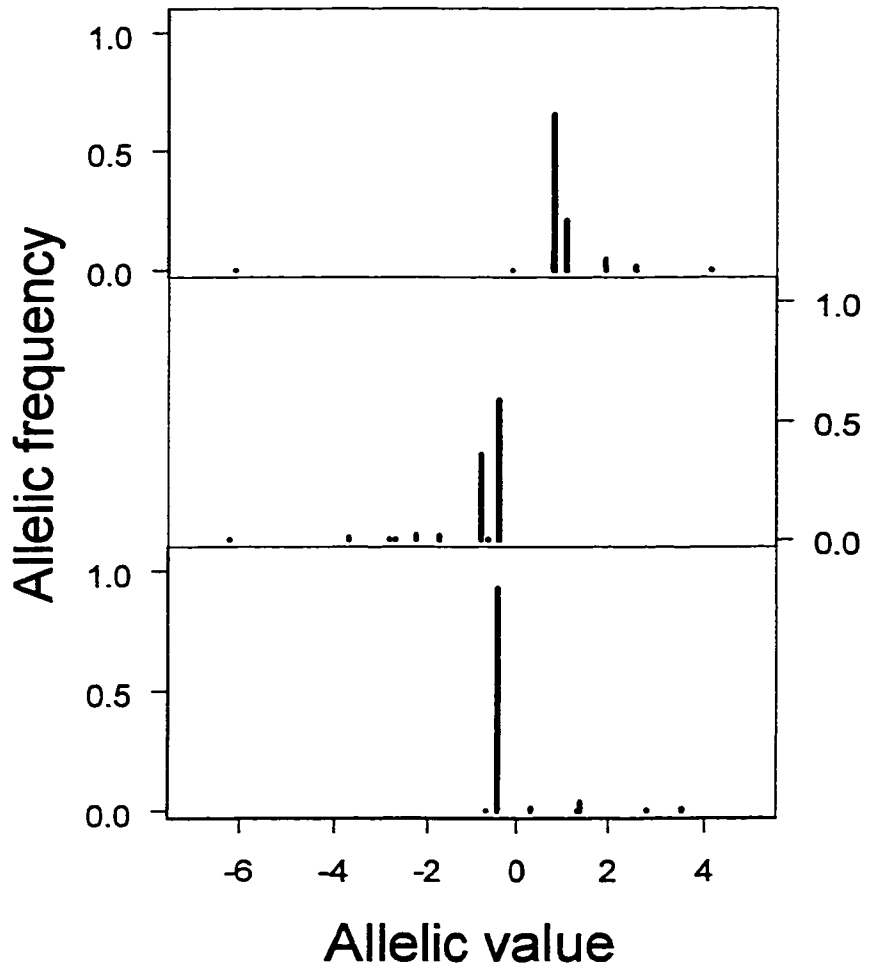

Fig. 2. Examples of the leptokurtic distribution of allelic values under the house-of-cards model. Shown are allelic distributions of three random, non sex-limited loci from a base population used in one of the replicates of Fig. 1A. Frequencies are based on a total of 8000 alleles per locus. Each bar represents a single allelic value, not the midpoint of a range of values.

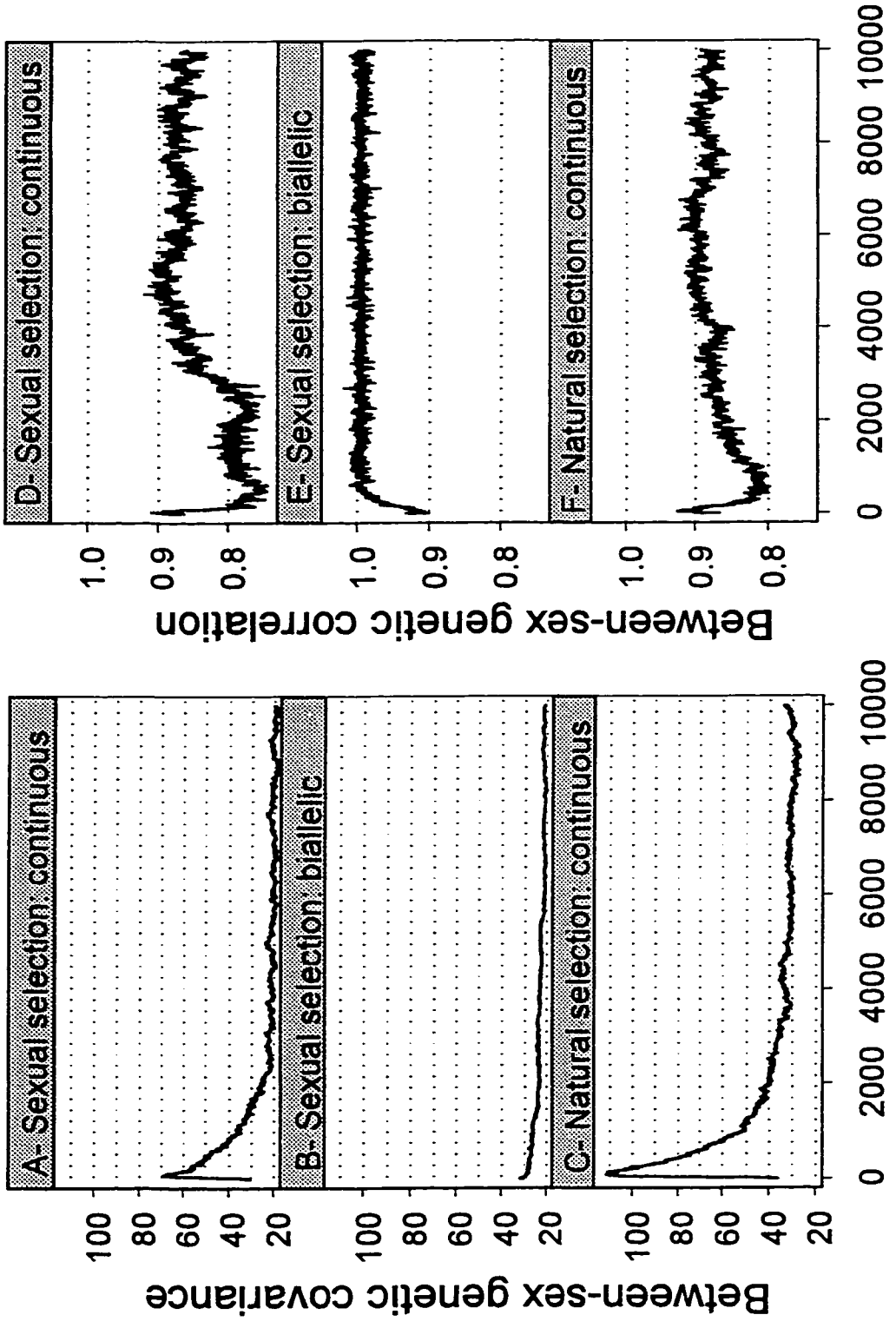

Fig. 3. Change in between-sex genetic covariances and correlations from the simulations of Fig. 1.

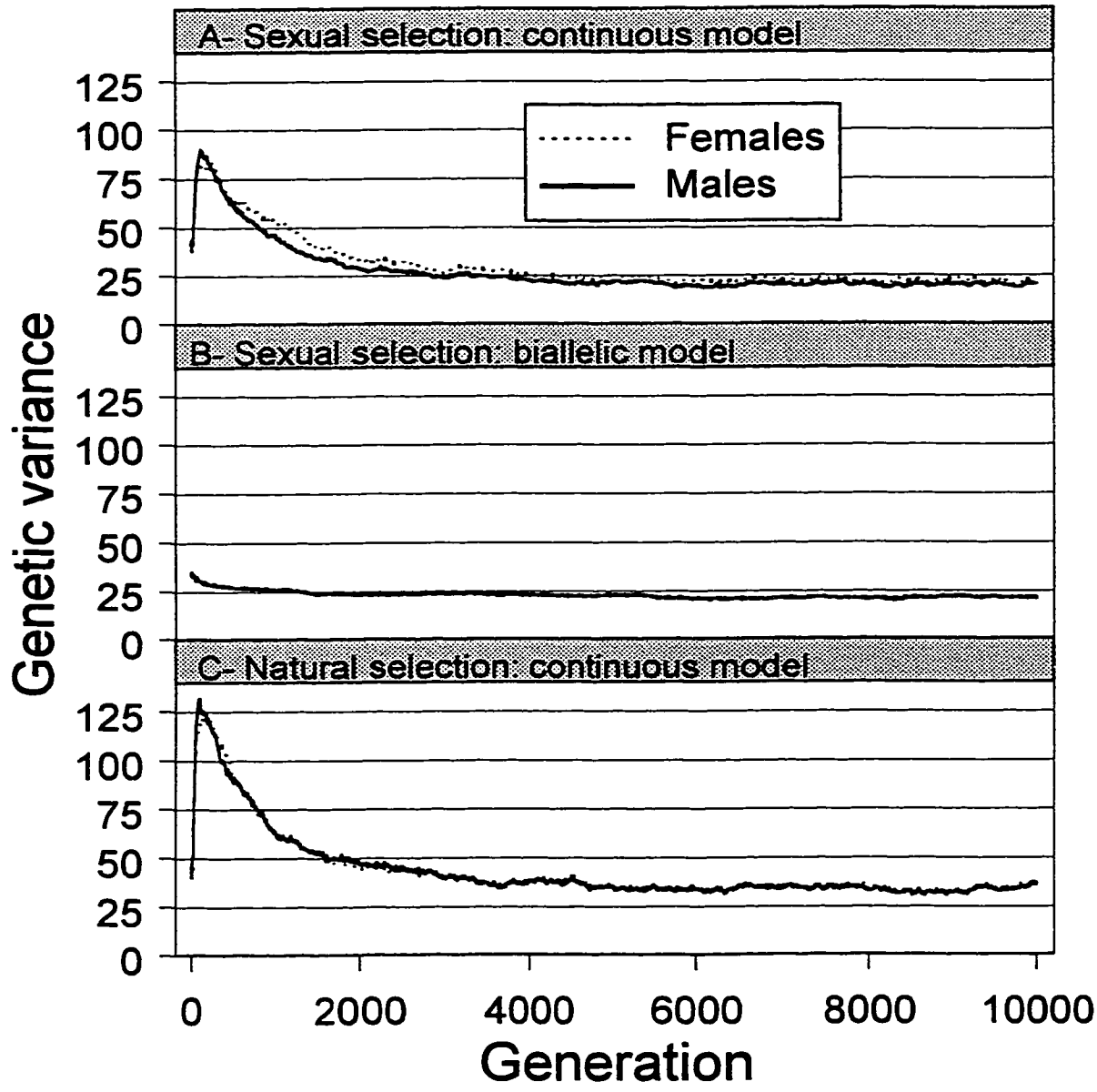

Fig. 4. Change in genetic variances of the two sexes from the simulations of Fig. 1.

between the sexes is defined as the genetic covariance between sexes divided by the product of the male and female standard deviations. Since changes in the genetic variances (Fig. 4A, C) are following a similar pattern to that of the covariances, large changes in dimorphism may be achieved without producing any permanent significant change in genetic correlations in the continuous models (Fig. 3D, F). Genetic covariance also decreases in the biallelic simulations (Fig. 3B), although here there is no initial increase in variance. The genetic correlation quickly increases to 1.0 (Fig. 3E), resulting in the lack of further response to selection seen in Fig. 1B. The correlation increases as the variance in the sex-specific loci becomes exhausted through fixation. At equilibrium, all the covariance between the sexes (Fig. 3B) is produced by variance at loci that are not sex-specific.

In the continuous simulations, the rapid increase in genetic variance (Fig. 4A, 4C) and genetic covariance between the sexes (Fig. 3A, 3C) seen in the early generations of directional selection, is due to the leptokurtic distribution of allelic effects at the initial equilibrium under house-of-cards conditions.

All further Figures and results refer to continuum-of-alleles models. We refer to the two strengths tested within each selection type as "strong" and "weak". Our two levels of sexual selection correspond to expected mating successes that differ between the largest and smallest members of the population by factors of 4.96 ($k = 0.8$ = "strong") and 1.66 ($k = 0.4$ = "weak") times. The strength of stabilizing selection is set at either 19 ("weak") or 9 ("strong") times the environmental variance. These values are within the range found in typical experimental estimates (Johnson, 1976; Turelli, 1984), and produced average initial (equilibrated base population) heritabilities in the continuous

simulations of 0.34 and 0.18 for weak and strong selection respectively. (That of the biallelic simulations of Fig. 1B was 0.33.)

Figure 5 compares the trajectories of predicted and observed SSD evolution for the two different strengths of stabilizing and sexual selection. Under strong sexual selection (Fig. 5A, C), equilibrium male size is larger than predicted, so the equilibrium SSD is also larger. Under weak sexual selection (Fig. 5B, D), the predictions for equilibrium SSD are very accurate. For all of the simulations, the rate at which SSD evolves in the early generations is faster than predicted. This accelerated rate of response is caused by an initial increase in genetic variance (Fig. 4A), and will be addressed in the Discussion.

Equilibrium size in males is determined by

$$\mathbf{Z}_m = \theta_m + (\mathbf{W}_m + \mathbf{P}_m)\mathbf{c}_m \qquad (7)$$

(Lande, 1980b), which can be rewritten as

$$\mathbf{Z}_m = \mathbf{W}_m\mathbf{P}_m^{-1}\mathbf{S}^*_m + \mathbf{S}^*_m \qquad (8)$$

since θ_m is equal to 0.0. Because $\mathbf{W}_m$ is constant, error in the predicted levels of SSD is caused by final values of the phenotypic variance ($\mathbf{P}_m$) and/or the sexual selection differential ($\mathbf{S}^*_m$) that are different from those estimated at generation zero. Figure 6 plots initial vs. final estimates of $\mathbf{P}_m$, $\mathbf{S}^*_m$, and $\mathbf{c}_m$, and observed vs. predicted values for equilibrium male body size ($\mathbf{Z}_m$). The final values are based on measurements taken every

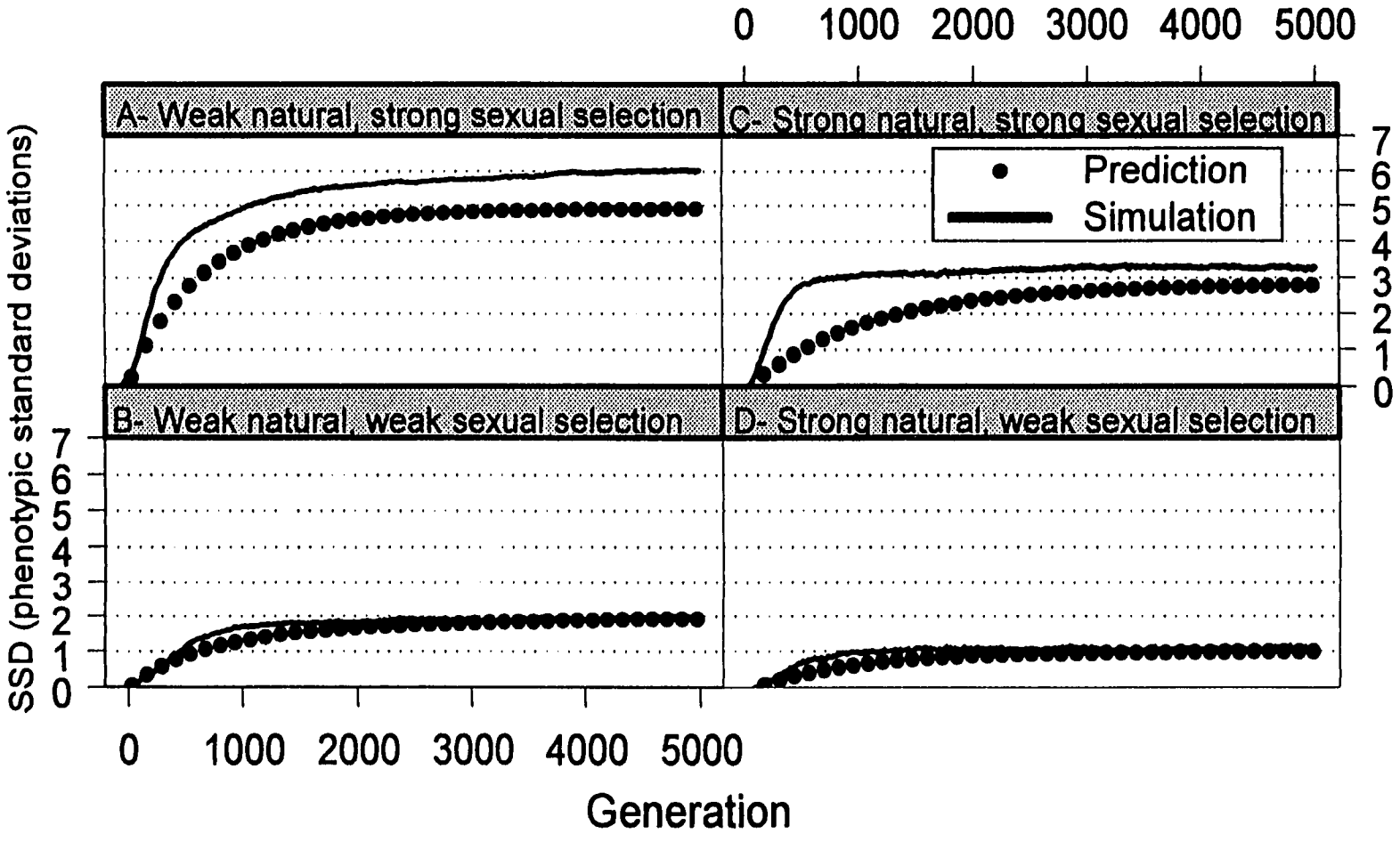

Fig. 5. Change in sexual size dimorphism (male size - female size) through frequency-dependent selection on male size. The sexual selection parameter (k) is either 0.8 (strong sexual selection, panels A, C) or 0.4 (weak sexual selection, panels B,D). Similary, the parameter for natural (stabilizing) selection (W) is either 9.0 (strong natural selection, panels C,D) or 19.0 (weak natural selection, A, B). Fig. 5A is from the simulations of Fig. 1A.

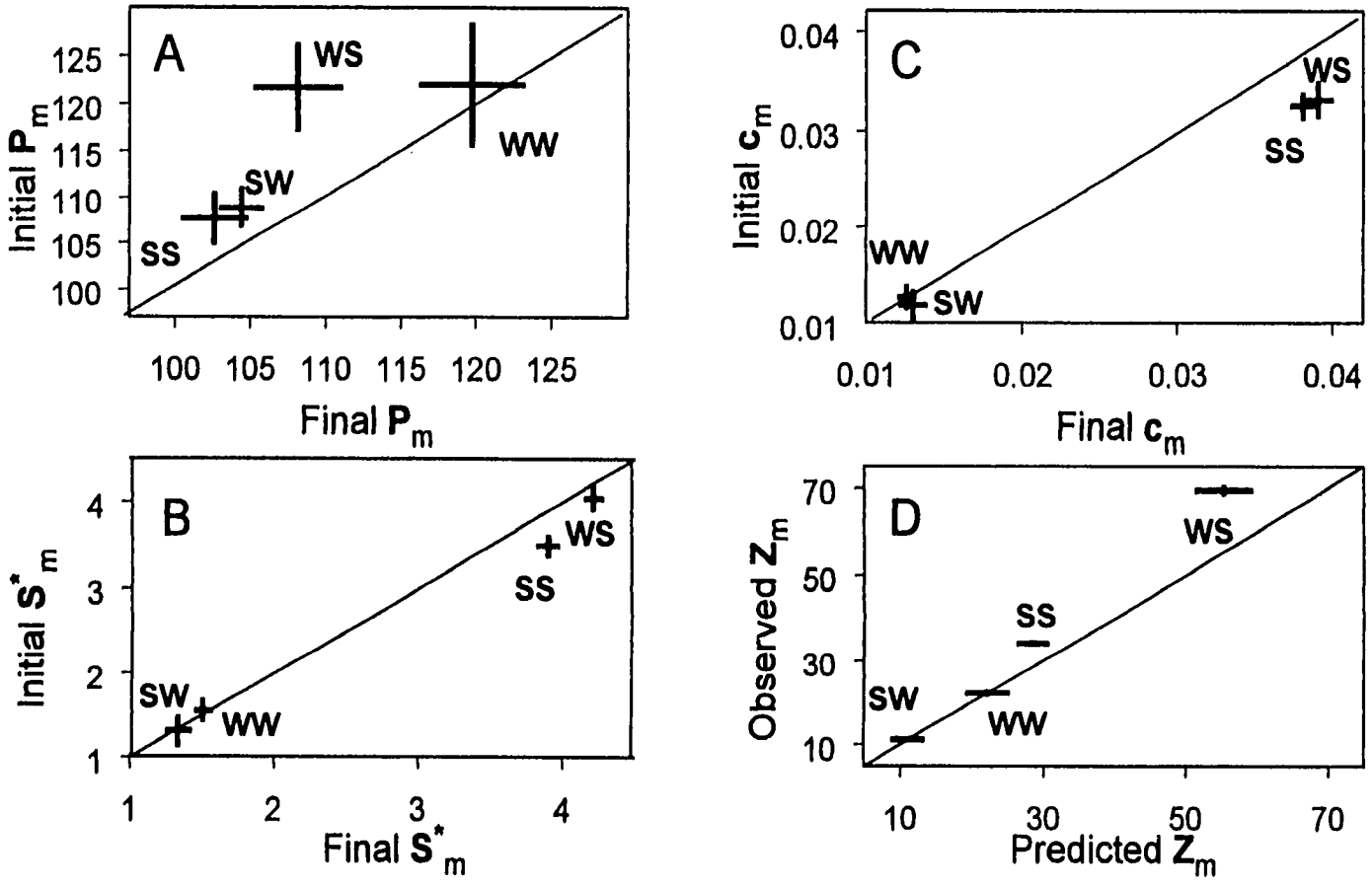

Fig. 6. Observed vs. predicted values for equilibrium male size, and initial vs. final estimates for variables that determine this equilibrium. $\mathbf{c}_m$ = intensity of sexual selection, $\mathbf{S}^*_m$ = sexual selection differential, $\mathbf{P}_m$ = phenotypic variance, and $\mathbf{Z}_m$ = equilibrium size. Data labels refer to the strength of natural selection, followed by the strength of sexual selection. w = weak selection (**W** = 19, k = 0.4), s = strong selection (**W** = 9, k = 0.8). These results are from the simulations shown in Fig. 5. Error bars indicate 1 standard deviation.

10 generations from 9900 to 10000, and are assumed to be estimates of the "true" equilibrium value of the variables. The initial values are based on the single set of estimates at generation zero, since that is the information on which the predictions are based. Here, the increased values for Z_{final} (and therefore equilibrium SSD) in the simulations with strong sexual selection (ws and ss in Fig. 6D) can be seen to be caused by an increase in $\mathbf{c}_m$ (Fig. 6C) compared to the generation zero estimate. This, in turn, is caused by the combination of an increase in $\mathbf{S}^*$ (Fig. 6B) and a decrease in $\mathbf{P}_m$ (Fig. 6A). The increase in $\mathbf{S}^*$ is due to the interaction between the two modes of selection. Frequency-dependent selection (equation 6) is an increasing function of rank, so its shape is independent of mean size. Frequency-independent selection (equation 5) changes from a concave downward to a near-linearly decreasing shape, as male mean size moves away from its optimum. The net result is that after the population mean has moved a small distance from its optimal value, for a constant $\mathbf{P}$, $\mathbf{S}^*$ increases as the population mean increases. This was also true for types of sexual selection other than the rank-selection used in our main simulations, e.g. truncation selection, and selection based on number of phenotypic standard deviations from the mean (results not shown). However, the phenotypic variance ($\mathbf{P}$) at the sexually selected equilibrium tends to decrease with increasing strength of sexual selection. This is because the two modes of selection produce balancing selection for the new equilibrium, rather than pure stabilizing selection as is found at generation zero. The resulting fitness function is steeper than that in the base population before sexual selection. Figure 7 shows this change for one of the replicates from Figure 5C. The strength of "stabilizing" selection at the sexually selected

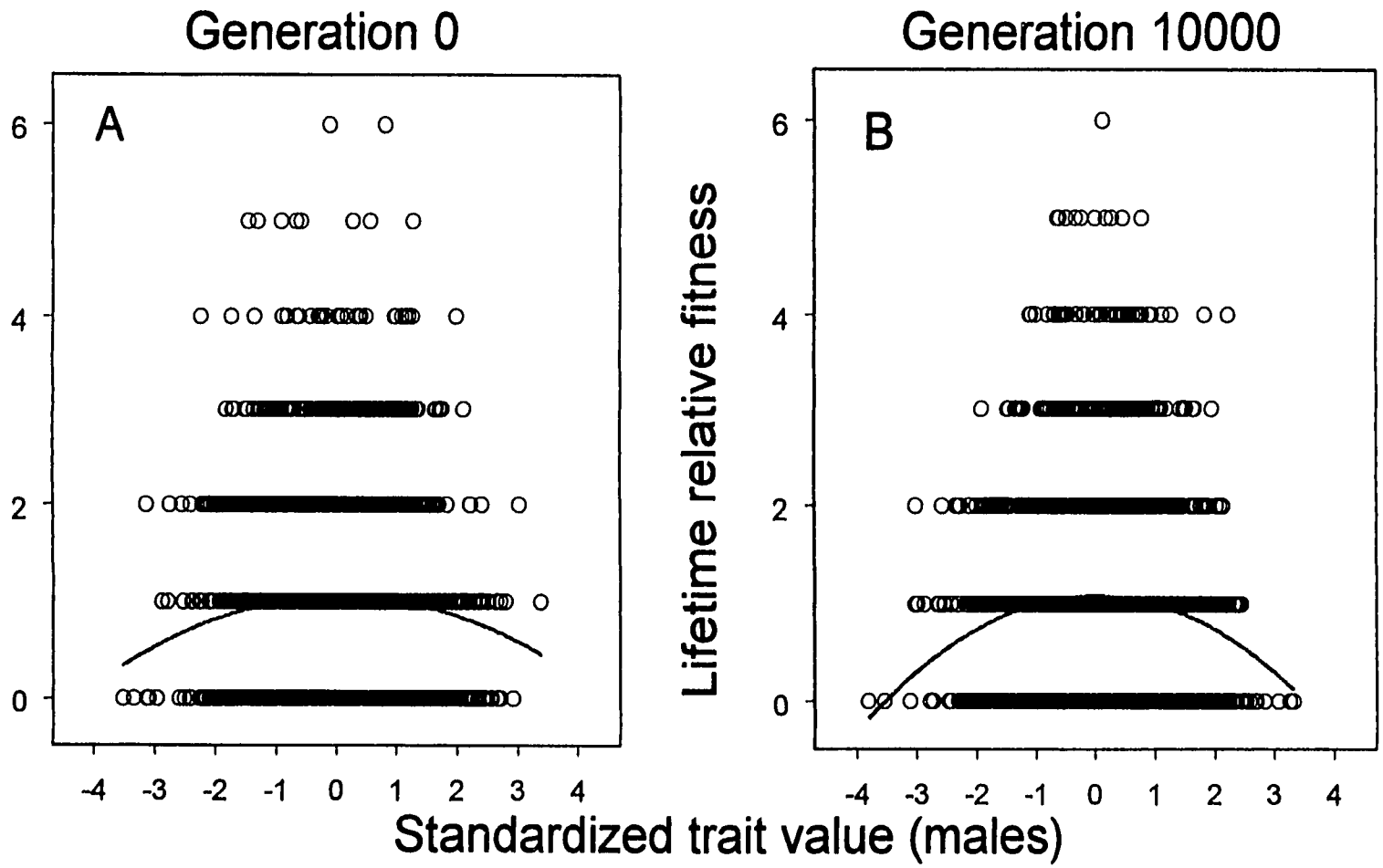

Fig. 7. Relative fitness vs. standardized male size for A) the base population, and B) the population at the sexually selected equilibrium, from a randomly chosen replicate from the simulations shown in Fig. 5C ($W = 9$, $k = 0.8$). The y-axis is also the number of times each male was selected to be the parent of two offspring (one of each sex). The solid line is the best fit quadratic approximation to the fitness surface.

equilibrium can be estimated through multiple regression methods (Lande and Arnold, 1983). In Fig. 7B, this estimate (-0.09) indicates stronger stabilizing selection at the sexually selected equilibrium than in the generation zero population of Fig. 7A (-0.05). (Note – these values are not equivalent to the coefficients of **W**, where larger magnitudes indicate weaker selection, and stabilizing selection has positive coefficients).

In the trajectories shown in Fig. 1, there is always a large displacement in female mean size, caused by the sudden increase in the male optimal size. However, if the optimal size in males is incremented by a small amount each generation (Fig. 8), female size shows much less displacement (less than one-sixth the amount seen in Fig. 1C). The trajectories produced when the intensity of sexual selection is incremented slightly each generation are very similar (results not shown). Thus, the maladaptive evolution of female size away from its optimum may largely be an artifact of the assumption (from Lande's (1980b) paper) that the shift in selective pressures in males is sudden and large.

DISCUSSION

Equilibrium SSD levels were predicted very accurately for weak sexual selection, but under stronger sexual selection, the discrepancy between predicted and simulated SSD increased. In addition, the rate of separation between the sexes in the initial generations of selection was much faster than predicted. Our results suggest that genetic correlations constrain both the short-term and long-term evolution of SSD less than predicted by the Lande model, if population sizes are relatively large. The overall quality of the predictions, however, is surprisingly good, given the sometimes large changes that are occurring in the genetic and phenotypic variances.

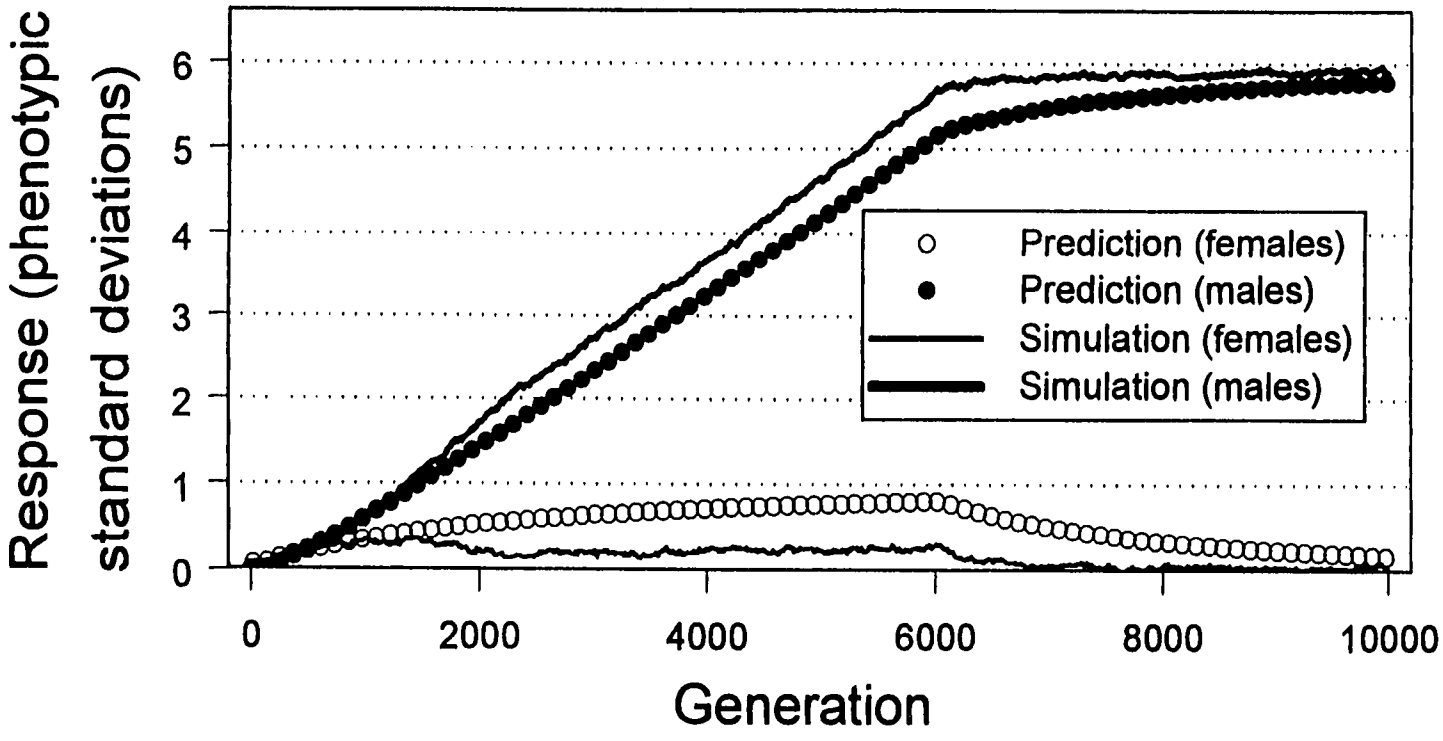

Fig. 8. Response to frequency-independent selection for increased male size. W = 19, k = 0.0, $\theta_f = 0.0$. The value of the natural selection optimum for males (θ_m) is 0.0 at generation zero, and is incremented by 0.001 phenotypic standard deviations each generation until generation 6000. The predictions are based on equations that have been given this information about the changing location of θ_m.

Since we know very little about the genetic basis of quantitative traits, little can be said about the relative likelihood of the alternative genetic models shown in Fig. 1A,B. Although the importance of the genetic assumptions has been discussed in great detail in the context of the maintenance of genetic variance under stabilizing selection, their significance in terms of predictions of long-term directional selection has received relatively little attention (Barton and Turelli, 1987; Burger, 1993; Reeve, 2000). If only a small number of alleles are possible at each locus, as in our biallelic simulations, permanent or effectively permanent constraints to phenotypic evolution are likely to be very common. However, in terms of testing predictions through simulation modelling, we can say only that if the genetic limit is reached in such models, no further response is possible. For this reason, we will confine our discussion primarily to the results of the continuous models.

The accelerated response in male size, and therefore SSD, in the early generations of selection is related to changes in the distributional properties of allelic effects. It is known from earlier studies (Barton and Turelli, 1987; Reeve, 2000) that under house-of-cards conditions, directional selection can lead to an initial increase in genetic variance. This occurs because, at mutation-stabilizing selection-drift balance (our populations at generation zero), the allelic effects at each locus will be normally distributed only if mutation rates are very high ("Gaussian" conditions). Most quantitative genetics equations for predicting response to selection rely on the assumption of normally distributed allelic effects (Barton and Turelli, 1987). However, under house-of-cards equilibrium conditions, these effects tend to be leptokurtically distributed. Most alleles have values near the mean value for that locus, but some rare alleles will have values quite

distant from the mean. With directional selection, favourable rare alleles are selected, and as they increase in frequency, genetic variance can increase dramatically. When a new equilibrium is reached, those once rare alleles (or their descendents) become the norm, and allelic effects again become leptokurtically distributed about the loci's new mean values. The initial rise in genetic variance causes an increase in phenotypic variance, and a response that is faster than expected from base population parameters. Such accelerated responses are seldom seen in selection experiments, but this could be due to a lack of rare alleles caused by small effective population sizes.

Fisher (1958) and Lande (1980b) suggest that the genetic correlation between the sexes for sexually selected traits should diminish greatly with time. However, this is not required (or permitted) in Lande's model, and it does not occur in our simulations. A reduction in the between-sex genetic correlation could occur through the "capture" of new genes by either sex, or through sex-limitation in genes that were previously expressed equally in both sexes. Wright (1993) has suggested that there may be a third phase to SSD evolution in some organisms, in addition to the two discussed in Lande (1980b). In this third phase, selection for newly evolved sex-specific genes (e.g. for sex-hormones or their receptors in vertebrates) could increase the rate at which SSD evolves. As our simulation model does not allow for the evolution of the genetic architecture itself (i.e. which genes control which sex), we cannot address this question. However, our simulations show that rapid rates of SSD evolution can occur even in a very simplistic model with no provisions for such genetic changes.

The consequences of non-constant **G** and **P** matrices in quantitative genetics models such as equation 4 have not been adequately explored. In our simulation, changes

in the genetic and phenotypic variance not only cause the accelerated initial response, but contribute to the attainment of an equilibrium male size, and hence SSD, that is greater than predicted. In populations undergoing strong sexual selection, the difference between the initial and equilibrium values of c_m was due not only to changes in the phenotypic variance, but to changes in the shape of the net fitness function acting over the two episodes of selection (viability and sexual). Therefore, even if we assume that the functional relationship between body size and mating success (k) remains constant over time, it is highly likely that the actual intensity of sexual selection (c_m), increases as the mean sizes of the two sexes diverge.

Whether the changes in variances are permanent depends on the mode of selection. For equilibrium shifts caused solely by frequency-independent natural selection, there is no permanent change in the magnitude of the net selective forces acting on the population. Therefore, the genetic variance can return to the value present before directional selection. In contrast, frequency-dependent sexual selection causes a decrease in genetic variance that is a function of the relative strengths of the two modes of selection and the distance that equilibrium male size has shifted. Although it is only the male lifetime fitness function that is permanently altered by sexual selection, female genetic and phenotypic variances at equilibrium are also lowered. This is because all the genes from fathers are passed on to both sons and daughters. However, in our simulations females have three loci that never directly experience the effects of the increased selection intensity on males, since they do not contribute to the male's phenotype. The larger equilibrium genetic variance in females (for instance, ~10% higher than in males for the simulations of Fig. 4A) is due entirely to the increased variance present at these female-specific loci.

In addition to the results shown in this paper, many simulations were run with different starting conditions. When population sizes of N= 400 were used, patterns of response were quite different. There are two main characteristics of small populations which combine to produce these differences. First, small populations have many fewer rare alleles segregating at equilibrium. Therefore the increase in variance and the accelerated response of the means, declines as population size decreases. Second, as there is a smaller amount of standing variance in small populations, it is more quickly eroded, and response to continued directional selection quickly becomes dependent on input from new mutations. This is much slower than the response due to standing variance. The net result is that for small populations, response is predicted fairly accurately for a small number of generations, but then tends to lag behind the predicted response if sexual selection is strong. Equilibrium SSD in such cases takes much longer to reach than predicted by equation 4. Even with large populations, this lag effect can be produced by increasing the intensity of sexual selection, although here the initial accelerated response is still present. For all population sizes tested, equilibrium SSD levels were underestimated by equation 4 when there was strong selection.

The effect of antagonistic pleiotropy was tested by including loci whose effect was of opposite sign in males and females. This produced no qualitative differences in the results. Likewise, no qualitative differences were found in simulations that included six traits (genetically correlated within and between sexes), with sexual selection acting on one of the male traits.

Even under conditions where the accuracy of predictions was very good, it must be remembered that the results are based on an average of ten replicates. It can be seen

from Fig. 6D that there is much more variance in the predicted than in the observed equilibrium values for male size. This is mainly caused by random fluctuations in the variances in the base populations, and the effect that these have on estimates of c_m. Even in a population of size N=4000, there is a great deal of variance in the variances, a problem that has been dealt with at length elsewhere (e.g. Burger et. al., 1989; Keightley and Hill, 1989).

For the simulated replicates, equilibrium values are a function of the characteristics of the genetic architecture of the population, not the value of the variables that happen to be present at generation zero. This architecture is determined by a combination of the fixed parameters (e.g. **W** , N, and the mutation rate), and the "developmental rules" that specify which genes control size in each sex. In multi-trait systems, these rules also determine the pleiotropic relationships between traits. In our simulations, these rules are determined by the number of loci that contribute to both sexes or to just males or just females. The rationale for not allowing mutations that change this architecture is that such developmental rules are likely to be more strongly canalized than are the loci that contribute to normal quantitative variation (Wagner, 1989).

Equation 4 cannot and should not be expected to predict long-term evolution in real populations. There are numerous reasons why this is an unrealistic goal. Most of the variables cannot be estimated with reasonable accuracy. Even if they could be, environmental changes probably cause **W** and θ to fluctuate constantly over time. The fitness surface itself is unlikely to ever be Gaussian, and may not always be unimodal. In addition, there will always be large numbers of traits under selection, and most of these will not be included in the analysis. Therefore, the results from our simulation analysis

should not be viewed as a test of how well the equation can predict evolution, but of how robust the equation's predictions are to violations of assumptions caused by house-of-cards equilibrium conditions in finite populations. This does not mean that the equations are of no practical importance. They, like other predictive models in quantitative genetics, should be viewed as a starting point for trying to understand and possibly quantify the evolutionary forces that were responsible for current patterns of trait distributions. The greatest impediment to predicting long term response in real populations is our ignorance of patterns of temporal variation in the forces of selection, rather than the relatively minor error caused by changes in the **G** matrix.

General Conclusions

In this thesis, I have used a selection experiment and simulation modelling in an attempt to answer several questions of relevance to the problem of SSD evolution.

Can fecundity selection alter SSD?

The selection experiment of Chapter 1 is the first experimental demonstration of Darwin's fecundity selection hypothesis of SSD evolution. Selection for increased fecundity caused an increase in body size (abdomen and thorax width) in both sexes, with the response being relatively greater in females. The change in SSD, which was statistically significant, occurred despite the fact that the direct response in fecundity was smaller than that of either width measurement (in phenotypic standard deviations), and was not itself significant.

What is the relationship between Lande's equations, the simulation models, and evolution in real populations?

Before summarizing the conclusions from the simulation work in this thesis, it is worth stating what should and should not be expected of the equations and simulations. Lande's equations for predicting the evolution of trait means cannot be adequately tested in populations of real organisms, since a lack of agreement between observed and predicted values could have several causes. These fall into three main categories:

(1) Insufficient data. Examples include inaccurate parameter estimates, selectively important traits missing from the analysis, changing environmental conditions, etc. A discrepancy between prediction and observation from these causes is not a shortcoming of the equations, other than in a practical sense. Given perfect information, the predictions could in theory be perfect. However, since it can never be guaranteed that all the relevant

information is known with sufficient accuracy, the predictions from the equations are largely non-falsifiable.

(2) Genetic complexity. Lande's equations assume additivity among alleles at the same locus (no dominance) and among loci (no epistasis). His equations are purely phenotypic, in the sense that the only variable that can change over time is the vector of trait means. Genetic details are not stated or required, other than that for additivity. In real populations, factors such as epistasis, genes of major effect, gene duplication, and heterochrony, to name only a few, are bound to complicate the course of long-term evolution.

(3) Constancy of genetic and phenotypic distributions. These distributions, assumed to be constant in the equations, are likely to vary over time in finite populations. The constancy required by the equations is assured by an assumption of infinite populations with infinite numbers of loci per trait. This allows use of the central limit theorem and normal-distribution theory to produce tractable analytical solutions.

My simulation models were specifically designed to test how well the response equations are able to cope with the violation of constancy assumptions described in (3) above. The problems associated with category (1) were eliminated, since all relevant variables and parameters were known, and could be estimated with great accuracy. Problems associated with category (2) were avoided because the simulation model was purely additive. Therefore, the simulations should be viewed not as a test of how well Lande's equations predict evolution, but of how robust the predictions are to complications brought about by finite populations. Equation (3) of chapter two, and equation (4) of chapter three have never been tested in real populations, and probably

never will be. Even the relatively simple equation (2) of chapter 2 has only been tested once (Grant and Grant, 1995) and that was only for a single generation of selection. For the reasons outlined in category (1) above, Lande's equations should be considered as explanatory rather than predictive models of phenotypic evolution if more than a single generation of selection is being considered.

Do Lande's equations for frequency-dependent sexual selection accurately predict equilibrium SSD?

When sexual selection is weak and there is little change in the value of the male trait, the predictions can be remarkably accurate. As selection becomes stronger, the predictions become increasingly inaccurate, and underestimate the total male response. This, in turn, results in predicted equilibrium SSD's that are smaller than those observed in the simulations. The cause of this discrepancy is the fact that the "strength" of sexual selection increases with movement away from the frequency-independent optimum, even when the relationship between mating success and relative phenotype remains constant. The increase in strength of sexual selection is caused by a shift in the difference between weighted phenotypes before and after sexual selection.

Do Lande's equations accurately predict response in the initial generations of selection?

The predictions generally underestimate the rate of simulated SSD evolution in the "initial" generations of sexual selection. This is because the distribution of allelic effects under a house-of-cards model is highly leptokurtic. For a given allelic variance, a leptokurtric distribution will have more evolutionary potential than will a normal distribution. However, depending on the exact parameters used, the accelerated response may not start until generation 15-40. Before this point, the response is not accelerated

because it takes some time for the initially rare alleles to increase to a frequency high enough to contribute greatly to the variance. This is true for sexual selection and for peak-shift frequency-independent selection. Therefore, even very large artificial selection experiments should not be expected to see such accelerated responses to directional selection unless they are run for at least 30 generations, which is seldom the case.

Should we expect constraints on SSD to be permanent?

Whether constraints on the evolution of SSD are likely to be permanent depends critically on the variability of the genes that control the traits. This can be seen in the simulations as the difference between final equilibrium values in the biallelic vs continuous models. If there are only a small number of possible states typically possible at each locus, then male and female body sizes may not be able to diverge to the extent predicted by the equations. If the continuous alleles model of Crow and Kimura (1964) is a reasonable representation of mutational potential, genetic constraints to the evolution of SSD should be less important than predicted, over both the short- and long-term.

Will female traits always be shifted greatly from their optima, when sexual selection acts on highly genetically correlated male traits?

Large shifts in female trait values are a consequence of the assumption (from the diagrams in Lande's [1980b] paper) that the onset of sexual selection in males is sudden and strong. This same assumption can be seen in the diagrams from Zeng (1988) for the correlated response of traits within a sex to a shift in the optimal values of one or more genetically correlated traits. In either case, if the onset of novel selection is instead gradual and weak, there is much less total movement of the unselected traits away from their optima. In the case of sexual selection, this means that it is possible to have a large shift in

male size without having a major effect on female size.

Summary

The results of all three chapters of this thesis suggest that SSD may be able to evolve much more quickly than suggested by theory. The selection experiment of chapter one showed that it may be possible to produce significant changes in SSD without any direct selection on body size itself. The simulation models of chapters two and three showed that given realistic estimates for mutation rates and mutational variance, directional selection causes an initial increase in genetic variance. This, in turn, leads to an accelerated response in the trait means, and a faster approach to equilibrium than predicted.

Literature Cited

Abouheif, E. and D. J. Fairbairn. 1997. A comparative analysis of allometry for sexual size dimorphism: assessing Rench's rule. *American Naturalist* 149: 540-562.

Alatalo, R. V., L. Gustafsson, and A. Lundberg. 1990. Phenotypic selection on heritable size traits, environmental variance and genetic response. *American Naturalist* 135: 464-471.

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

Andersson, M. 1994. *Sexual Selection*. Princeton University Press, Princeton, New Jersey.

Arak, A. 1988. Sexual dimorphism in body size: a model and a test. *Evolution* 42: 820-825.

Arnold, S. J. 1992. Constraints on phenotypic evolution. *American Naturalist* 140(Suppl.): s85-s107.

Atchley, W. R., and B. K. Hall. 1991. A model for development and evolution of complex morphological structures. *Biological Review* 66: 101-157.

Baatz, M., and G. P. Wagner. 1997. Adaptive inertia caused by hidden pleiotropic effects. *Theoretical Population Biology* 51: 49-66.

Barton, N. H. 1986. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genetical Research* 47: 209-216.

Barton, N. H. 1989. The divergence of a polygenic system subject to stabilizing selection, mutation and drift. *Genetical Research* 54: 59-77.

Barton, N. H., and M. Turelli. 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genetical Research* 49: 157-173.

Bell, A. E., C. H. Moore, and D. C. Warren. 1957. The evaluation of new methods for the improvement of quantitative characteristics. *Cold Spring Harbour Symposium of Quantitative Biology* 20: 197-212.

Bird, M. A., and H. E. Schaffer. 1972. A study of the genetic basis of the sexual dimorphism for wing length in *Drosophila melanogaster*. *Genetics* 72: 475-487.

Bjorklund, M. 1996. The importance of evolutionary constraints in ecological time scales. *Evolutionary Ecology* 10: 423-431.

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

Botella, L. M., A. Moya, A. C. Gonzalez, and J. L. Mensua. 1985. Larval stop, delayed development and survival in overcrowded cultures of *Drosophila melanogaster*: effect of urea and uric acid. *Journal of Insect Physiology* 31: 179-185.

Bulmer, M. G. 1972. The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genetical Research* 19: 17-25.

Bulmer, M. G. 1980. *The Mathematical Theory of Quantitative Genetics*. Oxford University Press. Oxford, U.K.

Bulmer, M. G. 1989. Maintenance of genetic variability by mutation-selection balance, a child's guide through the jungle. *Genome* 31: 761-767.

Burger, R. 1986. Constraints for the evolution of functionally coupled characters, a nonlinear analysis of a phenotypic model. *Evolution* 40: 182-193.

Burger, R. 1993. Predictions of the dynamics of a polygenic character under directional selection. *Journal of Theoretical Biology* 162: 487-513.

Burger, R., G. P. Wagner, and F. Stettinger. 1989. How much heritable variation can be maintained in finite populations by mutation-selection balance? *Evolution* 43: 1748-1766.

Burger, R., and R. Lande. 1994. On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. *Genetics* 138: 901-912.

Burger, R., and M. Lynch. 1995. Evolution and extinction in a changing environment, a quantitative-genetic analysis. *Evolution* 49: 151-163.

Charlesworth, B. 1990. Optimization models, quantitative genetics, and mutation. *Evolution* 44: 520-538.

Charlesworth, B. 1993. Directional selection and the evolution of sex and recombination. *Genetical Research* 61: 205-224.

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.

Conner, J. K., Rush, S., and Jennetten, P. 1996a. Measurements of natural selection on floral traits in Wild Radish (*Raphinus raphanistrum*) I. Selection through lifetime female fitness. *Evolution* 50: 1127-1136.

Conner, J. K., Rush, S., Kercher, S., and Jennetten, P. 1996b. Measurements of natural selection on floral traits in Wild Radish (*Raphanus raphanistrum*) II. Selection through lifetime male and total fitness. *Evolution* 50: 1137-1146.

Crow, J. F., and M. Kimura. 1964. The theory of genetic loads. *Proceedings of the XI*

International Congress of Genetics pp. 495-505.

Darwin, C. 1871. *The Descent of Man, and Selection in Relation to Sex.* John Murray, London.

De Moed, G. H., C. L. J. J. Kruitwagan, G. de Jong, and W. Scharloo.1999. Critical weight for the induction of pupariation in *Drosophila melanogaster*: genetic and environmental variation. *Journal of Evolutionary Biology* 12: 852-858.

Deng, H.-W., and T. T. Kibota. 1995. The importance of the environmental variance-covariance structure in predicting evolutionary responses. *Evolution* 49: 572-574.

Eisen, E. J., and J. P. Hanrahan. 1972. Selection for sexual dimorphism in body weight of mice. *Australian Journal of Biological Science* 25: 1015-1024.

Endler, J. A. 1986. *Natural Selection in the Wild.* Princeton University Press, Princeton, NJ.

Fairbairn, D. J. 1997. Allometry for sexual size dimorphism: pattern and process in the coevolution of body size in males and females. *Annual Review of Ecology and Sytematics* 28: 659-687.

Falconer, D. S. 1989. *Introduction to Quantitative Genetics.* Longmans, New York.

Ferguson, I. M., and Fairbairn, D. J. 2000. Sex-specific selection and sexual size dimorphism in the waterstrider *Aquarius remigis. Journal of Evolutionary Biology* 13: 160-170.

Fisher, R. A. 1958. *The Genetical Theory of Natural Selection.* Second edition. Dover Publ., New York.

Gowe, R. S., and R. W. Fairfull. 1990. Genetics controls in selection. In: Crawford, R.D. (ed.) *Poultry Breeding and Genetics 1990:* pp. 935-954. Elsevier, Amsterdam.

Gowen, J. W., and L. E. Johnson. 1946. On the mechanism of heterosis. I. Metabolic capacity of different races of *Drosophila melanogaster* for egg production. *American Naturalist* 80: 149-179.

Grant, P. R. and B. R. Grant. 1995. Predicting microevolutionary responses to directional selection on heritable variation. *Evolution* 49: 241-251.

Gromko, M. H. 1995. Unpredictability of correlated response to selection: pleiotropy and sampling interact. *Evolution* 49: 685-693.

Head, G. 1995. Selection on fecundity and variation in the degree of sexual size dimorphism among spider species (class Araneae). *Evolution* 49: 776-781.

Hedrick, A. V., and E. J. Temeles. 1989. The evolution of sexual dimorphism in animals: hypotheses and tests. *Trends in Ecology and Evolution* 4: 136-138.

Herrara, C. M. 1993. Selection on floral morphology and environmental determinants of fecundity in a hawk moth-pollinated violet. *Ecological Monographs* 63: 251-275.

Hill, W. 1982. Rates of change in quantitative traits from fixation of new mutations. *Proceedings of the National Academy of Sciences U.S.A.* 79: 142-145.

Hill, W. G., and P. D. Keightley. 1988. Interrelations of mutation, population size, artificial and natural selection. Pp. 57-70 In *Proceedings of the Second International Conference on Quantitative Genetics*. Edited by E.J. Eisen, M.M. Goodman, G. Namkoong, and B.S. Weir. Sinauer Associates, Inc., Sunderland, MA.

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

Houle, D. 1989. The maintenance of polygenic variation in finite populations. *Evolution*

43: 1767-1780.

Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. *Genetics* 143: 1367-1483.

Johnston, C. 1976. *Introduction to Natural Selection.* Baltimore: University Park Press.

Keightley, P. D., and W. G. Hill. 1987. Directional selection and variation in finite populations. *Genetics* 117: 573-582.

Keightley, P. D., and W. G. Hill. 1988. Quantitative genetic variability maintained by mutations-stabilizing selection balance in finite populations. *Genetical Research* 52: 33-43.

Keightley, P. D., and W. G. Hill. 1989. Quantitative genetic variability maintained by mutation-stabilizing selection balance: sampling variation and response to subsequent directional selection. *Genetical Research* 54: 45-57.

Kirkpatrick, M., and R. Lande. 1989. The evolution of maternal effects. *Evolution* 43: 485-503.

Lande, R. 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical Research* 26: 221-235.

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

———. 1980a. Genetic variation and phenotypic evolution during allopatric speciation. *American Naturalist* 116: 463-479.

———. 1980b. Sexual dimorphism, sexual selection and adaptation in polygenic characters. *Evolution* 34: 292-305.

———. 1987. Genetic correlations between the sexes in the evolution of sexual size

dimorphism and mating preferences. Pp. 83-94 *in* J. W. Bradbury and M. B. Anderson, editors. *Sexual selection: testing the alternatives*. Wiley, New York.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37: 1210-1226.

Lande, R., and S. J. Arnold. 1985. Evolution of mating preference and sexual dimorphism. *Journal of Theoretical Population Biology* 117: 651-664.

Lascoux, M. 1997. Unpredictability of correlated response to selection: linkage and initial frequency also matter. *Evolution* 51: 1394-1400.

Latter, B. D. H. 1960. Natural selection for an intermediate optimum. *Australian Journal of Biological Sciences* 13: 30-35.

Loeschcke, V. 1987. (Editor) *Genetic constraints on adaptive evolution*. Springer-Verlag, New York.

Lynch, M. 1988. The rate of polygenic mutation. *Genetical Research* 51: 137-148.

Lynch, M., and R. Lande. 1993. Evolution and extinction in response to environmental change, pp. 234-250 in Biotic Interactions and Global Change edited by P. M. Kareiva, J. G. Kingsolver and R. B. Huey. Sinauer, Sunderland, Mass.

Maynard Smith, J., R. Burian, S. Kauffman, P. Albrech, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert. 1985. Developmental constraints and evolution. *Quarterly Review of Biology*: 60 265-287.

Meagher, T. R. 1994. The quantitative genetics of sexual size dimorphism in *Silene latifolia* (Caryophyllaceae). II. Response to sex-specific selection. *Evolution* 48: 939-951.

Moore, A. J. 1990. The evolution of sexual dimorphism by sexual selection: the separate

of effects of intrasexual selection and interspecific selection. *Evolution* 44:315-331.

Narain, P., C. Joshi, and S. S. Prabhu. 1962. Response to selection for fecundity in *Drosophila melanogaster*. *Drosophila Information Service* 36: 97-99.

Nunney, L. 1996. The response to selection for fast larval development in *Drosophila melanogaster* and its effect on adult weight: an example of a fitness trade-off. *Evolution* 50: 1193-1204.

Orozco, F., and A. E. Bell. 1974. Reciprocal recurrent selection compared to within-strain selection for increasing rate of egg lay of *Tribolium* under optimal and stress conditions. *Genetics* 77: 143-161.

Parker, G. A. 1992. The evolution of sexual size dimorphism in fish. *Journal of Fish Biology* 41: 1-20.

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

Partridge, L., R. Langelan, K. Fowler, and V. French. 1999. Correlated responses to selection on body size in *Drosophila melanogaster*. *Genetical Research* 74: 43-54.

Preziosi, R. F., D. J. Fairbairn, D. A. Roff, and J. M. Brennan. 1996. The functional influence of body size on reproduction in the waterstrider *Aquarius remigis*. *Oecologia* 108: 424-431.

Preziosi, R. F., and D. J. Fairbairn. 1997. Sexual size dimorphism and selection in the wild in the waterstrider *Aquarius remigis*: lifetime fecundity selection on female total length and its components. *Evolution* 51: 467-474.

Preziosi, R. F., and D. J. Fairbairn. *in press*. Lifetime selection on adult body size in a

waterstrider: opposing selection and maintenance of sexual size dimorphism. *Evolution*.

Preziosi, R. F., and D. A. Roff. 1998. Evidence of genetic isolation between sexually monomorphic and sexually dimorphic traits in the water strider *Aquarius remigis*. *Heredity* 81: 92-99.

Price, D. K. 1996. Sexual selection, selection load and quantitative genetics of zebra finch bill colour. *Proceedings of the Royal Society of London B* 263: 217-221.

Price, D. K., and N. T. Burley. 1993. Constraints on the evolution of attractive traits: genetic (co)variance of zebra finch bill colour. *Heredity* 71: 405-412.

Price, T. D. 1984. Sexual selection on body size, territory and plumage variables in a population of Darwin's finches. *Evolution* 38: 327-341.

Price, T., M. Kirkpatrick, and S. J. Arnold. 1988. Directional selection and the evolution of breeding date in birds. *Science* 240: 798-799.

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

Rausher, M. D. 1992. The measurement of selection on quantitative traits, biases due to environmental covariances between traits and fitness. *Evolution* 46: 616-626.

Reiss, M. J. 1989. *The Allometry of Growth and Reproduction*. Cambridge University Press, Cambridge.

Reeve, J. P. 2000. Predicting long term response to selection. *Genetical Research* 75: 83-94.

Reeve, J. P., and D. J. Fairbairn. 1996. An empirical test of models for the evolution of sexual size dimorphism as a correlated response to selection on body size.

Evolution 50: 1927-1938.

Richardson, R. H., and K.-I. Kojima. 1965. The kinds of genetic variability in relation to selection responses in *Drosophila* fecundity. *Genetics* 52: 583-598.

Robertson, A. 1970. Some optimum problems in individual selection. *Theoretical Population Biology* 1: 120-127.

Robertson, F. W. 1957. Studies in quantitative inheritance. XI. Genetic and environmental correlation between body size and egg production in *Drosophila melanogaster*. *Journal of Genetics* 55: 428-443.

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

Roff, D. A. 1997. *Evolutionary Quantitative Genetics*. Chapman & Hall, New York.

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

Schlichting, C. D., and M. Pigliucci. 1998. *Phenotypic Evolution: A Reaction Norm Perspective*. Sinauer, New York.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50: 1766-1774.

Selander, R. K. 1972. Sexual selection and dimorphism in birds. *In* B. G. Campbell (editor). *Sexual selection and the descent of man, 1871-1971*. Aldine, Chicago.

Shaw, F. H., R. G. Shaw, G. S. Wilkinson, and M. Turelli. 1995. Changes in genetic variances and covariances, G whiz! *Evolution* 49: 1260-1267.

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

Shine, R. 1989. Ecological causes for the evolution of sexual dimorphism: a review of the evidence. *Quarterly Review of Biology* 64: 419-461.

Shiotsugu, J., A. M. Leroi, H. Yashiro, M. R. Rose, and L. D. Mueller. 1997. The symmetry of correlated selection responses in adaptive evolution: an experimental study using *Drosophila*. *Evolution* 51: 163-172.

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

Slatkin, M. 1987a. Heritable variation and heterozygosity under a balance between mutation and stabilizing selection. *Genetical Research* 50: 53-62.

Slatkin, M. 1987b. Quantitative genetics of heterochrony. *Evolution* 41: 799-811.

Sokal, R. R., and F. J. Rohlf. 1995. *Biometry*, 3rd edn. W. H. Freeman and Company. New York

SPSS Advanced Statistics 7.5. 1997. SPSS Inc., Chicago.

Tait, W. M., and S. S. Prabhu. 1970. Fecundity and hatchability in *Drosophila melanogaster* with a second chromosome held unchanged. *Journal of Genetics* 60: 152-158.

Turelli, M. 1984. Heritable genetic variation through mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology* 25: 138-193.

Turelli, M., and N. H. Barton. 1990. Dynamics of polygenic characters under selection. *Theoretical Population Biology* 38: 1-57.

van Tienderen, P. H., and G. de Jong. 1994. A general model of the relation between phenotypic selection and genetic response. *Journal of Evolutionary Biology* 7: 1-12.

Via, S., and R. Lande. 1985. Genotype-environment interaction and the evolution of

phenotypic plasticity. *Evolution* 39: 505-522.

Wagner, G. P. 1988. The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *Journal of Evolutionary Biology* 1: 45-66.

Wagner, G. P. 1989. Multivariate mutation-selection balance with constrained pleiotropic effects. *Genetics* 122: 223-234.

Webster, M. S. 1992. Sexual dimorphism, mating system and body size in New World blackbirds (Icterinae). *Evolution* 46: 1621-1641.

Weis, A. E. 1996. Variable selection on Eurosta's gall size, III, Can an evolutionary response to selection be detected? *Journal of Evolutionary Biology* 9: 623-640.

Wickman, P.-O., and B. Karlsson. 1989. Abdomen size, body size and the reproductive effort of insects. *Oikos* 56: 209-214.

Wikelski, M. and F. Trillmich. 1997. Bod size and sexual size dimorphism in marine iguanas as a result of opposing natural and sexual selection: an island comparison. *Evolution* 51: 922-936.

Williams, G. C. 1966. *Adaptation and Natural Selection*. Princeton University Press, Princeton, NJ.

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

Zeng, Z.-B. 1988. Long-term correlated response, interpopulation covariation, and interspecific allometry. *Evolution* 42: 363-374.

Wikelski, M., and F. Trillmich. 1997. Body size and sexual size dimorphism in marine iguanas fluctuate as a result of opposing natural and sexual selection: an island

comparison. Evolution 51: 922-936.

Zeng, Z.-B., and W. G. Hill. 1986. The selection limit due to the conflict between truncation and stabilizing selection with mutation. *Genetics* 114: 1313-1328.

Zwann, B. J., R. Bijlsma, and R. F. Hoekstra. 1992. On the developmental theory of ageing. II. The effect of developmental temperature on longevity in relation to adult body size in *D. melanogaster*. *Heredity* 68: 123-130.